Introduction to Optimal Transport

Matthew Thorpe

Centre for Mathematical Sciences
University of Cambridge
Email: m.thorpe@maths.cam.ac.uk

Lent 2019
Current Version: Friday 10th May, 2019

Foreword

These notes have been written to supplement my lectures given at the University of Cambridge in the Lent terms of 2017-2018 and 2018-2019. The purpose of the lectures is to provide an introduction to optimal transport. Optimal transport dates back to Gaspard Monge in 1781 [16], with significant advancements by Leonid Kantorovich in 1942 [12] and Yann Brenier in 1987 [5]. The latter in particular lead to connections with partial differential equations, fluid mechanics, geometry, probability theory and functional analysis. Currently optimal transport enjoys applications in image retrieval, signal and image representation, inverse problems, cancer detection, texture and colour modelling, shape and image registration, and machine learning, to name a few. The purpose of this course is to introduce the basic theory that surrounds optimal transport, rather than focus on any specific application.

I can recommend the following references. My lectures and notes are mainly based on *Topics in Optimal Transportation* [23]. Two other accessible introductions are *Optimal Transport: Old and New* [24] (available for free online) and *Optimal Transport for Applied Mathemacians* [20] (also freely available online). For a more technical treatment of optimal transport I refer to *Gradient Flows in Metric Spaces and in the Space of Probability Measures* [2]. For numerical aspects I refer to *Computational Optimal Transport* [19] and gradient flows I refer to *Lecture Notes on Gradient Flows and Optimal Transport* [7]. For a short review of applications in optimal transport see the article *Optimal Mass Transport for Signal Processing and Machine Learning* [13].

Please let me know of any mistakes in the text. I will also be updating the notes as the course progresses.

Some Notation:

1. $C_b^0(Z)$ is the space of all continuous and bounded functions on Z.
2. A sequence of probability measures $\pi_n \in \mathcal{P}(Z)$ converges weak* to π, and we write $\pi_n \stackrel{*}{\rightharpoonup} \pi$, if for any $f \in C_b^0(Z)$ we have $\int_Z f \, \mathrm{d}\pi_n \to \int_Z f \, \mathrm{d}\pi$.
3. A *Polish space* is a separable completely metrizable topological space (i.e. a complete metric space with a countable dense subset).
4. $\mathcal{P}(Z)$ is the set of probability measures on Z, i.e. the subset of $\mathcal{M}_+(Z)$ with unit mass.
5. $\mathcal{M}_+(Z)$ is the set of positive Borel measures on Z.

6. $P^X : X \times Y \to X$ is the projection onto X, i.e. $P(x,y) = x$, similarly $P^Y : X \times Y \to Y$ is given by $P^Y(x,y) = y$.

7. A function $\Theta : E \to \mathbb{R} \cup \{+\infty\}$ is *convex* if for all $(z_1, z_2, t) \in E \times E \times [0,1]$ we have $\Theta(tz_1 + (1-t)z_2) \leq t\Theta(z_1) + (1-t)\Theta(z_2)$.

8. If E is a normed vector space then E^* is its dual space, i.e. the space of all bounded and linear functions on E.

9. For a set A in a topological space Z the *interior* of A, which we denote by $\mathrm{int}(A)$, is the set of points $a \in A$ such that there exists an open set $\mathcal{O}$ with the property $a \in \mathcal{O} \subseteq A$.

10. All vector spaces are assumed to be over $\mathbb{R}$.

11. The *closure* of a set A in a topological space Z, which we denote by $\overline{A}$, is the set of all points $a \in Z$ such that for any open set $\mathcal{O}$ with $a \in \mathcal{O}$ we have $\mathcal{O} \cap A \neq \emptyset$.

12. The *graph* of a function $\varphi : X \to \mathbb{R}$ which we denote by $\mathrm{Gra}(\varphi)$, is the set $\{(x,y) : x \in X, y = \varphi(x)\}$.

13. A function $\varphi : X \to \mathbb{R} \cup \{\pm\infty\}$ is *proper* if it is nowhere equal to $-\infty$ and there exists $x_0 \in X$ such that $\varphi(x_0) < +\infty$.

14. The *domain* of a proper function $\varphi : X \to \mathbb{R}$ is the set $\mathrm{Dom}(\varphi) = \{x \in X : \varphi(x) < +\infty\}$.

15. The k^{th} *moment* of $\mu \in \mathcal{P}(X)$ is defined as $\int_X |x|^k \, \mathrm{d}\mu(x)$.

16. The *support* of a probability measure $\mu \in \mathcal{P}(X)$ is the smallest closed set A such that $\mu(A) = 1$, we denote the support of a measure $\mathrm{supp}(\mu)$.

17. $\mathcal{L}$ is the Lebesgue measure on $\mathbb{R}^d$ (the dimension d should be clear by context).

18. We write $\mu\lfloor_A$ for the measure μ restricted to A, i.e. $\mu\lfloor_A(B) = \mu(A \cap B)$ for all measurable B.

19. Given a probability measure μ we say a property holds μ-almost surely if it holds on a set of probability one. If μ is the Lebesgue measure we will just say that it holds almost surely.

I would like to thank David Bourne for his input to the course and in particular for sketching out Chapter 5 and providing many of the exercises. I would also like to thank Olly Crook who (unknowingly) helped in the preparation of Chapter 9. In and earlier version of these noted mistakes and typos were found by Bobby He.

Contents

Chapter 1

Some Background on Measure Theory

We start by recalling the definition of a σ-algebra. A *σ-algebra*, Σ, on a space X is a collection of subsets of X with the following properties:

1. $X \in \Sigma$;
2. (closure under complements) if $A \in \Sigma$ then $A^c = X \setminus A \in \Sigma$;
3. (closure under countable unions) if $\{A_i\}_{i=1}^\infty \subset \Sigma$ then $\cup_{i=1}^\infty A_i \in \Sigma$.

A measure μ is a function from Σ to $\mathbb{R} \cup \{+\infty\}$ satisfying the following properties:

1. (non-negativity) for all $A \in \Sigma$, $\mu(A) \geq 0$;
2. (null-empty set) $\mu(\emptyset) = 0$;
3. (countable additivity) for any $\{A_i\}_{i=1}^\infty \subset \Sigma$ where $\{A_i\}_{i=1}^\infty$ are pairwise disjoint, $\mu(\cup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$.

A measure μ is a *probability measure* if $\mu(X) = 1$. We denote the set of probability measures on X by $\mathcal{P}(X)$.

The *Borel σ-algebra* on a topological space is the smallest σ-algebra that contains all the open sets in X. A *Borel measure* is any measure $\mu : \mathcal{B}(X) \to [0, +\infty]$. Unless otherwise stated it will always be assumed that the measures are Borel measures; in particular $\mathcal{P}(X)$ is the set of Borel probability measures over X. A Borel measure μ is said to be *inner regular* if $\mu(A)$ can be written as the supremum over all compact sets $K \subseteq A$, i.e.

$$\mu(A) = \sup \{\mu(K) \, : \, K \subseteq A \text{ and } K \text{ is compact}\},$$

and *outer regular* if $\mu(A)$ can be written as the infimum over all open sets $O \supseteq A$, i.e.

$$\mu(A) = \inf \{\mu(O) \, : \, O \supseteq A \text{ and } O \text{ is open}\}.$$

A measure μ is called *locally finite* if every point of X has a neighborhood U for which $\mu(U)$ is finite. A *Radon measure* is a inner and outer regular locally finite measure.

Throughout these notes the underlying space X will be a *Polish space*, that is a complete and separable metric space. An important consequence of working in Polish spaces is that all Borel probability measures are Radon measures.

The weak* topology on $\mathcal{P}(X)$ is defined as follows.[1]

Definition 1.1. *Let X be a metric space and $C_b^0(X)$ the space of continuous and bounded functions on X. Let $\{\mu_n\}_{n=1}^\infty \subset \mathcal{P}(X)$ and $\mu \in \mathcal{P}(X)$. We say μ_n converges weak* to μ, and we write $\mu_n \xrightharpoonup{*} \mu$, if*

$$\int_X f(x) \, \mathrm{d}\mu_n(x) \to \int_X f(x) \, \mathrm{d}\mu(x)$$

as $n \to \infty$ for all $f \in C_b^0(X)$.

The Portmanteau theorem provides equivalent definitions of the weak* convergence.

Theorem 1.2. Portmanteau Theorem. *Let X be a metric space, $\{\mu_n\}_{n=1}^\infty \subset \mathcal{P}(X)$ and $\mu \in \mathcal{P}(X)$. The following are equivalent:*

$\lim_{n\to\infty} \int_X f(x) \, \mathrm{d}\mu_n(x) = \int_X f(x) \, \mathrm{d}\mu(x)$ *for all continuous and bounded functions* f;

$\lim_{n\to\infty} \int_X f(x) \, \mathrm{d}\mu_n(x) = \int_X f(x) \, \mathrm{d}\mu(x)$ *for all Lipschitz and bounded functions* f;

$\limsup_{n\to\infty} \int_X f(x) \, \mathrm{d}\mu_n(x) \leq \int_X f(x) \, \mathrm{d}\mu(x)$ *for all usc functions f bounded from above*;

$\liminf_{n\to\infty} \int_X f(x) \, \mathrm{d}\mu_n(x) \geq \int_X f(x) \, \mathrm{d}\mu(x)$ *for all lsc functions f bounded from below*;

$\limsup_{n\to\infty} \mu_n(C) \leq \mu(C)$ *for all closed sets* C;

$\liminf_{n\to\infty} \mu_n(O) \geq \mu(O)$ *for all open sets* O;

$\lim_{n\to\infty} \mu_n(A) \geq \mu(A)$ *for all continuity sets A of μ.*

where lsc stands for lower semi-continuous, usc stands for upper semi-continuous and continuity sets of μ are sets A satisfying $\mu(\partial A) = 0$.

The following theorem allows us to characterise compactness in the weak* convergence; this is similar to the "closed plus bounded is equivalent to compact" in Euclidean spaces however boundedness is replaced by the notion of tightness. A set of probability measures $\mathcal{K} \subset \mathcal{P}(X)$ is said to be *tight* if for all $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subseteq X$ such that $\mu(X \setminus K_\varepsilon) \leq \varepsilon$ for all $\mu \in \mathcal{K}$. If $\mathcal{K} = \{\mu\}$ and μ is inner regular then $\mathcal{K}$ is tight (in this case we will often just say that μ is tight).

Theorem 1.3. Prokhorov's Theorem. *Let X be a Polish space. A set $\mathcal{K} \subset \mathcal{P}(X)$ is tight if and only if the closure of $\mathcal{K}$ is sequentially compact in the weak* topology.*

[1] Sometimes this is called the weak topology, however in functional analysis it is more usual to think of $\mathcal{P}(X)$ as the dual space of $C_b^0(X)$ and therefore this would be called weak* rather than weak.

Given a measure $\mu \in \mathcal{P}(X)$ and a map $T : X \to Y$ we can define the pushforward of μ by T as follows.

Definition 1.4. *Let $\mu \in \mathcal{P}(X)$ and $T : X \to Y$ be measurable, the pushforward of μ by T, which we denote $T_\# \mu$, is the measure ν defined by*

$$\nu(B) = \mu(T^{-1}(B)) \qquad \textit{for all measurable sets } B. \tag{1.1}$$

To visualise the pushforward measure see Figure 1.1. For greater generality we work with the inverse of T rather than T itself; the inverse is treated in the general set valued sense, i.e. $x \in T^{-1}(y)$ if $T(x) = y$, if the function T is injective then we can equivalently say that $\nu(T(A)) = \mu(A)$ for all μ-measurable A. What Figure 1.1 shows is that for any ν-measurable B, and $A = \{x \in X : T(x) \in B\}$ that $\mu(A) = \nu(B)$. For future use we state some properties of the pushforward measure.

Proposition 1.5. *Let $\mu \in \mathcal{P}(X)$, $T : X \to Y$, $S : Y \to Z$ and $f \in L^1(Y)$. Then*

1. **change of variables formula**

$$\int_Y f(y)\,\mathrm{d}(T_\#\mu)(y) = \int_X f(T(x))\,\mathrm{d}\mu(x); \tag{1.2}$$

2. **composition of maps**

$$(S \circ T)_\# \mu = S_\#(T_\# \mu).$$

Proof. Recall that, for non-negative $f : Y \to \mathbb{R}$

$$\int_Y f(y)\,\mathrm{d}(T_\#\mu)(y) := \sup\left\{ \int_Y s(y)\,\mathrm{d}(T_\#\mu)(y) : 0 \le s \le f \text{ and } s \text{ is simple} \right\}.$$

Now if $s(y) = \sum_{i=1}^N a_i \delta_{U_i}(y)$ where $a_i = s(y)$ for any $y \in U_i$ then

$$\int_Y s(y)\,\mathrm{d}(T_\#\mu)(y) = \sum_{i=1}^N a_i T_\#\mu(U_i) = \sum_{i=1}^N a_i \mu(V_i) = \int_X r(x)\,\mathrm{d}\mu(x)$$

for $V_i = T^{-1}(U_i)$ and $r = \sum_{i=1}^N a_i \delta_{V_i}$. For $x \in V_i$ we have $T(x) \in U_i$ and therefore $r(x) = a_i = s(T(x)) \le f(T(x))$. From this it is not hard to see that

$$\sup_{0 \le s \le f} \int_Y s(y)\,\mathrm{d}(T_\#\mu)(y) = \sup_{0 \le r \le f \circ T} \int_X r(x)\,\mathrm{d}\mu(x)$$

where both supremums are taken over simple functions. Hence (1.2) holds for non-negative functions. By treating signed functions as $f = f^+ - f^-$ we can prove the proposition for $f \in L^1(Y)$.

For the second statement let $A \subset Z$ and observe that $T^{-1}(S^{-1}(A)) = (S \circ T)^{-1}(A)$. Then

$$S_\#(T_\#\mu)(A) = T_\#\mu(S^{-1}(A)) = \mu(T^{-1}(S^{-1}(A))) = \mu((S \circ T)^{-1}(A)) = (S \circ T)_\#\mu(A).$$

Hence $S_\#(T_\#\mu) = (S \circ T)_\#\mu$. $\square$

Given two measures μ and ν the existence of a map T such that $T_\#\mu = \nu$ is not only non-trivial, but it may also be false. For example, consider two discrete measures $\mu = \delta_{x_1}$, $\nu = \frac{1}{2}\delta_{y_1} + \frac{1}{2}\delta_{y_2}$ where $y_1 \neq y_2$. Then $\nu(\{y_1\}) = \frac{1}{2}$ but $\mu(T^{-1}(y_1)) \in \{0, 1\}$ depending on whether $x_1 \in T^{-1}(y_1)$. Hence no such map exists.

There are three important cases where maps exist:

1. the uniform discrete case when $\mu = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\nu = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$;

2. the absolutely continuous case when $\mathrm{d}\mu(x) = f(x)\,\mathrm{d}x$ and $\mathrm{d}\nu(y) = g(y)\,\mathrm{d}y$;

3. the semi-discrete case when μ is absolutely continuous, i.e. $\mathrm{d}\mu(x) = f(x)\,\mathrm{d}x$, and ν is discrete, i.e. $\nu = \sum_{j=1}^n m_j \delta_{y_j}$.

It is important that in the uniform discrete case that μ and ν are supported on the same number of points; the supports do not have to be the same but they do have to be of the same size. We will revisit the each case later (the discrete case Chapter 3.2, the semi-discrete case in Chapter 5 and the absolutely continuous case in Chapter 6).

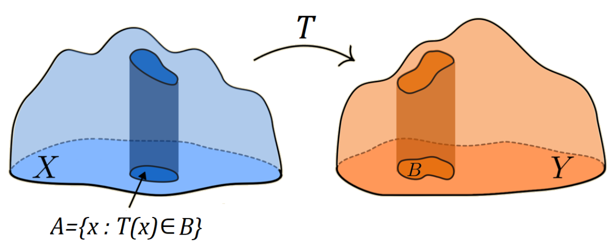

Figure 1.1: The pushforward measure, μ acts on the space X, ν acts on the space Y. Figure modified from Figure 1 in [13].

Our final piece of background knowledge will be to recall the disintegration of measures.

Theorem 1.6. Disintegration of Measures. *Let $\mathbf{X}, Z$ be Polish spaces and $P : \mathbf{X} \to Z$ a measurable map. Let $\pi \in \mathcal{P}(\mathbf{X})$ and define $\omega = P_\#\pi \in \mathcal{P}(Z)$. Then there exists a ω-almost everywhere uniquely determined family of probability measures $\{\pi(\cdot|z)\}_{z\in Z}$ such that $\pi(\cdot|z) \in \mathcal{P}(P^{-1}(z))$ and*

$$\int_{\mathbf{X}} f(\mathbf{x})\,\mathrm{d}\pi(\mathbf{x}) = \int_Z \int_{P^{-1}(z)} f(\mathbf{x})\,\mathrm{d}\pi(\mathbf{x}|z)\,\mathrm{d}\omega(z)$$

for all measurable $f : \mathbf{X} \to [0, +\infty]$.

In terms of conditional probability we can understand $\pi(A|z)$ as the conditional probability of A given the event z.

Our usual application of the above theorem will be to $\mathbf{X} = X \times Y$ and $\pi \in \Pi(\mu, \nu)$ (the set of couplings between μ and ν see Chapter 2.2). In this case we define $P(x, y) = y$ and note that

$\nu = P_{\#}\pi$. Since $P^{-1}(y) = X \times \{y\}$ we can with an abuse of notation write $\pi(\cdot|y) \in \mathcal{P}(X)$. In particular, we can write

$$\int_{X \times Y} f(x,y)\,\mathrm{d}\pi(x,y) = \int_Y \int_X f(x,y)\,\mathrm{d}\pi(x|y)\,\mathrm{d}\nu(y)$$

for all measurable $f : X \times Y \to [0, +\infty]$.

Chapter 2

Formulation of Optimal Transport

There are two ways to formulate the optimal transport problem: the Monge and Kantorovich formulations. We explain both these formulations in this chapter. Historically the Monge formulation comes before Kantorovich which is why we introduce Monge first. The Kantorovich formulation can be seen as a generalisation of Monge. Both formulations have their advantages and disadvantages. My experience is that Monge is more useful in applications, whilst Kantorovich is more useful theoretically. In a later chapter (see Chapter 6) we will show sufficient conditions for the two problems to be considered equivalent. After introducing both formulations we give an existence result for the Kantorovich problem; existence results for Monge are considerably more difficult. We look at special cases of the Monge and Kantorovich problems in the next chapter, a more general treatment is given in Chapters 4 and 6.

2.1 The Monge Formulation

Optimal transport gives a framework for comparing measures μ and ν in a Lagrangian framework. Essentially one pays a cost for transporting one measure to another. To illustrate this consider the first measure μ as a pile of sand and the second measure ν as a hole we wish to fill up. We assume that both measures are probability measures on spaces X and Y respectively. Let $c : X \times Y \to [0, +\infty]$ be a cost function where $c(x, y)$ measures the cost of transporting one unit of mass from $x \in X$ to $y \in Y$. The optimal transport problem is how to transport μ to ν whilst minimizing the cost c.[1] We call $T : X \to Y$ a transport map if it satisfies the following definition.

Definition 2.1. *We say that $T : X \to Y$ transports $\mu \in \mathcal{P}(X)$ to $\nu \in \mathcal{P}(Y)$, and we call T a* transport map, *if $\nu = T_{\#}\mu$.*

The Monge formulation of optimal transport is to find the transport map T between μ and ν that moves as little mass as possible.

[1]Some time ago I either read or was told that the original motivation for Monge was how to design defences for Napoleon. In this case the pile of sand is a wall and the hole a moat. Obviously one wishes to to make the wall using the dirt dug out to form the moat. In this context the optimal transport problem is how to transport the dirt from the moat to the wall.

Definition 2.2. ***Monge's Optimal Transport Problem:*** *given $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$,*

$$\text{minimise } \mathbb{M}(T) = \int_X c(x, T(x)) \, \mathrm{d}\mu(x)$$

over measurable maps $T : X \to Y$ subject to $\nu = T_{\#}\mu$.

Monge originally considered the problem with L^1 cost, i.e. $c(x, y) = |x - y|$. This problem is significantly harder than with L^2 cost, i.e. $c(x, y) = |x - y|^2$. In fact the first correct proof with L^1 cost dates back only a few years to 1999 (see Evans and Gangbo [9]) and required stronger assumptions than the L^2 cost, Sudakov thought to have proven the result in 1979 [22] however this was found to contain a mistake which was later fixed by Ambrosio and Pratelli [1, 3].

In general Monge's problem is difficult due to the non-linearity in the constraint (1.1). If we assume that μ and ν are absolutely continuous with respect to the Lebegue measure on $\mathbb{R}^d$, i.e. $\mathrm{d}\mu(x) = f(x) \, \mathrm{d}x$ and $\mathrm{d}\nu(y) = g(y) \, \mathrm{d}y$, and assume T is a C^1 diffeomorphism (T is bijective and T, T^{-1} are differentiable) then one can show that (1.1) is equivalent to

$$f(x) = g(T(x)) \left| \det(\nabla T(x)) \right|.$$

The above constraint is highly non-linear and difficult to handle with standard techniques from the calculus of variations.

2.2 The Kantorovich Formulation

Observe that in the Monge formulation mass is mapped $x \mapsto T(x)$. In particular, this means that mass is not split. In the discrete case this causes difficulties concerning the existence of maps T such that $T_{\#}\mu = \nu$, see the example $\mu = \delta_{x_1}$, $\nu = \frac{1}{2}\delta_{y_1} + \frac{1}{2}\delta_{y_2}$ in the previous section. Observe that if we allow mass to be split, i.e. half of the mass from x_1 goes to y_1 and half the mass goes to y_2, then we have a natural relaxation. This is in effect what the Kantorovich formulation does. To formalise this we consider a measure $\pi \in \mathcal{P}(X \times Y)$ and think of $\mathrm{d}\pi(x, y)$ as the amount of mass transferred from x to y; this way mass can be transferred from x to multiple locations. Of course the total amount of mass removed from any measurable set $A \subset X$ has to equal to $\mu(A)$, and the total amount of mass transferred to any measurable set $B \subset Y$ has to be equal to $\nu(B)$. In particular, we have the constraints:

$$\pi(A \times Y) = \mu(A), \qquad \pi(X \times B) = \nu(B) \quad \text{for all measurable sets } A \subseteq X, \ B \subseteq Y.$$

We say that any π which satisfies the above has first marginal μ and second marginal ν, we denote the set of such π by $\Pi(\mu, \nu)$. We will call $\Pi(\mu, \nu)$ the set of *transport plans* between μ and ν. Note that $\Pi(\mu, \nu)$ is never non-empty (in comparison with the set of transport plans) since $\mu \otimes \nu \in \Pi(\mu, \nu)$ which is the trivial transport plan which transports every grain of sand at x to y proportional to $\nu(y)$. We can now define Kantorovich's formulation of optimal transport.

Definition 2.3. ***Kantorovich's Optimal Transport Problem:*** *given $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$,*

$$\textit{minimise } \mathbb{K}(\pi) = \int_{X\times Y} c(x,y)\,\mathrm{d}\pi(x,y)$$

over $\pi \in \Pi(\mu,\nu)$.

By the example with discrete measures, where we showed there did not exist any transport maps, we know that Kantorovich's and Monge's optimal transport problems do not always coincide. However, let us assume that there exists a transport map $T^\dagger : X \to Y$ that is optimal for Monge, then if we define $\mathrm{d}\pi(x,y) = \mathrm{d}\mu(x)\delta_{y=T^\dagger(x)}$ a quick calculation shows that $\pi \in \Pi(\mu,\nu)$:

$$\pi(A\times Y) = \int_A \delta_{T^\dagger(x)\in Y}\,\mathrm{d}\mu(x) = \mu(A)$$
$$\pi(X\times B) = \int_X \delta_{T^\dagger(x)\in B}\,\mathrm{d}\mu(x) = \mu((T^\dagger)^{-1}(B)) = T^\dagger_{\#}\mu(B) = \nu(B).$$

Since,

$$\int_{X\times Y} c(x,y)\,\mathrm{d}\pi(x,y) = \int_X\int_Y c(x,y)\delta_{y=T^\dagger(x)}\,\mathrm{d}y\,\mathrm{d}\mu(x) = \int_X c(x,T^\dagger(x))\,\mathrm{d}\mu(x)$$

it follows that

$$\inf \mathbb{K}(\pi) \leq \inf \mathbb{M}(T). \tag{2.1}$$

In fact one does not need minimisers of Monge's problem to exist. If $\mathbb{M}(T^\dagger) \leq \min \mathbb{M}(T) + \varepsilon$ for some $\varepsilon > 0$ then $\inf \mathbb{K}(\pi) \leq \inf \mathbb{M}(T) + \varepsilon$. Since $\varepsilon > 0$ was arbitrary then (2.1) holds.

When the optimal plan $\pi^\dagger$ can be written in the form $\mathrm{d}\pi^\dagger(x,y) = \mathrm{d}\mu(x)\delta_{y=T(x)}$ it follows that T is an optimal transport map and $\inf \mathbb{K}(\pi) = \inf \mathbb{M}(T)$. Conditions sufficient for such a condition will be explored in Chapter 6. For now we just say that if $c(x,y) = |x-y|^2$, μ,ν both have finite second moments, and μ does not give mass to small sets[2] then there exists an optimal plan $\pi^\dagger$ which can be written as $\mathrm{d}\pi^\dagger(x,y) = \mathrm{d}\mu(x)\delta_{y=T^\dagger(x)}$ where $T^\dagger$ is an optimal map.

Let us observe the advantages of both Monge and Kantorovich formulation. Transport maps give a natural method of interpolation between two measures, in particular we can define

$$\mu_t = ((1-t)\mathrm{Id} + tT^\dagger)_{\#}\mu$$

then μ_t interpolates between μ and ν. In fact this line of reasoning will lead us directly to geodesics that we consider in greater detail in Chapter 7. In Figure 2.1 we compare the optimal transport interpolation with the Euclidean interpolation defined by $\mu_t^E = (1-t)\mu + t\nu$. In many applications the Lagrangian nature of optimal transport will be more realistic than Euclidean formulations.

Notice that the Kantorovich problem is convex (the constraints are convex and one usually has that the cost function $c(x,y) = d(x-y)$ where d is convex). In particular let us consider

[2] $\mu \in \mathcal{P}(\mathbb{R}^d)$ does not give mass to small sets if for all sets A of Hausdorff dimension at most $d-1$ that $\mu(A) = 0$.

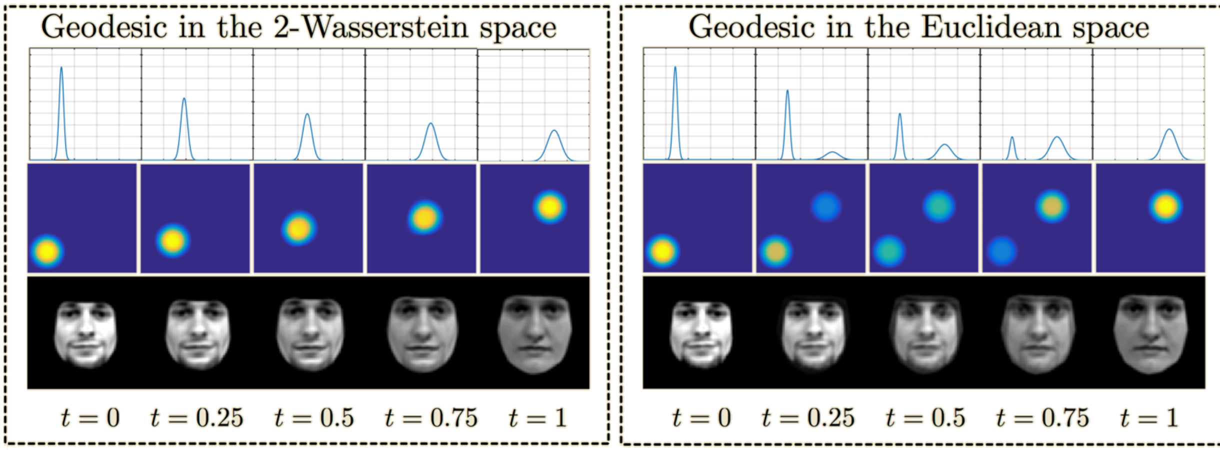

Figure 2.1: Interpolation in the optimal transport framework (left) and Euclidean space (right), figure modified from Figure 2 in [13].

the Kantorovich problem between discrete measures $\mu = \sum_{i=1}^m \alpha_i \delta_{x_i}$, $\nu = \sum_{j=1}^n \beta_j \delta_{y_j}$ where $\sum_{i=1}^m \alpha_i = 1 = \sum_{j=1}^n \beta_j$, $\alpha_i \geq 0, \beta_j \geq 0$. Let $c_{ij} = c(x_i, y_j)$ and $\pi_{ij} = \pi(x_i, x_j)$. Then the Kantorovich problem is to solve

$$\text{minimise } \sum_{i=1}^m \sum_{j=1}^n c_{ij}\pi_{ij} \text{ over } \pi \text{ subject to } \pi_{ij} \geq 0, \sum_{i=1}^m \pi_{ij} = \beta_j, \sum_{j=1}^n \pi_{ij} = \alpha_i.$$

This is a linear programme! In fact Kantorovich is considered as the inventor of linear programming. Not only does this provide a method for solving optimal transport problems (either through off the shelf linear programming algorithms, or through more recent advances such as entropy regularised approaches see [6]) but the dual formulation:

$$\inf_{\pi \geq 0, \mathcal{C}^\top \pi = (\mu^\top, \nu^\top)^\top} c \cdot \pi = \sup_{\mathcal{C}(\varphi^\top, \phi^\top)^\top \leq c} (\mu \cdot \varphi + \nu \cdot \phi)$$

is an important stepping stone in establishing important properties such as the characterisation of optimal transport maps and plans. We study the dual formulation in the Chapter 4. In the next section we prove the existence of transport plans under fairly general conditions.

2.3 Existence of Transport Plans

Section references: Proposition 2.4 is taken from [23, Proposition 2.1].

We complete this chapter by proving the existence of a minimizer to Kantorovich's optimal transport problem. For the proof we use the direct method from the calculus of variations. Approximately the direct method is *compactness* plus *lower semi-continuity*. More precisely if we are considering a variational problem $\inf_{v \in V} F(v)$ then we first show that the set V is compact (or at least a set which contains the minimizer is compact). Then, let v_n be a minimising sequence, i.e.

$F(v_n) \to \inf F$. Upon extracting a subsequence we can assume that $v_n \to v^\dagger \in V$. This gives a candidate minimizer. If we can show that F is lower semi-continuous then $\lim_{n\to\infty} F(v_n) \geq F(v^\dagger)$ and hence $v^\dagger$ is a minimiser.

Proposition 2.4. *Let $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ where X, Y are Polish spaces, and assume $c : X \times Y \to [0,\infty)$ is lower semi-continuous. Then there exists $\pi^\dagger \in \Pi(\mu,\nu)$ that minimises $\mathbb{K}$ (defined in Definition 2.3) over all $\pi \in \Pi(\mu,\nu)$.*

Proof. Note that $\Pi(\mu,\nu)$ is non-empty. Let us see that $\Pi(\mu,\nu)$ is compact in the weak* topology. Let $\delta > 0$ and choose compact sets $K \subset X$, $L \subset Y$ such that

$$\mu(X \setminus K) \leq \delta, \qquad \nu(Y \setminus L) \leq \delta.$$

(Existence of sets follows directly since by definition Radon measures are inner regular.) If $(x,y) \in (X \times Y) \setminus (K \times L)$ then either $x \notin K$ or $y \notin L$, hence $(x,y) \in X \times (Y \setminus L)$ or $(x,y) \in (X \setminus K) \times Y$. So, for any $\pi \in \Pi(\mu,\nu)$

$$\begin{aligned} \pi((X \times Y) \setminus (K \times L)) &\leq \pi(X \times (Y \setminus L)) + \pi((X \setminus K) \times Y) \\ &= \nu(Y \setminus L) + \mu(X \setminus K) \\ &\leq 2\delta. \end{aligned}$$

Hence, $\Pi(\mu,\nu)$ is tight. By Prokhorov's theorem the closure of $\Pi(\mu,\nu)$ is sequentially compact in the topology of weak* convergence.

To check that $\Pi(\mu,\nu)$ is (weak*) closed let $\pi_n \in \Pi(\mu,\nu)$ be a sequence weakly* converging to $\pi \in \mathcal{M}(X \times Y)$, i.e.

$$\int_{X\times Y} f(x,y)\,\mathrm{d}\pi_n(x,y) \to \int_{X\times Y} f(x,y)\,\mathrm{d}\pi(x,y) \quad \forall f \in C_b^0(X \times Y).$$

We choose $f(x,y) = \tilde{f}(x)$, where $\tilde{f}$ is continuous and bounded. We have,

$$\int_X \tilde{f}(x)\,\mathrm{d}\mu(x) \to \int_{X\times Y} \tilde{f}(x)\,\mathrm{d}\pi(x,y) = \int_X \tilde{f}(x)\,\mathrm{d}P^X_\#\pi(x)$$

where $P^X(x,y) = x$ is the projection onto X (so $P^X_\#\pi$ is the X marginal). Since this is true for all $\tilde{f} \in C_b^0(X)$ it follows that $P^X_\#\pi = \mu$. Similarly, $P^Y_\#\pi = \nu$. Hence, $\pi \in \Pi(\mu,\nu)$ and $\Pi(\mu,\nu)$ is weakly* closed.

Let $\pi_n \in \Pi(\mu,\nu)$ be a minimising sequence, i.e. $\mathbb{K}(\pi_n) \to \inf_{\pi\in\Pi(\mu,\nu)} \mathbb{K}(\pi)$. Since $\Pi(\mu,\nu)$ is compact we can assume that $\pi_n \overset{*}{\rightharpoonup} \pi^\dagger \in \Pi(\mu,\nu)$. Our candidate for a minimiser is $\pi^\dagger$. Note that c is lower semi-continuous and bounded from below. Then,

$$\inf_{\pi\in\Pi(\mu,\nu)} \mathbb{K}(\pi) = \lim_{n\to\infty} \int_{X\times Y} c(x,y)\,\mathrm{d}\pi_n(x,y) \geq \int_{X\times Y} c(x,y)\,\mathrm{d}\pi^\dagger(x,y)$$

where we use the Portmanteau theorem which provides equivalent characterisations of weak* convergence. Hence $\pi^\dagger$ is a minimiser. □

Chapter 3

Special Cases

In this section we look at some special cases where we can prove existence and characterisation of optimal transport maps and plans. Generalising these results requires some work and in particular a duality theorem. On the other hand the results in this chapter require less background and are somehow "lower hanging fruit". Chapters 4 and 6 are essentially the results of this chapter generalised to more abstract settings. The two special cases we consider here are when measures μ, ν are on the real line, and when measures μ, ν are discrete. We start with the real line.

3.1 Optimal Transport in One Dimension

Section references: a version of Theorem 3.1 can be found in [23, Theorem 2.18] and [20, Theorem 2.9 and Proposition 2.17], versions of Corollary 3.2 can be found in [23, Remark 2.19] and [20, Lemma 2.8 and Proposition 2.17], Proposition 3.3 can be found in [10, Theorem 2.3].

Let us consider two measures $\mu, \nu \in \mathcal{P}(\mathbb{R})$ with cumulative distribution functions F and G respectively. We recall that

$$F(x) = \int_{-\infty}^{x} \mathrm{d}\mu = \mu((-\infty, x])$$

and F is right-continuous, non-decreasing, $F(-\infty) = 0$ and $F(+\infty) = 1$. We define the generalised inverse of F on $[0, 1]$ by

$$F^{-1}(t) = \inf \{x \in \mathbb{R} \,:\, F(x) \geq t\} .$$

If F is invertible then $F^{-1}(F(x)) = x$ and $F(F^{-1}(t)) = t$. The main result of this section is the following theorem.

Theorem 3.1. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ (where $\mathbb{R}$ is equipped with the Euclidean metric) with cumulative distributions F and G respectively. Assume $c(x, y) = d(x - y)$ where $d : \mathbb{R} \to [0, +\infty)$ is convex and continuous. Let $\pi^\dagger$ be the measure on $\mathbb{R}^2$ with cumulative distribution function $H(x, y) = \min\{F(x), G(y)\}$. Then $\pi^\dagger \in \Pi(\mu, \nu)$ and furthermore $\pi^\dagger$ is optimal for Kantorovich's optimal transport problem with cost function c. Moreover the optimal transport cost is*

$$\min_{\pi \in \Pi(\mu,\nu)} \mathbb{K}(\pi) = \int_0^1 d\left(F^{-1}(t) - G^{-1}(t)\right) \, \mathrm{d}t.$$

Before proving the theorem we state a corollary.

Corollary 3.2. *Under the assumptions of Theorem 3.1 the following holds.*

1. *If $c(x,y) = |x-y|$ then the optimal transport distance is the L^1 distance between cumulative distribution functions, i.e.*

$$\inf_{\pi\in\Pi(\mu,\nu)} \mathbb{K}(\pi) = \int_{\mathbb{R}} |F(x) - G(x)| \, \mathrm{d}x.$$

2. *If μ does not give mass to atoms then $\min_{\pi\in\Pi(\mu,\nu)} \mathbb{K}(\pi) = \min_{T\,:\,T_{\#}\mu=\nu} \mathbb{M}(T)$ and furthermore $T^\dagger = G^{-1}\circ F$ is a minimiser to Monge's optimal transport problem, i.e. $T^\dagger_{\#}\mu = \nu$ and*

$$\inf_{T\,:\,T_{\#}\mu=\nu} \mathbb{M}(T) = \mathbb{M}(T^\dagger).$$

Proof. For the first part, by Theorem 3.1, it is enough to show that

$$\int_0^1 \left|F^{-1}(t) - G^{-1}(t)\right| \, \mathrm{d}t = \int_{\mathbb{R}} |F(x) - G(x)| \, \mathrm{d}x.$$

Define $\mathcal{A} \subset \mathbb{R}^2$ by

$$\mathcal{A} = \{(x,t) \,:\, \min\{F(x), G(x)\} \leq t \leq \max\{F(x), G(x)\},\, x\in\mathbb{R}\}.$$

From Figure 3.1 we notice that we can equivalently write

$$\mathcal{A} = \left\{(x,t) \,:\, \min\{F^{-1}(t), G^{-1}(t)\} \leq x \leq \max\{F^{-1}(t), G^{-1}(t)\},\, t\in[0,1]\right\}$$

up to a $\mathcal{L}$ negligible set. By Fubini's theorem

$$\mathcal{L}(\mathcal{A}) = \int_{\mathbb{R}} \int_{\min\{F(x),G(x)\}}^{\max\{F(x),G(x)\}} \mathrm{d}t \, \mathrm{d}x = \int_0^1 \int_{\min\{F^{-1}(t),G^{-1}(t)\}}^{\max\{F^{-1}(t),G^{-1}(t)\}} \mathrm{d}x \, \mathrm{d}t$$

where $\mathcal{L}$ is the Lebesgue measure. Since $\max\{a,b\} - \min\{a,b\} = |a-b|$ then

$$\begin{aligned}\int_{\mathbb{R}} \int_{\min\{F(x),G(x)\}}^{\max\{F(x),G(x)\}} \mathrm{d}t \, \mathrm{d}x &= \int_{\mathbb{R}} \min\{F(x), G(x)\} - \max\{F(x), G(x)\} \, \mathrm{d}x \\ &= \int_{\mathbb{R}} |F(x) - G(x)| \, \mathrm{d}x.\end{aligned}$$

Similarly

$$\int_0^1 \int_{\min\{F^{-1}(t),G^{-1}(t)\}}^{\max\{F^{-1}(t),G^{-1}(t)\}} \mathrm{d}x \, \mathrm{d}t = \int_0^1 \left|F^{-1}(t) - G^{-1}(t)\right| \, \mathrm{d}t.$$

This proves the first part of the corollary.

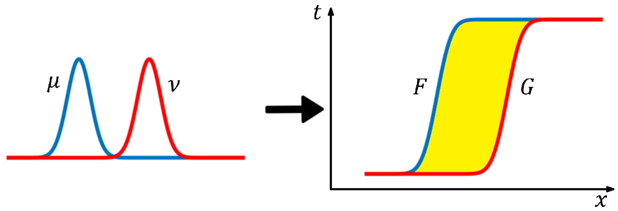

Figure 3.1: Optimal transport distance in 1D with cost $c(x,y) = |x-y|$, figure is taken from [14].

For the second part we recall by Proposition 1.5 that $T^\dagger_\# \mu = G^{-1}_\#(F_\# \mu)$. We show that (i) $G^{-1}_\# \mathcal{L}\lfloor_{[0,1]} = \nu$ and (ii) $\mathcal{L}\lfloor_{[0,1]} = F_\# \mu$. This is enough to show that $T_\# \mu = \nu$. For (i),

$$
\begin{aligned}
G^{-1}_\# \mathcal{L}\lfloor_{[0,1]}((-\infty, y]) &= \mathcal{L}\lfloor_{[0,1]}\left(\left\{t \,:\, G^{-1}(t) \leq y\right\}\right) \\
&= \mathcal{L}\lfloor_{[0,1]}(\{t \,:\, G(y) \geq t\}) \\
&= G(y) \\
&= \nu\left((-\infty, y]\right)
\end{aligned}
$$

where we used $G^{-1}(t) \leq y \Leftrightarrow G(y) \geq t$. For (ii) we note that F is continuous (as μ does not give mass to atoms). So for all $t \in (0,1)$ the set $F^{-1}([0,t])$ is closed, in particular $F^{-1}([0,t]) = (-\infty, x_t]$ for some x_t with $F(x_t) = t$. Now, for $t \in (0,1)$,

$$
\begin{aligned}
F_\# \mu\left([0,t]\right) &= \mu\left(\{x \,:\, F(x) \leq t\}\right) \\
&= \mu\left(\{x \,:\, F(x) \leq F(x_t)\}\right) \\
&= \mu\left(\{x \,:\, x \leq x_t\}\right) \\
&= F(x_t) \\
&= t.
\end{aligned}
$$

Hence $F_\# \mu = \mathcal{L}\lfloor_{[0,1]}$.

Now we show that $T^\dagger$ is optimal. By Theorem 3.1

$$
\begin{aligned}
\inf_{\pi \in \Pi(\mu,\nu)} \mathbb{K}(\pi) &= \int_0^1 d\left(F^{-1}(t) - G^{-1}(t)\right) \, \mathrm{d}t \\
&= \int_{\mathbb{R}} d\left(x - G^{-1}(F(x))\right) \, \mathrm{d}\mu(x) \qquad \text{since } F_\# \mu = \mathcal{L}\lfloor_{[0,1]} \text{ and by Proposition 1.5} \\
&= \int_{\mathbb{R}} d\left(x - T^\dagger(x)\right) \, \mathrm{d}\mu(x) \\
&\geq \inf_{T \,:\, T_\# \mu = \nu} \mathbb{M}(T).
\end{aligned}
$$

Since $\inf_{T \,:\, T_\# \mu = \nu} \mathbb{M}(T) \geq \min_{\pi \in \Pi(\mu,\nu)} \mathbb{K}(\pi)$ then the minimum of the Monge and Kantorovich optimal transport problems coincide and $T^\dagger$ is an optimal map for Monge. $\square$

Before we prove Theorem 3.1 we give some basic ideas in the proof. The key is the idea of monotonicity. We say that a set $\Gamma \subset \mathbb{R}^2$ is *monotone* (with respect to d) if for all $(x_1, y_1), (x_2, y_2) \in \Gamma$

that

$$d(x_1 - y_1) + d(x_2 - y_2) \le d(x_1 - y_2) + d(x_2 - y_1).$$

For example, if $\Gamma = \{(x, y) : f(x) = y\}$ and f is increasing, then Γ is monotone (assuming that d is increasing). The definition generalises to higher dimensions and often appears in convex analysis (for example the subdifferential of a convex function satisfies a monotonicity property). As a result, this concept can also be used to prove analogous results to Theorem 3.1 in higher dimensions. The definition should be natural for optimal transport. In particular, let Γ be the support of $\pi^\dagger$, which is a solution of Kantorovich's optimal transport problem. If $\pi^\dagger$ transports mass from x_1 to y_1 and from $x_2 > x_1$ to y_2 we expect $y_2 > y_1$, else it would have been cheaper to transport from x_1 to y_2, and from x_2 to y_1. The following proposition formalises this reasoning.

Proposition 3.3. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ (where $\mathbb{R}$ is equipped with the Euclidean metric). Assume $\pi^\dagger \in \Pi(\mu, \nu)$ is an optimal transport plan in the Kantorovich sense for cost function $c(x, y) = d(x - y)$ where d is continuous. Then for all $(x_1, y_1), (x_2, y_2) \in \mathrm{supp}(\pi^\dagger)$ we have*

$$d(x_1 - y_1) + d(x_2 - y_2) \le d(x_1 - y_2) + d(x_2 - y_1).$$

Proof. Let $\Gamma = \mathrm{supp}(\pi^\dagger)$ and $(x_1, y_1), (x_2, y_2) \in \Gamma$. Assume there exists $\eta > 0$ such that

$$d(x_1 - y_1) + d(x_2 - y_2) - d(x_1 - y_2) - d(x_2 - y_1) \ge \eta.$$

Let I_1, I_2, J_1, J_2 be closed intervals with the following properties:

1. $x_i \in I_i$, $y_i \in J_i$, $i = 1, 2$;
2. $|d(x - y) - d(x_i - y_j)| \le \varepsilon$ for $x \in I_i$, $y \in J_j$, $i, j = 1, 2$, where $\varepsilon < \frac{\eta}{4}$;
3. $I_i \times J_j$ are disjoint;
4. $\pi^\dagger(I_1 \times J_1) = \pi^\dagger(I_2 \times J_2) = \delta > 0$.

Properties 1-3 can be satisfied by choosing the intervals I_i, J_j sufficiently small. It may not be possible to satisfy property 4, however since $(x_i, y_i) \in \Gamma$ then we can find set I_i, J_j that satisfy 1-3 and $\pi^\dagger(I_1 \times J_1) > 0$, $\pi^\dagger(I_2 \times J_2) > 0$. It makes the notation in the proof easier to assume that $\pi^\dagger(I_1 \times J_1) = \pi^\dagger(I_2 \times J_2)$ however if not the proof can be adapted and we briefly describe how at the end.

The idea of the proof is to, instead of transferring mass from x_1 to y_1, and from x_2 to y_2, transfer mass from x_1 to y_2, and from x_2 to y_1. To make the argument rigorous we talk about the mass around each of x_i, y_i (hence the need for the intervals I_i, J_i).

Let

$$\tilde{\mu}_1 = P^X_{\#} \pi^\dagger \lfloor_{I_1 \times J_1}, \qquad \tilde{\mu}_2 = P^X_{\#} \pi^\dagger \lfloor_{I_2 \times J_2},$$
$$\tilde{\nu}_1 = P^Y_{\#} \pi^\dagger \lfloor_{I_1 \times J_1}, \qquad \tilde{\nu}_2 = P^Y_{\#} \pi^\dagger \lfloor_{I_2 \times J_2}.$$

And choose any $\tilde{\pi}_{12} \in \Pi(\tilde{\mu}_1, \tilde{\nu}_2)$, $\tilde{\pi}_{21} \in \Pi(\tilde{\mu}_2, \tilde{\nu}_1)$. We define $\tilde{\pi}$ to satisfy

$$\tilde{\pi}(A \times B) = \begin{cases} \pi^\dagger(A \times B) & \text{if } (A \times B) \cap (I_i \times J_j) = \emptyset \text{ for all } i, j \\ 0 & \text{if } A \times B \subseteq I_i \times J_i \text{ for some } i \\ \pi^\dagger(A \times B) + \tilde{\pi}_{12}(A \times B) & \text{if } A \times B \subseteq I_1 \times J_2 \\ \pi^\dagger(A \times B) + \tilde{\pi}_{21}(A \times B) & \text{if } A \times B \subseteq I_2 \times J_1. \end{cases}$$

For sets $(A \times B) \cap (I_i \times J_j) \neq \emptyset$ but $A \times B \not\subseteq (I_i \times J_j)$ then we define $\tilde{\pi}(A \times B)$ by

$$\tilde{\pi}(A \times B) = \tilde{\pi}((A \times B) \cap (I_i \times J_j)) + \tilde{\pi}((A \times B) \cap (I_i \times J_j)^c).$$

By construction, for $B \cap (J_1 \cup J_2) = \emptyset$,

$$\tilde{\pi}(\mathbb{R} \times B) = \pi^\dagger(\mathbb{R} \times B) = \nu(B).$$

If $B \subseteq J_1$ then

$$\begin{aligned} \tilde{\pi}(\mathbb{R} \times B) &= \tilde{\pi}((\mathbb{R} \setminus (I_1 \cup I_2)) \times B) + \tilde{\pi}(I_1 \times B) + \tilde{\pi}(I_2 \times B) \\ &= \pi^\dagger((\mathbb{R} \setminus (I_1 \cup I_2)) \times B) + 0 + \pi^\dagger(I_2 \times B) + \tilde{\pi}_{21}(I_2 \times B) \\ &= \pi^\dagger((\mathbb{R} \setminus I_1) \times B) + \pi^\dagger(I_1 \times B) \\ &= \pi^\dagger(\mathbb{R} \times B) \\ &= \nu(B) \end{aligned}$$

since $\tilde{\pi}_{21}(I_2 \times B) = \tilde{\nu}_1(B) = \pi^\dagger(I_1 \times (B \cap J_1)) = \pi^\dagger(I_1 \times B)$. Similarly for $B \subseteq J_2$. Hence we have $\tilde{\pi}(\mathbb{R} \times B) = \nu(B)$ for all measurable B. Analogously $\tilde{\pi}(A \times \mathbb{R}) = \mu(A)$ for all measurable A. Therefore $\tilde{\pi} \in \Pi(\mu, \nu)$.

Now,

$$\begin{aligned} &\int_{\mathbb{R}\times\mathbb{R}} d(x-y)\, \mathrm{d}\pi^\dagger(x,y) - \int_{\mathbb{R}\times\mathbb{R}} d(x-y)\, \mathrm{d}\tilde{\pi}(x,y) \\ &\qquad = \int_{(I_1\times J_1)\cup(I_2\times J_2)} d(x-y)\, \mathrm{d}\pi^\dagger(x,y) - \int_{I_1\times J_2} d(x-y)\, \mathrm{d}\tilde{\pi}_{12}(x,y) \\ &\qquad\qquad - \int_{I_2\times J_1} d(x-y)\, \mathrm{d}\tilde{\pi}_{21}(x,y) \\ &\qquad \geq \delta\left(d(x_1-y_1)-\varepsilon\right) + \delta\left(d(x_2-y_2)-\varepsilon\right) - \delta\left(d(x_1-y_2)+\varepsilon\right) - \delta\left(d(x_2-y_1)+\varepsilon\right) \\ &\qquad \geq \delta(\eta - 4\varepsilon) \\ &\qquad > 0 \end{aligned}$$

since $\tilde{\pi}_{12}(I_1 \times J_2) = \tilde{\mu}_1(I_1) = \pi^\dagger(I_1 \times J_1) = \delta$, and similarly $\tilde{\pi}_{21}(I_2 \times J_1) = \delta$. This contradicts the assumption that $\pi^\dagger$ is optimal, hence no such η can exist.

Finally we remark that if $\pi^\dagger(I_1 \times J_1) > \pi^\dagger(I_2 \times J_2)$ then one can adapt the constructed plan $\tilde{\pi}$ by transporting some mass with the original plan $\pi^\dagger$. In particular the new constructed plan is chosen to satisfy

$$\tilde{\pi}(A \times B) = \pi^\dagger(A \times B)\left(1 - \frac{\pi^\dagger(I_2 \times J_2)}{\pi^\dagger(I_1 \times J_1)}\right)$$

if $A \times B \subset I_1 \times J_1$, and $\tilde{\mu}_1, \tilde{\nu}_1$ are rescaled:

$$\tilde{\mu}_1 = \frac{\pi^\dagger(I_2 \times J_2)}{\pi^\dagger(I_1 \times J_1)} P^X_\# \pi^\dagger \lfloor_{I_1 \times J_1}, \qquad \tilde{\nu}_1 = \frac{\pi^\dagger(I_2 \times J_2)}{\pi^\dagger(I_1 \times J_1)} P^Y_\# \pi^\dagger \lfloor_{I_1 \times J_1}.$$

All other definitions remain unchanged. One can go through the argument above and reach the same conclusion. □

We now prove Theorem 3.1.

Proof of Theorem 3.1. Assume d is continuous and strictly convex. By Proposition 2.4 there exists $\pi^* \in \Pi(\mu, \nu)$ that is an optimal transport plan in the Kantorovich sense. We will show that $\pi^* = \pi^\dagger$ By Proposition 3.3 $\Gamma = \mathrm{supp}(\pi^*)$ is monotone, i.e.

$$d(x_1 - y_1) + d(x_2 - y_2) \leq d(x_1 - y_2) + d(x_2 - y_1)$$

for all $(x_1, y_1), (x_2, y_2) \in \Gamma$. We claim that for any x_1, x_2, y_1, y_2 satisfying the above and $x_1 < x_2$ that $y_1 \leq y_2$. Assume that $y_2 < y_1$ and let $a = x_1 - y_1$, $b = x_2 - y_2$ and $\delta = x_2 - x_1$. We know that

$$d(a) + d(b) \leq d(b - \delta) + d(a + \delta).$$

Let $t = \frac{\delta}{b-a}$, it is easy to check that $t \in (0,1)$ if and only if $y_2 < y_1$, and $b - \delta = (1-t)b + ta$, $a + \delta = tb + (1-t)a$. Then, by strict convexity of d,

$$d(b - \delta) + d(a + \delta) < (1-t)d(b) + td(a) + td(b) + (1-t)d(a) = d(b) + d(a).$$

This is a contradiction, hence $y_2 \geq y_1$.

Now we show that $\pi^\dagger = \pi^*$. More precisely we show that

$$\pi^* \left((-\infty, x], (-\infty, y] \right) = \min\{F(x), G(y)\}.$$

Pick $(x, y) \in \mathbb{R}^2$. Let $A = (-\infty, x] \times (y, +\infty)$, $B = (x, +\infty) \times (-\infty, y]$. We claim $\pi(A)$ and $\pi(B)$ cannot both be non-zero. Indeed, assume there exists $(x_1, y_1) \in \Gamma \cap A$ and pick any $(x_2, y_2) \in B$. By definition of A and B we have $x_1 \leq x < x_2$ and $y_2 \leq y < y_1$. Since $X_1 < x_2$, if $(x_2, y_2) \in \Gamma$ we would necessarily have $y_1 \leq y_2$ by the argument above, but this is a contradiction. This implies that if there exists $(x_1, y_1) \in \Gamma \cap A$ then $B \subset \Gamma^c$, hence $\pi(B) = 0$. Analogously, if there exists $(x_2, y_2) \in \Gamma \cap B$ then $\pi(A) = 0$.

It follows that

$$\begin{aligned}\pi^* \left((-\infty, x] \times (-\infty, y] \right) = \min \big\{ &\pi^* \left(((-\infty, x] \times (-\infty, y]) \cup A \right), \\ &\pi^* \left(((-\infty, x] \times (-\infty, y]) \cup B \right) \big\}.\end{aligned}$$

But,

$$\begin{aligned}\pi^* \left(((-\infty, x] \times (-\infty, y]) \cup A \right) &= \pi \left((-\infty, x] \times \mathbb{R} \right) = F(x) \\ \pi^* \left(((-\infty, x] \times (-\infty, y]) \cup B \right) &= \pi \left(\mathbb{R} \times (-\infty, y] \right) = G(y).\end{aligned}$$

Hence $\pi^*\left((-\infty,x]\times(-\infty,y]\right)=\min\{F(x),G(y)\}$.

Now we generalise to d not strictly convex. Since d is convex it can be bounded below by an affine function. Let $d(x)\geq(ax+b)_+$. One can check that $f(x)=\frac{1}{2}\sqrt{4+(ax+b)^2}+\frac{1}{2}(ax+b)$ is strictly convex and satisfies $0\leq f(x)\leq 1+d(x)$. Then $d_\varepsilon:=d+\varepsilon f$ is strictly convex and satisfies $d\leq d_\varepsilon\leq(1+\varepsilon)d+\varepsilon$. Now let $\pi\in\Pi(\mu,\nu)$, then

$$\begin{aligned}\int_{\mathbb{R}\times\mathbb{R}} d(x-y)\,\mathrm{d}\pi^\dagger(x,y) &\leq \int_{\mathbb{R}\times\mathbb{R}} d_\varepsilon(x-y)\,\mathrm{d}\pi^\dagger(x,y)\\ &\leq \int_{\mathbb{R}\times\mathbb{R}} d_\varepsilon(x-y)\,\mathrm{d}\pi(x,y)\\ &\leq (1+\varepsilon)\int_{\mathbb{R}\times\mathbb{R}} d(x-y)\,\mathrm{d}\pi(x,y)+\varepsilon.\end{aligned}$$

Taking $\varepsilon\to 0$ proves that $\pi^\dagger$ is an optimal plan in the sense of Kantorovich.

Now we show that $\int_{\mathbb{R}\times\mathbb{R}} d(x-y)\,\mathrm{d}\pi^\dagger(x,y)=\int_0^1 d(F^{-1}(t)-G^{-1}(t))\,\mathrm{d}t$. We claim that $\pi^\dagger=(F^{-1},G^{-1})_\#\mathcal{L}\lfloor_{[0,1]}$. Assuming so, then

$$\begin{aligned}\int_{\mathbb{R}\times\mathbb{R}} d(x-y)\,\mathrm{d}\pi^\dagger(x,y) &= \int_{\mathbb{R}\times\mathbb{R}} d(x-y)\,\mathrm{d}\left((F^{-1},G^{-1})_\#\mathcal{L}\lfloor_{[0,1]}\right)(x,y)\\ &= \int_0^1 d(F^{-1}(t)-G^{-1}(t))\,\mathrm{d}t\end{aligned}$$

by the change of variable formula (Proposition 1.5).

To prove the claim we have

$$\begin{aligned}(F^{-1},G^{-1})_\#\mathcal{L}\lfloor_{[0,1]}((-\infty,x]\times(-\infty,y]) &= \mathcal{L}\lfloor_{[0,1]}\Big((F^{-1},G^{-1})^{-1}\left((-\infty,x]\times(-\infty,y]\right)\Big)\\ &= \mathcal{L}\lfloor_{[0,1]}\left(\left\{t\,:\,F^{-1}(t)\leq x \text{ and } G^{-1}(t)\leq y\right\}\right)\\ &= \mathcal{L}\lfloor_{[0,1]}(\{t\,:\,F(x)\geq t \text{ and } G(y)\geq t\})\\ &= \min\{F(x),G(y)\}\\ &= \pi^\dagger\left((-\infty,x]\times(-\infty,y]\right).\end{aligned}$$

where we used $F^{-1}(t)\leq x\Leftrightarrow F(x)\geq t$. □

Remark 3.4. Note that we actually showed that if d is continuous and strictly convex then minimisers of the Kantorovich optimal transport problem are unique.

3.2 Existence of Transport Maps for Discrete Measures

Section references: the discrete special case is based on the proof outlined in the introduction to [23]. The proof of the Minkowski-Carathéodory Theorem comes from [21, Theorem 8.11]

Proving the existence of a transport map $T^\dagger$ that are optimal for Monge's optimal transport problem, i.e. $T^\dagger$ minimises $\mathbb{M}(T)$ over all T satisfying $T_\#\mu=\nu$, is difficult and in fact for general measures we will only consider this problem for a specific cost function $c(x,y)=|x-y|^2$. Here we consider general cost functions but restrict to discrete measures $\mu=\frac{1}{n}\sum_{i=1}^n\delta_{x_i}$ and $\nu=\frac{1}{n}\sum_{j=1}^n\delta_{y_j}$. Note that since all points $X=\{x_i\}_{i=1}^n, Y=\{y_j\}_{j=1}^n$ have equal mass that the map $T:X\to Y$ defined by $T(x_i)=y_{\sigma(i)}$ where $\sigma:\{1,\dots,n\}\to\{1,\dots,n\}$ is a permutation is a transport map (i.e. satisfies (1.1)). Hence the set of transport maps is non-empty.

For a convex and compact set B in a Banach space M we define the set of extreme points, which we denote by $\mathcal{E}(B)$, as the set of points in B that cannot be written as nontrivial convex combinations of points in B. I.e. if $B\ni\pi=\sum_{i=1}^m\alpha_i\pi_i$ (where $\sum_{i=1}^m\alpha_i=1$, $\alpha_i\geq 0$, $\pi_i\in B$) then $\pi\in\mathcal{E}(B)$ if and only if $\alpha_i\in\{0,1\}$. We recall two results. The first is the Minkowski–Carathéodory Theorem. The theorem is set in Euclidean spaces but can be generalised to Banach spaces where it is known as Choquet's theorem.

Theorem 3.5. Minkowski–Carathéodory Theorem. *Let $B\subset\mathbb{R}^M$ be a non-empty, convex and compact set and equip $\mathbb{R}^M$ with the Euclidean metric. Then, for any $\pi^\dagger\in B$ there exists a measure η supported on $\mathcal{E}(B)$ such that for any affine function f*

$$f(\pi^\dagger)=\int f(\pi)\,\mathrm{d}\eta(\pi).$$

Proof. Let $d=\dim(B)$. It is enough to show that there exists $\{a_i\}_{i=0}^d$ and $\{\pi^{(i)}\}_{i=0}^d$ such that $\pi^\dagger=\sum_{i=0}^d a_i\pi^{(i)}$ where $\sum_{i=0}^n a_i=1$ and $\{\pi^{(i)}\}_{i=0}^d\subset\mathcal{E}(B)$. We prove the result by induction. The case when $d=0$ is trivial since B is just a point.

Now assume the result is true for all sets of dimension at most $d-1$. Pick $\pi^\dagger\in B$ and assume $\pi^\dagger\notin\mathcal{E}(B)$. Pick $\pi^{(0)}\in\mathcal{E}(B)$ and take the line segment $[\pi^{(0)},\pi^\dagger]$ and extend it until it intersects with the boundary of B, i.e. let θ parametrise the line then $\{\theta\,:\,(1-\theta)\pi^{(0)}+\theta\pi^\dagger\in B\}=[0,\alpha]$ for some $\alpha\geq 1$ (where α exists and is finite by convexity and compactness of B). Let $\xi=(1-\alpha)\pi^{(0)}+\alpha\pi^\dagger$ then $\pi^\dagger=(1-\theta_0)\xi+\theta_0\pi^{(0)}$ where $\theta_0=1-\frac{1}{\alpha}$. Now since $\xi\in F$ for some proper face F of B[1] then by the induction hypothesis there exists $\{\pi^{(i)}\}_{i=1}^d$ such that $\xi=\sum_{i=1}^d\theta_i\pi^{(i)}$ with $\sum_{i=1}^d\theta_i=1$. Hence, $\pi=\sum_{i=1}^d(1-\theta_0)\theta_i\pi^{(i)}+\theta_0\pi^{(0)}$. Since $(1-\theta_0)\sum_{i=1}^d\theta_i+\theta_0=1$ then π is a convex combination of $\{\pi^{(i)}\}_{i=0}^d$. □

Theorem 3.6. Birkhoff's theorem. *Let B be the set of $n\times n$ bistochastic matrices, i.e.*

$$B=\left\{\pi\in\mathbb{R}^{n\times n}\,:\,\forall ij,\,\pi_{ij}\geq 0;\,\forall j,\,\sum_{i=1}^n\pi_{ij}=1;\forall i,\,\sum_{j=1}^n\pi_{ij}=1\right\}.$$

Then the set of extremal points $\mathcal{E}(B)$ of B is exactly the set of permutation matrices, i.e.

$$\mathcal{E}(B)=\left\{\pi\in\{0,1\}^{n\times n}\,:\,\forall j,\,\sum_{i=1}^n\pi_{ij}=1;\forall i,\,\sum_{j=1}^n\pi_{ij}=1\right\}.$$

[1] A face F of a convex set B is any set with the property that if $\pi^{(1)},\pi^{(2)}\in B$, $t\in(0,1)$ and $t\pi^{(1)}+(1-t)\pi^{(2)}\in F$ then $\pi^{(1)},\pi^{(2)}\in F$. A proper face is a face which has dimension at most $\dim(B)-1$. A result we use without proof is that the boundary of a convex set is the union of all proper faces.

Proof. We start by showing that every permutation matrix is an extremal point. Let $\pi_{ij} = \delta_{j=\sigma(i)}$ where σ is a permutation. Assume that $\pi \notin \mathcal{E}(B)$. Then there exists $\pi^{(1)}, \pi^{(2)} \in B$, with $\pi^{(1)} \neq \pi \neq \pi^{(2)}$, and $t \in (0,1)$ such that $\pi = t\pi^{(1)} + (1-t)\pi^{(2)}$. Let ij be such that $0 = \pi_{ij} \neq \pi_{ij}^{(1)}$, then

$$0 = \pi_{ij} = t\pi_{ij}^{(1)} + (1-t)\pi_{ij}^{(2)} \quad \implies \quad \pi_{ij}^{(2)} = -\frac{\pi_{ij}^{(1)}}{1-t} < 0.$$

This contradicts $\pi_{ij}^{(2)} \geq 0$, hence $\pi \in \mathcal{E}(B)$.

Now we show that every $\pi \in \mathcal{E}(B)$ is a permutation matrix. We do this in two parts: we (i) show that $\pi \in \mathcal{E}(B)$ implies that $\pi_{ij} \in \{0,1\}$, then (ii) show $\pi = \delta_{j=\sigma(i)}$ for a permutation σ.

For (i) let $\pi \in \mathcal{E}(B)$ and assume there exists $i_1 j_1$ such that $\pi_{i_1 j_1} \in (0,1)$. Since $\sum_{i=1}^n \pi_{ij_1} = 1$ then there exists $i_2 \neq i_1$ such that $\pi_{i_2 j_1} \in (0,1)$. Similarly, since $\sum_{j=1}^n \pi_{i_2 j} = 1$ there exists $j_2 \neq j_1$ such that $\pi_{i_2 j_2} \in (0,1)$. Continuing this procedure until $i_m = i_1$ we obtain two sequences:

$$\mathcal{I} = \{i_k j_k \,:\, k \in \{1,\dots,m-1\}\} \qquad \mathcal{I}^+ = \{i_{k+1} j_k \,:\, k \in \{1,\dots,m-1\}\}$$

with $i_{k+1} \neq i_k$ and $j_{k+1} \neq j_k$. Define $\pi^{(\delta)}$ by the following

$$\pi_{ij}^{(\delta)} = \begin{cases} \pi_{i_k j_k} + \delta & \text{if } ij = i_k j_k \text{ for some } k \\ \pi_{i_{k+1} j_k} - \delta & \text{if } ij = i_{k+1} j_k \text{ for some } k \\ \pi_{ij} & \text{else.} \end{cases}$$

Then,

$$\sum_{i=1}^n \pi_{ij}^{(\delta)} = \sum_{i=1}^n \pi_{ij} + \delta \left|\{ij \in \mathcal{I} \,:\, i \in \{1,\dots,n\}\}\right| - \delta \left|\{ij \in \mathcal{I}^+ \,:\, i \in \{1,\dots,n\}\}\right|.$$

Now if $ij \in \mathcal{I}$ then there exists i' such that $i'j \in \mathcal{I}^+$, and likewise, if $ij \in \mathcal{I}^+$ then there exists i' such that $i'j \in \mathcal{I}$. Hence,

$$\left|\{ij \in \mathcal{I} \,:\, i \in \{1,\dots,n\}\}\right| = \left|\{ij \in \mathcal{I}^+ \,:\, i \in \{1,\dots,n\}\}\right|.$$

It follows that $\sum_{i=1}^n \pi_{ij}^{(\delta)} = 1$ and analogously $\sum_{j=1}^n \pi_{ij}^{(\delta)} = 1$.

Choose $\delta = \min\{\min\{\pi_{ij}, 1-\pi_{ij}\} \,:\, ij \in \mathcal{I} \cup \mathcal{I}^+\} \in (0,1)$. Define $\pi^{(1)} = \pi^{(-\delta)}$, $\pi^{(\delta)}$. We have that $\pi_{ij}^{(1)}, \pi_{ij}^{(2)} \geq 0$ and therefore $\pi^{(1)}, \pi^{(2)} \in B$ with $\pi^{(1)} \neq \pi^{(2)}$. Moreover we have $\pi = \frac{1}{2}\pi^{(1)} + \frac{1}{2}\pi^{(2)}$. Hence, $\pi \notin \mathcal{E}(B)$. The contradiction implies that there does not exist $i_1 j_1$ such that $\pi_{i1 j_1} \in (0,1)$. We have shown that if $\pi \in \mathcal{E}(B)$ then $\pi_{ij} \in \{0,1\}$.

We're left to show (ii): that $\pi_{ij} = \delta_{j=\sigma(i)}$. Since $\pi \in B$ then for each i there exists j^* such that $\pi_{ij^*} = 1$ (else $\sum_{j=1}^n \pi_{ij} \neq 1$). We let $\sigma(i) = j^*$ so by construction we have $\pi_{i\sigma(i)} = 1$. We claim σ is a permutation. It is enough to show that σ is injective. Now if $j = \sigma(i_1) = \sigma(i_2)$ where $i_1 \neq i_2$ then

$$1 = \sum_{i=1}^n \pi_{ij} \geq \pi_{i_1 j} + \pi_{i_2 j} = 2.$$

The contradiction implies that $i_1 = i_2$ and therefore σ is injective. $\square$

We now show that the existence of optimal transport maps between discrete measures $\mu = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\nu = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$.

Theorem 3.7. *Let $\mu = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\nu = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$. Assume that $\{x_i\}_{i=1}^n$ are unique. Then, for any $c : \{x_i\}_{i=1}^n \times \{y_j\}_{j=1}^n \to \mathbb{R}$ there exists a solution to Monge's optimal transport problem between μ and ν and furthermore*

$$\min_{\pi\in\Pi(\mu,\nu)} \mathbb{K}(\pi) = \min_{T\,:\,T_{\#}\mu=\nu} \mathbb{M}(T).$$

Proof. Let $c_{ij} = c(x_i, y_j)$ and B be the set of bistochastic $n \times n$ matrices, i.e.

$$B = \left\{ \pi \in \mathbb{R}^{n\times n} \,:\, \forall ij,\, \pi_{ij} \geq 0;\, \forall j,\, \sum_{i=1}^n \pi_{ij} = 1; \forall i,\, \sum_{j=1}^n \pi_{ij} = 1 \right\}.$$

With an abuse of notation we can say $\pi \in B$ if and only if $\frac{1}{n}\pi \in \Pi(\mu,\nu)$. The Kantorovich problem reads as

$$\text{minimise } \frac{1}{n}\sum_{i,j=1}^n c_{ij}\pi_{ij} \quad \text{over } \pi \in B.$$

Although, by Proposition 2.4, there exists a minimiser to the Kantorovich optimal transport problem we do not use this fact here. Let M be the minimum of the Kantorovich optimal transport problem, $\varepsilon > 0$ and find an approximate minimiser $\pi^\varepsilon \in B$ such that

$$M \geq \frac{1}{n}\sum_{i,j=1}^n c_{ij}\pi^\varepsilon - \varepsilon.$$

If we let $f(\pi) = \frac{1}{n}\sum_{i,j=1}^n c_{ij}\pi_{ij}$ then assuming that B is compact and convex we have that there exists a measure η supported on $\mathcal{E}(B)$ such that

$$f(\pi^\varepsilon) = \int f(\pi)\,\mathrm{d}\eta(\pi).$$

Hence

$$M \geq \int \frac{1}{n}\sum_{i,j=1}^n c_{ij}\pi_{ij}\,\mathrm{d}\eta(\pi) - \varepsilon \geq \inf_{\pi\in\mathcal{E}(B)} \frac{1}{n}\sum_{i,j=1}^n c_{ij}\pi_{ij} - \varepsilon \geq M - \varepsilon.$$

Since this is true for all ε it holds that $\inf_{\pi\in\mathcal{E}(B)} \frac{1}{n}\sum_{i,j=1}^n c_{ij}\pi_{ij} = M$. Since $\mathcal{E}(B)$ is finite ($\#\{\pi \in \mathcal{E}(B)\} \leq n!$) there exists a minimiser $\pi^\dagger \in \mathcal{E}(B)$ of f, i.e. $\frac{1}{n}\sum_{i,j=1}^n c_{ij}\pi_{ij}^\dagger = M$. Note that we have also shown (independently from Proposition 2.4) that since $\inf_{\pi\in\mathcal{E}(B)} \frac{1}{n}\sum_{i,j=1}^n c_{ij}\pi_{ij} = \inf_{\pi\in B} \frac{1}{n}\sum_{i,j=1}^n c_{ij}\pi_{ij}$ there exists a solution to Kantorovich's optimal transport problem.

By Birkhoff's theomem $\pi^\dagger$ is a permutation matrix, that is there exists a permutation $\sigma^\dagger : \{1,\dots,n\} \to \{1,\dots,n\}$ such that $\pi_{ij}^\dagger = \delta_{j=\sigma^\dagger(i)}$. Let $T^\dagger : X \to Y$ be defined by $T^\dagger(x_i) = y_{\sigma(i)}$. Since $x_i \neq x_j$ for all $i \neq j$ then we have that $T^\dagger$ is well-defined.

We already know that the set of transport maps is non-empty. Let T be any transport map and define $\pi_{ij} = \delta_{y_j = T(x_i)}$, (it is an exercise to show that $\pi \in B$) then

$$\mathbb{M}(T) = \frac{1}{n}\sum_{i=1}^n c(x_i, T(x_i)) = \frac{1}{n}\sum_{ij} c_{ij}\pi_{ij} \geq \frac{1}{n}\sum_{ij} c_{ij}\pi_{ij}^\dagger = \frac{1}{n}\sum_{i=1}^n c(x_i, T^\dagger(x_i)) = \mathbb{M}(T^\dagger).$$

Hence $T^\dagger$ is a solution to Monge's optimal transport problem.

We are left to show that B is compact and convex. To show B is compact we consider the ℓ^1 norm: $\|\pi\|_1 := \sum_{ij} |\pi_{ij}|$ (since all norms are equivalent on finite dimensional spaces it does not really matter which norm we choose). Clearly B is bounded as for all $\pi \in B$ we have $\|\pi\|_1 = 1$. For closure, we consider a sequence $\pi^{(m)} \in B$ with $\pi^{(m)} \to \pi$. Trivially $\pi_{ij}^{(m)} \to \pi_{ij}$ for all ij and therefore $\pi_{ij} \geq 0$, likewise $\sum_{i=1}^n \pi_{ij} = \lim_{m\to\infty} \sum_{i=1}^n \pi_{ij}^{(m)} = 1$ and $\sum_{j=1}^n \pi_{ij} = 1$. Hence $\pi \in B$ and B is closed. Therefore B is compact.

Convexity of B is easy to check by considering $\pi^{(1)}, \pi^{(2)} \in B$ and $\pi = t\pi^{(1)} + (1-t)\pi^{(2)}$ for $t \in [0,1]$ then clearly $\pi_{ij} \geq 0$,

$$\sum_{i=1}^n \pi_{ij} = t\sum_{i=1}^n \pi_{ij}^{(1)} + (1-t)\sum_{i=1}^n \pi_{ij}^{(2)} = t + (1-t) = 1,$$

and similarly $\sum_{j=1}^n \pi_{ij} = 1$. Hence $\pi \in B$ and B is convex. □

Remark 3.8. Note that the set of transport maps $T : \{x_i\}_{i=1}^n \to \{y_j\}_{j=1}^n$. Indeed, there are at most $n!$ such maps. Since the Monge problem is taken over a finite set then it immediately follows that there exists a minimiser.

Chapter 4

Kantorovich Duality

We saw in the previous chapter how Kantorovich's optimal transport problem resembles a linear programme. It should not therefore be surprising that Kantorovich's optimal transport problem admits a dual formulation. In the following section we state the duality result and give an intuitive but non-rigorous proof. In Section 4.2 we give a general minimax principle upon which we can base the proof of Kantorovich duality. In Section 4.3 we can then rigorously prove duality. With additional assumptions such as restricting X, Y to Euclidean spaces we prove the existence of solutions to the dual problem in Section 4.4. In Section 4.5 we prove the Kantorovich-Rubinstein theorem.

4.1 Kantorovich Duality

Section references: The statement and proof of the main result, Theorem 4.1, come from [23, Theorem 1.3].

We start by stating Kantorovich then give an intuitive proof with one key step missing. The proof is made rigorous in Section 4.3.

Theorem 4.1. Kantorovich Duality. *Let $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ where X, Y are Polish spaces. Let $c : X \times Y \to [0, +\infty]$ be a lower semi-continuous cost function. Define $\mathbb{K}$ as in Definition 2.3 and $\mathbb{J}$ by*

$$\mathbb{J} : L^1(\mu) \times L^1(\nu) \to \mathbb{R}, \qquad \mathbb{J}(\varphi, \psi) = \int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu. \tag{4.1}$$

Let Φ_c be defined by

$$\Phi_c = \left\{ (\varphi, \psi) \in L^1(\mu) \times L^1(\nu) \, : \, \varphi(x) + \psi(y) \leq c(x, y) \right\} \tag{4.2}$$

where the inequality is understood to hold for μ-almost every $x \in X$ and ν-almost every $y \in Y$. Then,

$$\min_{\pi \in \Pi(\mu,\nu)} \mathbb{K}(\pi) = \sup_{(\varphi,\psi) \in \Phi_c} \mathbb{J}(\varphi, \psi).$$

Let us give an informal interpretation of the result which originally comes from Caffarelli and I take from Villani [23]. Consider the *shippers problem*. Suppose we own a number of coal mines and a number of factories, we wish to transport the coal from mines to factories. The amount each mine produces and each factory requires is fixed (and we assume equal). The cost for us to transport from mine x to factory y is $c(x,y)$. The total optimal cost is the solution to Kantorovich's optimal transport problem. Now a clever shipper comes to you and says they will ship for you and you just pay a price $\varphi(x)$ for loading and $\psi(y)$ for unloading. To make it in your interest the shipper makes sure that $\varphi(x)+\psi(y)\leq c(x,y)$ that is the cost is no more than what you would have spent transporting the coal yourself. Kantorovich duality tells us that one can find φ and ψ such that this price scheme costs just as much as paying for the cost of transport yourself.

We now give an informal proof that will later be made rigorous. Let $M=\inf_{\pi\in\Pi(\mu,\nu)}\mathbb{K}(\pi)$. Observe that

$$M=\inf_{\pi\in\mathcal{M}_+(X\times Y)}\sup_{(\varphi,\psi)}\left(\int_{X\times Y}c(x,y)\,\mathrm{d}\pi+\int_X\varphi\,\mathrm{d}\left(\mu-P^X_\#\pi\right)+\int_Y\psi\,\mathrm{d}\left(\nu-P^Y_\#\pi\right)\right)\tag{4.3}$$

where we take the supremum on the right hand side over $(\varphi,\psi)\in C^0_b(X)\times C^0_b(Y)$. This follows since

$$\sup_{\varphi\in C^0_b(X)}\int_X\varphi\,\mathrm{d}\left(\mu-P^X_\#\pi\right)=\begin{cases}+\infty & \text{if } \mu\neq P^X_\#\pi\\ 0 & \text{else.}\end{cases}$$

Hence, the infimum over π of the right hand side of (4.3) is on the set where $P^X_\#\pi=\mu$ and, similarly, $P^Y_\#\pi=\nu$ (which means that $\pi\in\Pi(\mu,\nu)$). We can rewrite (4.3) more conveniently as

$$M=\inf_{\pi\in\mathcal{M}_+(X\times Y)}\sup_{(\varphi,\psi)}\left(\int_{X\times Y}c(x,y)-\varphi(x)-\psi(y)\,\mathrm{d}\pi+\int_X\varphi\,\mathrm{d}\mu+\int_Y\psi\,\mathrm{d}\nu\right).$$

Assuming a minimax principle we switch the infimum and supremum to obtain

$$M=\sup_{(\varphi,\psi)}\left(\int_X\varphi\,\mathrm{d}\mu+\int_Y\psi\,\mathrm{d}\nu+\inf_{\pi\in\mathcal{M}_+(X\times Y)}\int_{X\times Y}c(x,y)-\varphi(x)-\psi(y)\,\mathrm{d}\pi\right).\tag{4.4}$$

Now if there exists $(x_0,y_0)\in X\times Y$ and $\varepsilon>0$ such that $\varphi(x_0)+\psi(y_0)-c(x_0,y_0)=\varepsilon>0$ then by letting $\pi_\lambda=\lambda\delta_{(x_0,y_0)}$ for some $\lambda>0$ we have

$$\inf_{\pi\in\mathcal{M}_+(X\times Y)}\int_{X\times Y}c(x,y)-\varphi(x)-\psi(y)\,\mathrm{d}\pi\leq-\lambda\varepsilon\to-\infty\quad\text{as }\lambda\to\infty.$$

Hence the infimum on right hand side of (4.4) can be restricted to when $\varphi(x)+\psi(y)\leq c(x,y)$ for all $(x,y)\in X\times Y$, i.e. $(\varphi,\psi)\in\Phi_c$ (this heuristic argument actually used $(\varphi,\psi)\in C^0_b(X)\times C^0_b(Y)$ not $L^1(\mu)\times L^1(\nu)$ and there is a difference between the constraint $\varphi(x)+\psi(y)\leq c(x,y)$ holding everywhere and holding *almost* everywhere, these are technical details that are not important at this stage). When $(\varphi,\psi)\in\Phi_c$ then

$$\inf_{\pi\in\mathcal{M}_+(X\times Y)}\int_{X\times Y}c(x,y)-\varphi(x)-\psi(y)\,\mathrm{d}\pi=0$$

which is achieved for $\pi \equiv 0$ for example. Hence,

$$\inf_{\pi \in \Pi(\mu,\nu)} \mathbb{K}(\pi) = \sup_{(\varphi,\psi)\in\Phi_c} \int_X \varphi(x)\,\mathrm{d}\mu(x) + \int_Y \psi(y)\,\mathrm{d}\nu(y).$$

This is the statement of Kantorovich duality. To complete this argument we need to make the minimax principle rigorous. In the next section we prove a minimax principle, in the section after we apply it to Kantorovich duality and provide a complete proof.

4.2 Fenchel-Rockafeller Duality

Section references: I take the duality theorem (Theorem 4.3) from [23, Theorem 1.9]. Lemma 4.4 is hopefully obvious and the Hahn-Banach theorem is well known.

To rigorously prove the Kantorovich duality theorem we need a minimax principle, i.e. conditions sufficient to interchange the infimum and supremum when we introduced the Lagrange multipliers φ, ψ in (4.3). The minimax principle is specific to convex functions; at this stage it is perhaps not clear how to apply it to Kantorovich's optimal transport problem when we made no convexity assumption on c. In this section we will make use of the Legendre-Fenchel transform.

Definition 4.2. *We define the Legendre-Fenchel transform (also called the convex conjugate) of a convex function $\Theta : E \to \mathbb{R} \cup \{+\infty\}$, where E is a normed vector space, by*

$$\Theta^* : E^* \to \mathbb{R} \cup \{+\infty\}, \qquad \Theta^*(z^*) = \sup_{z\in E} \left(\langle z^*, z\rangle - \Theta(z)\right).$$

Convex analysis will play a greater role in the sequel, in particular in Chapter 6 where we will provide a more in-depth review. We now state the minimax principle taken from Villani [23].

Theorem 4.3. Fenchel-Rockafellar Duality. *Let E be a normed vector space and $\Theta, \Xi : E \to \mathbb{R} \cup \{+\infty\}$ two convex functions. Assume there exists $z_0 \in E$ such that $\Theta(z_0) < \infty$, $\Xi(z_0) < \infty$ and Θ is continuous at z_0. Then,*

$$\inf_E (\Theta + \Xi) = \max_{z^*\in E^*} \left(-\Theta^*(-z^*) - \Xi^*(z^*)\right).$$

In particular the supremum on the right hand side is attained.

We recall a couple of preliminary results (that we do not prove) before we prove the theorem.

Lemma 4.4. *Let E be a normed vector space.*

1. *If $\Theta : E \to \mathbb{R} \cup \{+\infty\}$ is convex then the epigraph A defined by*

   $$A = \{(z,t) \in E \times \mathbb{R} \,:\, t \geq \Theta(z)\}$$

   is also convex.

2. *If* $\Theta : E \to \mathbb{R} \cup \{+\infty\}$ *is concave then the hypograph* B *defined by*

$$B = \{(z,t) \in E \times \mathbb{R} \,:\, t \le \Theta(z)\}$$

is convex.

3. *If* $C \subset E$ *is convex then* $\mathrm{int}(C)$ *is convex.*

4. *If* $D \subset E$ *is convex and* $\mathrm{int}(D) \neq \emptyset$ *then* $\overline{D} = \overline{\mathrm{int}(D)}$*.*

The following theorem, the Hahn-Banach theorem can be stated in multiple different forms. The most convenient form for us is in terms of separation of convex sets.

Theorem 4.5. Hahn-Banach Theorem. *Let* E *be a topological vector space. Assume* A, B *are convex, non-empty and disjoint subsets of* E*, and that* A *is open. Then there exists a closed hyperplane separating* A *and* B*.*

We now prove Theorem 4.3.

Proof of Theorem 4.3. By writing

$$-\Theta^*(-z^*) - \Xi^*(z^*) = \inf_{x,y\in E} \left(\Theta(x) + \Xi(y) + \langle z^*, x-y\rangle\right)$$

and choosing $y = x$ on the right hand side we see that

$$\inf_{x\in E} \left(\Theta(x) + \Xi(x)\right) \ge \sup_{z^*\in E^*} \left(-\Theta^*(-z^*) - \Xi^*(z^*)\right).$$

Let $M = \inf\,(\Theta + \Xi)$, and define the sets A, B by

$$\begin{aligned} A &= \{(x,\lambda) \in E \times \mathbb{R} \,:\, \lambda \ge \Theta(x)\} \\ B &= \{(y,\sigma) \in E \times \mathbb{R} \,:\, \sigma \le M - \Xi(y)\}. \end{aligned}$$

By Lemma 4.4 A and B are convex. By continuity and finiteness of Θ at z_0 the interior of A is non-empty and by finiteness of Ξ at z_0 B is non-empty. Let $C = \mathrm{int}(A)$ (which is convex by Lemma 4.4). Now, if $(x,\lambda) \in C$ then $\lambda > \Theta(x)$, therefore $\lambda + \Xi(x) > \Theta(x) + \Xi(x) \ge M$. Hence $(x,\lambda) \notin B$. In particular $B \cap C = \emptyset$. By the Hahn-Banach theorem there exists a hyperplane $H = \{\Phi = \alpha\}$ that separates B and C, i.e. if we write $\Phi(x,\lambda) = f(x) + k\lambda$ (where f is linear) then

$$\begin{aligned} \forall (x,\lambda) \in C, &\qquad f(x) + k\lambda \ge \alpha \\ \forall (x,\lambda) \in B, &\qquad f(x) + k\lambda \le \alpha. \end{aligned}$$

Now if $(x,\lambda) \in A$ then there exists a sequence $(x_n,\lambda_n) \in C$ such that $(x_n,\lambda_n) \to (x,\lambda)$. Hence $f(x) + k\lambda \leftarrow f(x_n) + k\lambda_n \ge \alpha$. Therefore

$$\forall (x,\lambda) \in A, \qquad f(x) + k\lambda \ge \alpha \tag{4.5}$$

$$\forall (x,\lambda) \in B, \qquad f(x) + k\lambda \le \alpha. \tag{4.6}$$

We know that $(z_0, \lambda) \in A$ for λ sufficiently large, hence $k \geq 0$. We claim $k > 0$. Assume $k = 0$. Then

$$\begin{aligned} \forall (x,\lambda) \in A, \quad f(x) \geq \alpha \qquad &\Longrightarrow \qquad f(x) \geq \alpha \quad \forall x \in \mathrm{Dom}(\Theta) \\ \forall (x,\lambda) \in B, \quad f(x) \leq \alpha \qquad &\Longrightarrow \qquad f(x) \leq \alpha \quad \forall x \in \mathrm{Dom}(\Xi). \end{aligned}$$

As $\mathrm{Dom}(\Xi) \ni z_0 \in \mathrm{Dom}(\Theta)$ then $f(z_0) = \alpha$. Since Θ is continuous at z_0 there exists $r > 0$ such that $B(z_0, r) \subset \mathrm{Dom}(\Theta)$, hence for all z with $\|z\| < r$ and $\delta \in \mathbb{R}$ with $|\delta| < 1$ we have

$$f(z_0 + \delta z) \geq \alpha \quad \Longrightarrow \quad f(z_0) + \delta f(z) \geq \alpha \quad \Longrightarrow \quad \delta f(z) \geq 0.$$

This is true for all $\delta \in (-1, 1)$ and therefore $f(z) = 0$ for $z \in B(0, r)$. Hence $f \equiv 0$ on E. It follows that $\Phi \equiv 0$ which is clearly a contradiction (either $H = E \times \mathbb{R}$ if $\alpha = 0$ or $H = \emptyset$). It must be that $k > 0$.

By (4.5) we have

$$\begin{aligned} \Theta^*\left(-\frac{f}{k}\right) &= \sup_{z\in E}\left(-\frac{f(z)}{k} - \Theta(z)\right) \\ &= -\frac{1}{k}\inf_{z\in E}\left(f(z) + k\Theta(z)\right) \\ &\leq -\frac{\alpha}{k} \end{aligned}$$

since $(z, \Theta(z)) \in A$. Similarly, by (4.6) we have

$$\begin{aligned} \Xi^*\left(\frac{f}{k}\right) &= \sup_{z\in E}\left(\frac{f(z)}{k} - \Xi(z)\right) \\ &= -M + \frac{1}{k}\sup_{z\in E}\left(f(z) + k(M - \Xi(z))\right) \\ &\leq -M + \frac{\alpha}{k} \end{aligned}$$

since $(z, M - \Xi(z)) \in B$. It follows that

$$M \geq \sup_{z^*\in E^*}\left(-\Theta^*(-z^*) - \Xi^*(z^*)\right) \geq -\Theta^*\left(-\frac{f}{k}\right) - \Xi^*\left(\frac{f}{k}\right) \geq \frac{\alpha}{k} + M - \frac{\alpha}{k} = M.$$

So

$$\inf_{x\in E}\left(\Theta(x) + \Xi(x)\right) = M = \sup_{z^*\in E^*}\left(-\Theta^*(-z^*) - \Xi^*(z^*)\right).$$

Furthermore $z^* = \frac{f}{k}$ must achieve the supremum. $\square$

Remark 4.6. To explain where the choice of $z^* = \frac{f}{k}$ in the above proof let us consider the Euclidean setting and assume that Θ and Ξ are differentiable. In terms of a Lagrange multiplier we can write

$$\inf_{x\in E}\left(\Theta(x) + \Xi(x)\right) = \inf_{x,y\in E}\sup_{z^*\in E^*}\left(\Theta(x) + \Xi(y) + \langle z^*, x - y\rangle\right).$$

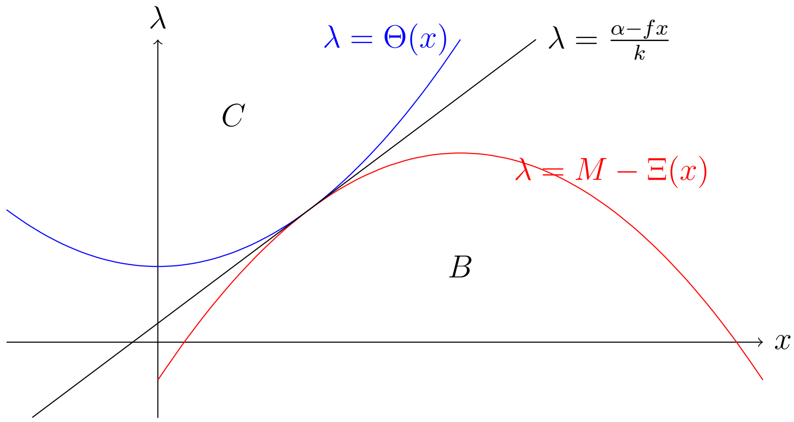

Figure 4.1: Illustration of the Fenchel-Rockafellar Duality.

For fixed z^* the infimum of $\Theta(x) + \Xi(y) + \langle z^*, x - y\rangle$ is achieved when

$$\nabla\Theta(x) = -z^* \quad \text{and} \quad \nabla\Xi(x) = z^*.$$

From the Lagrange formulation we expect (and hope) that $x = y$, hence we want to find a z^* and x such that $z^* = -\nabla\Theta(x) = \nabla\Xi(x)$. From Figure 4.1 we see that the candidate z^* is $\frac{f}{k}$.

4.3 Proof of Kantorovich Duality

Section references: The two lemmas in this section together prove the Kantorovich duality theorem, both lemmas come from [23].

Finally we can prove Kantorovich dualiy as stated in Theorem 4.1. We break the theorem into two parts.

Lemma 4.7. *Under the same conditions as Theorem 4.1 we have*

$$\sup_{(\varphi,\psi)\in\Phi_c} \mathbb{J}(\varphi,\psi) \leq \inf_{\pi\in\Pi(\mu,\nu)} \mathbb{K}(\pi).$$

Proof. Let $(\varphi,\psi) \in \Phi_c$ and $\pi \in \Pi(\mu,\nu)$. Let $A \subset X$ and $B \subset Y$ be sets such that $\mu(A) = 1$, $\nu(B) = 1$ and

$$\varphi(x) + \psi(y) \leq c(x,y) \qquad \forall (x,y) \in A \times B.$$

Now $\pi(A^c \times B^c) \leq \pi(A^c \times Y) + \pi(X \times B^c) = \mu(A^c) + \nu(B^c) = 0$. Hence,

$$\begin{aligned}
\pi(A \times B) &= \pi(X \times B) - \pi(A^c \times B)\\
&= \nu(B) - \pi(A^c \times Y) + \pi(A^c \times B^c)\\
&= 1 - \mu(A^c) + \pi(A^c \times B^c)\\
&= 1.
\end{aligned}$$

So it follows that $\varphi(x)+\psi(y)\leq c(x,y)$ for π-almost every (x,y). Then,

$$\mathbb{J}(\varphi,\psi)=\int_X \varphi\,\mathrm{d}\mu+\int_Y \psi\,\mathrm{d}\nu=\int_{X\times Y}\varphi(x)+\psi(y)\,\mathrm{d}\pi(x,y)\leq\int_{X\times Y}c(x,y)\,\mathrm{d}\pi(x,y).$$

The result of the lemma follows by taking the supremum over $(\varphi,\psi)\in\Phi_c$ on the right hand side and the infimum over $\pi\in\Pi(\mu,\nu)$ on the left hand side. $\square$

To complete the proof of Theorem 4.1 we need to show that the opposite inequality in Lemma 4.7 is also true.

Lemma 4.8. *Under the same conditions as Theorem 4.1 we have*

$$\sup_{(\varphi,\psi)\in\Phi_c}\mathbb{J}(\varphi,\psi)\geq\inf_{\pi\in\Pi(\mu,\nu)}\mathbb{K}(\pi).$$

Proof. The proof is completed in three steps in increasing generality:

1. we assume X,Y are compact and c is continuous;
2. the assumption that X,Y are compact is relaxed, c is still continuous;
3. c is only assumed to be lower semi-continuous.

1. Let $E=C_b^0(X\times Y)$ equipped with the supremum norm. The dual space of E is the space of Radon measures $E^*=\mathcal{M}(X\times Y)$ (by the Riesz–Markov–Kakutani representation theorem[1]). Define

$$\Theta(u)=\begin{cases}0 & \text{if } u(x,y)\geq -c(x,y)\\ +\infty & \text{else,}\end{cases}$$
$$\Xi(u)=\begin{cases}\int_X\varphi(x)\,\mathrm{d}\mu(x)+\int_Y\psi(y)\,\mathrm{d}\nu(y) & \text{if } u(x,y)=\varphi(x)+\psi(y)\\ +\infty & \text{else.}\end{cases}$$

Note that although the representation $u(x,y)=\varphi(x)+\psi(y)$ is not unique (φ and ψ are only unique upto a constant) Ξ is still well defined. We claim that Θ and Ξ are convex. For Θ consider u,v with $\Theta(u),\Theta(v)<+\infty$, then $u(x,y)\geq -c(x,y)$ and $v(x,y)\geq -c(x,y)$, hence $tu(x,y)+(1-t)v(x,y)\geq c(x,y)$ for any $t\in[0,1]$. It follows that

$$\Theta(tu+(1-t)v)=0=t\Theta(u)+(1-t)\Theta(v).$$

If either $\Theta(u)=+\infty$ or $\Theta(v)=+\infty$ then clearly

$$\Theta(tu+(1-t)v)\leq t\Theta(u)+(1-t)\Theta(v).$$

[1] The Riesz–Markov–Kakutani representation theorem states that the dual space of $C_c^0(Z)$ is $\mathcal{M}(Z)$ for any locally compact space Z. If we apply this to a non-compact set $Z=X\times Y$ then we must have that any $u\in C_c^0(Z)$ converges to zero at infinity, but Exercise 2.6 shows that unless $u\equiv 0$ then $\Xi(u)=+\infty$, this is where the proof breaks down for non-compact sets X and Y

Hence Θ is convex. For Ξ if either $\Xi(u) = +\infty$ or $\Xi(v) = +\infty$ then clearly

$$\Xi(tu + (1-t)v) \leq t\Xi(u) + (1-t)\Xi(v).$$

Assume $u(x,y) = \varphi_1(x) + \psi_1(y)$, $v(x,y) = \varphi_2(x) + \psi_2(y)$ then

$$tu(x,y) + (1-t)v(x,y) = t\varphi_1(x) + (1-t)\varphi_2(x) + t\psi_1(y) + (1-t)\psi_2(y)$$

and therefore

$$\Xi(tu + (1-t)v) = \int_X t\varphi_1 + (1-t)\varphi_2 \, \mathrm{d}\mu + \int_Y t\psi_1 + (1-t)\psi_2 \, \mathrm{d}\nu = t\Xi(u) + (1-t)\Xi(v).$$

Hence Ξ is convex.

Let $u \equiv 1$ then $\Theta(u), \Xi(u) < +\infty$ and Θ is continuous at u. By Theorem 4.3

$$\inf_{u \in E} \left(\Theta(u) + \Xi(u)\right) = \max_{\pi \in E^*} \left(-\Theta^*(-\pi) - \Xi^*(\pi)\right). \tag{4.7}$$

First we calculate the left hand side of (4.7). We have

$$\inf_{u\in E} \left(\Theta(u) + \Xi(u)\right) \geq \inf_{\substack{\varphi(x)+\psi(y)\geq -c(x,y) \\ \varphi \in L^1(\mu), \psi \in L^1(\nu)}} \int_X \varphi(x) \, \mathrm{d}\mu(x) + \int_Y \psi(y) \, \mathrm{d}\nu(y) = - \sup_{(\varphi,\psi)\in\Phi_c} \mathbb{J}(\varphi,\psi).$$

We now consider the right hand side of (4.7). To do so we need to find the convex conjugates of Θ and Ξ. For Θ^* we compute

$$\Theta^*(-\pi) = \sup_{u\in E} \left(- \int_{X\times Y} u \, \mathrm{d}\pi - \Theta(u) \right) = \sup_{u \geq -c} - \int_{X\times Y} u \, \mathrm{d}\pi = \sup_{u\leq c} \int_{X\times Y} u \, \mathrm{d}\pi.$$

Then we find

$$\Theta^*(-\pi) = \begin{cases} \int_{X\times Y} c(x,y) \, \mathrm{d}\pi & \text{if } \pi \in \mathcal{M}_+(X\times Y) \\ +\infty & \text{else.} \end{cases}$$

For Ξ^* we have

$$\begin{aligned} \Xi^*(\pi) &= \sup_{u\in E} \left(\int_{X\times Y} u \, \mathrm{d}\pi - \Xi(u) \right) \\ &= \sup_{u(x,y)=\varphi(x)+\psi(y)} \left(\int_{X\times Y} u \, \mathrm{d}\pi - \int_X \varphi(x) \, \mathrm{d}\mu - \int_Y \psi(y) \, \mathrm{d}\nu \right) \\ &= \sup_{u(x,y)=\varphi(x)+\psi(y)} \left(\int_X \varphi \, \mathrm{d}(P_\#^X \pi - \mu) + \int_Y \psi \, \mathrm{d}(P_\#^Y \pi - \nu) \right) \\ &= \begin{cases} 0 & \text{if } \pi \in \Pi(\mu,\nu) \\ +\infty & \text{else.} \end{cases} \end{aligned}$$

Hence, the right hand side of (4.7) reads

$$\max_{\pi\in E^*} \left(-\Theta^*(-\pi) - \Xi^*(\pi)\right) = - \min_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} c(x,y) \, \mathrm{d}\pi = - \min_{\pi\in\Pi(\mu,\nu)} \mathbb{K}(\pi).$$

This completes the proof of part 1.

Parts 2 and 3 and more complicated (part 2 takes some work, part 3 is actually quite straghtforward) and are omitted; both parts can be found in [23, pp 28-32]. $\square$

4.4 Existence of Maximisers to the Dual Problem

Section references: Theorem 4.10 is adapted from the special case $X = Y = \mathbb{R}^d$, $c(x,y) = |x-y|^2$ in [23, Theorem 2.9], the other results in this section, Lemmas 4.11 and 4.12 are adapted from [23, Lemma 2.10].

The results of this section rely on similar concepts as the proof of duality, in particular we will need to use c-transforms defined below.

Definition 4.9. *For $\varphi : X \to \overline{\mathbb{R}}$, the c-transforms φ^c, φ^{cc} are defined by*

$$\begin{aligned} \varphi^c &: Y \to \overline{\mathbb{R}}, & \varphi^c(y) &= \inf_{x \in X} \left(c(x,y) - \varphi(x) \right) \\ \varphi^{cc} &: X \to \overline{\mathbb{R}}, & \varphi^{cc}(x) &= \inf_{y \in Y} \left(c(x,y) - \varphi^c(y) \right). \end{aligned}$$

One should compare this to the Legendre-Fenchel transform defined in the previous section; in particular if $Y = X^*$ and $c(x,y) = \langle y, x \rangle$ then $(-\varphi)^c(-y) = \varphi^*(y)$.

The objective of this section is to prove the existence of a maximiser to the dual problem. We state the theorem before giving a preliminary result followed by the proof of the theorem.

Theorem 4.10. *Let $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, where X and Y are Polish, and $c : X \times Y \to [0,\infty)$. Assume that there exists $c_X \in L^1(\mu)$, $c_Y \in L^1(\nu)$ such that $c(x,y) \leq c_X(x) + c_Y(y)$ for μ-almost every $x \in X$ and ν-almost every $y \in Y$. Then there exists $(\varphi, \psi) \in \Phi_c$ such that*

$$\sup_{\Phi_c} \mathbb{J} = \mathbb{J}(\varphi, \psi).$$

Furthermore we can choose $(\varphi, \psi) = (\eta, \eta^c)$ for some $\eta \in L^1(\mu)$ ***with*** *$\eta = \eta^{cc}$.*

The conditions $\sigma_X \in L^1(\mu)$ and $\sigma_Y \in L^1(\nu)$ are effectively moment conditions on μ and ν. In particular, if $c(x,y) = |x-y|^p$ then $c(x,y) \leq C(|x|^p + |y|^p)$ and the requirement that $\sigma_X \in L^1(\mu)$ and $\sigma_Y \in L^1(\nu)$ is exactly the condition that μ, ν have finite p^{th} moments.

We first give a result which implies we only need to consider c-transform pairs.

Lemma 4.11. *Let $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, where X and Y are Polish. For any $a \in \mathbb{R}$ and $(\tilde{\varphi}, \tilde{\psi}) \in \Phi_c$ we have $(\varphi, \psi) = (\tilde{\varphi}^{cc} - a, \tilde{\varphi}^c + a)$ satisfies $\mathbb{J}(\varphi, \psi) \geq \mathbb{J}(\tilde{\varphi}, \tilde{\psi})$ and $\varphi(x) + \psi(y) \leq c(x,y)$ for μ-almost every $x \in X$ and ν-almost every $y \in Y$.*

Furthermore, if $\mathbb{J}(\tilde{\varphi}, \tilde{\psi}) > -\infty$ and there exists $c_X \in L^1(\mu)$ and $c_Y \in L^1(\nu)$ such that $\varphi \leq c_X$ and $\psi \leq c_Y$, then $(\varphi, \psi) \in \Phi_c$.

Proof. Clearly $\mathbb{J}(\varphi - a, \psi + a) = \mathbb{J}(\varphi, \psi)$ for all $a \in \mathbb{R}$, $\varphi \in L^1(\mu)$ and $\psi \in L^1(\nu)$, so it is enough to show that $\varphi = \tilde{\varphi}^{cc} \geq \tilde{\varphi}$, $\psi = \tilde{\varphi}^c \geq \tilde{\psi}$, $\varphi(x) + \psi(y) \leq c(x,y)$.

Note that

$$\psi(y) = \inf_{x \in X} \left(c(x,y) - \tilde{\varphi}(x) \right) \geq \tilde{\psi}(y)$$

since $\tilde{\varphi}(x) + \tilde{\psi}(y) \leq c(x,y)$, and

$$\varphi(x) = \inf_{y \in Y} \sup_{z \in X} \left(c(x,y) - c(z,y) + \tilde{\varphi}(z) \right) \geq \tilde{\varphi}(x)$$

by choosing $z = x$.

We easily see that

$$\varphi(x) + \psi(y) = \inf_{z\in Y} \left(c(x,z) - \tilde{\varphi}^c(z) + \tilde{\varphi}^c(y)\right) \leq c(x,y)$$

by choosing $z = y$.

For the furthermore part of the lemma it is left to show integrability of φ, ψ. Let

$$M := \int_X c_X(x)\,\mathrm{d}\mu(x) + \int_Y c_Y(y)\,\mathrm{d}\nu(y) < \infty. \tag{4.8}$$

Note that

$$\int_X \varphi(x) - c_X(x)\,\mathrm{d}\mu(x) + \int_Y \psi(y) - c_Y(y)\,\mathrm{d}\nu(y) = \mathbb{J}(\varphi,\psi) - M \geq \mathbb{J}(\tilde{\varphi},\tilde{\psi}) - M$$

and since $\varphi - c_X \leq 0$, $\psi - c_Y \leq 0$ then both integrals on the left hand side are negative. In particular

$$\begin{aligned}\|\varphi - c_X\|_{L^1(\mu)} + \|\psi - c_Y\|_{L^1(\nu)} &= -\int_X \varphi(x) - c_X(x)\,\mathrm{d}\mu(x) - \int_Y \psi(y) - c_Y(y)\,\mathrm{d}\nu(y) \\ &\leq M - \mathbb{J}(\tilde{\varphi},\tilde{\psi}).\end{aligned}$$

Hence $\varphi - c_X \in L^1(\mu)$, $\psi - c_Y \in L^1(\nu)$ from which it follows $\varphi \in L^1(\mu)$, $\psi \in L^1(\nu)$. □

The next result gives an upper bound on maximising sequences.

Lemma 4.12. *Let $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, , where X and Y are Polish, and $c : X \times Y \to \mathbb{R}$. Assume that $c(x,y) \leq c_X(x) + c_Y(y)$ where $c_X \in L^1(\mu)$ and $c_Y \in L^1(\nu)$. Then there exists a sequence $(\varphi_k,\psi_k) \in \Phi_c$ such that $\mathbb{J}(\varphi_k,\psi_k) \to \sup_{\Phi_c}\mathbb{J}$ and satisfying the bounds*

$$\begin{aligned}\varphi_k(x) &\leq c_X(x) \quad \forall x \in X,\, \forall k \in \mathbb{N} \\ \psi_k(y) &\leq c_Y(y) \quad \forall y \in Y,\, \forall k \in \mathbb{N}.\end{aligned}$$

Proof. Let $(\tilde{\varphi}_k,\tilde{\psi}_k) \in \Phi_c$ be a maximising sequence. Notice that since $0 \leq \sup_{\Phi_c}\mathbb{J} \leq \inf_{\Pi(\mu,\nu)}\mathbb{K} \leq M < \infty$ that $\tilde{\varphi}_k, \tilde{\psi}_k$ must be proper functions (in fact not equal to $\pm\infty$ anywhere). Let $(\varphi_k,\psi_k) = (\tilde{\varphi}_k^{cc} - a_k, \tilde{\varphi}_k^c + a_k)$ where we will choose

$$a_k = \inf_{y\in Y}\left(c_Y(y) - \tilde{\varphi}_k^c(y)\right).$$

By Lemma 4.11 $(\varphi_k,\psi_k) \in \Phi_c$ and (φ_k,ψ_k) is a maximising sequence once we have shown that $\varphi_k \leq c_X$ and $\psi_k \leq c_Y$.

We start by showing $a_k \in (-\infty,+\infty)$. Since $(\tilde{\varphi}_k,\tilde{\psi}_k) \in \Phi_c$ then $\tilde{\varphi}_k(x) \leq c(x,y) - \tilde{\psi}_k(y)$ for all $y \in Y$. Hence there exists $y_0 \in Y$ and $b_0 \in \mathbb{R}$ (possibly depending on k) such that $\tilde{\varphi}_k(x) \leq c(x,y_0) + b_0$. Then,

$$\tilde{\varphi}_k^c(y_0) = \inf_{x\in X}\left(c(x,y_0) - \tilde{\varphi}_k(x)\right) \geq -b_0.$$

Hence, $a_k \leq c_Y(y_0) - \varphi_k^c(y_0) \leq c_Y(y_0) + b_0 < \infty$. We also have

$$c_Y(y) - \tilde{\varphi}_k^c(y) = \sup_{x\in X} \left(c_Y(y) - c(x,y) + \tilde{\varphi}_k(x)\right) \geq \sup_{x\in X} \left(\tilde{\varphi}_k(x) - c_X(x)\right) \geq \tilde{\varphi}_k(x_0) - c_X(x_0)$$

for any $x_0 \in X$. Hence, $a_k \geq \tilde{\varphi}(x_0) - c_X(x_0)$ which is greater than $-\infty$. We have shown that $a_k \in (-\infty, +\infty)$ and the pair (φ_k, ψ_k) are well defined.

Clearly $\psi_k(y) = \tilde{\varphi}_k^c(y) + a_k \leq c_Y(y)$. And,

$$\varphi_k(x) - c_X(x) = \inf_{y\in Y} \left(c(x,y) - \tilde{\varphi}_k^c(y) - a_k - c_X(x)\right) \leq \inf_{y\in Y} \left(c_Y(y) - \tilde{\varphi}_k^c(y) - a_k\right) = 0.$$

So, (φ_k, ψ_k) satisfy the bounds stated in the lemma. □

With the help of the preceding lemma we can prove Theorem 4.10.

Proof of Theorem 4.10. Note that

$$\sup_{(\varphi,\psi)\in\Phi_c} \mathbb{J}(\varphi,\psi) \leq \inf_{\pi\in\Pi(\mu,\nu)} \mathbb{K}(\pi) \leq M < \infty$$

by Lemma 4.7. Let $(\varphi_k, \psi_k) \in \Phi_c$ be a maximising sequence as in Lemma 4.12. Define $\varphi_k^{(\ell)}$, $\psi_k^{(\ell)}$ by

$$\begin{aligned} \varphi_k^{(\ell)}(x) &= \max\{\varphi_k(x) - c_X(x), -\ell\} + c_X(x) \\ \psi_k^{(\ell)}(y) &= \max\{\psi_k(y) - c_Y(y), -\ell\} + c_Y(y). \end{aligned}$$

Note that $\varphi_k \leq \varphi_k^{(\ell)}$, $\psi_k \leq \psi_k^{(\ell)}$,

$$\begin{aligned} -\ell \leq \varphi_k^{(\ell)}(x) - c_X(x) &\leq 0 \qquad \forall x \in X,\ \forall k \in \mathbb{N},\ \forall \ell \in \mathbb{N} \\ -\ell \leq \psi_k^{(\ell)}(y) - c_Y(y) &\leq 0 \qquad \forall y \in Y,\ \forall k \in \mathbb{N},\ \forall \ell \in \mathbb{N} \\ \varphi_k^{(1)} \geq \varphi_k^{(2)} &\geq \dots \\ \psi_k^{(1)} \geq \psi_k^{(2)} &\geq \dots \end{aligned}$$

and

$$\begin{aligned} \varphi_k^{(\ell)}(x) + \psi_k^{(\ell)}(y) &\leq \max\left\{\varphi_k(x) - c_X(x) + \psi_k(y) - c_Y(y), -\ell\right\} + c_X(x) + c_Y(y) \\ &\leq \max\left\{c(x,y) - c_X(x) - c_Y(y), -\ell\right\} + c_X(x) + c_Y(y). \end{aligned} \tag{4.9}$$

For each ℓ the sequence $\varphi_k^{(\ell)} - c_X$ is bounded in L^∞ so $\overline{\{\varphi_k^{(\ell)} - c_X\}_{k\in\mathbb{N}}}$ is weakly compact in $L^p(\mu)$ for any $p \in (1,\infty)$ (for reflexive Banach spaces boundedness plus closure is equivalent to weak compactness). Let's choose $p = 2$ then, after extracting a subsequence, we have that $\varphi_k^{(\ell)} - c_X \overset{L^2(\mu)}{\rightharpoonup} \varphi^{(\ell)} - c_X \in L^1(\mu)$ (since $L^2(\mu) \subset L^1(\mu)$) for some $\varphi^{(\ell)} \in L^1(\mu)$. Furthermore, as weak convergence in $L^2(\mu)$ implies weak convergence in $L^1(\mu)$ (easy exercise) then we have that $\varphi_k^{(\ell)} - c_X \overset{L^1(\mu)}{\rightharpoonup} \varphi^{(\ell)} - c_X$, and therefore $\varphi_k^{(\ell)} \overset{L^2(\mu)}{\rightharpoonup} \varphi^{(\ell)}$. By a diagonalisation argument we can assume that $\varphi_k^{(\ell)} \overset{L^1(\mu)}{\rightharpoonup} \varphi^{(\ell)}$ for all $\ell \in \mathbb{N}$. We can apply the same argument to $\psi_k^{(\ell)}$ to imply the existence of weak limits $\psi^{(\ell)} \in L^1(\nu)$. Since weak convergence preserves the ordering we have that

$$c_X \geq \varphi^{(1)} \geq \varphi^{(2)} \geq \dots$$
$$c_Y \geq \psi^{(1)} \geq \psi^{(2)} \geq \dots$$

Since $\varphi^{(\ell)}, \psi^{(\ell)}$ are bounded above by an L^1 function and monotonically decreasing we can apply the Monotone Convergence Theorem to infer

$$\lim_{\ell\to\infty} \int_X \varphi^{(\ell)}(x)\,\mathrm{d}\mu(x) = \int_X \varphi^\dagger(x)\,\mathrm{d}\mu(x)$$
$$\lim_{\ell\to\infty} \int_Y \psi^{(\ell)}(y)\,\mathrm{d}\nu(y) = \int_Y \psi^\dagger(y)\,\mathrm{d}\nu(y)$$

where $\varphi^\dagger$, $\psi^\dagger$ are the pointwise limits of $\varphi^{(\ell)}$, $\psi^{(\ell)}$:

$$\varphi^\dagger(x) = \lim_{\ell\to\infty} \varphi^{(\ell)}(x), \qquad \psi^\dagger(y) = \lim_{\ell\to\infty} \psi^{(\ell)}(y).$$

The functions $(\varphi^\dagger, \psi^\dagger)$ are our candidate maximisers. We are required to show that $(\varphi^\dagger, \psi^\dagger) \in \Phi_c$ and $\mathbb{J}(\varphi, \psi) \leq \mathbb{J}(\varphi^\dagger, \psi^\dagger)$ for all $(\varphi, \psi) \in \Phi_c$.

Since $\sup_{\Phi_c} \mathbb{J} = \lim_{k\to\infty} \mathbb{J}(\varphi_k, \psi_k) \leq \lim_{k\to\infty} \mathbb{J}(\varphi_k^{(\ell)}, \psi_k^{(\ell)}) = \mathbb{J}(\varphi^{(\ell)}, \psi^{(\ell)})$ for any $\ell \in \mathbb{N}$ then

$$\mathbb{J}(\varphi^\dagger, \psi^\dagger) = \lim_{\ell\to\infty} \mathbb{J}(\varphi^{(\ell)}, \psi^{(\ell)}) \geq \sup_{\Phi_c} \mathbb{J}.$$

Hence $(\varphi^\dagger, \psi^\dagger)$ maximises $\mathbb{J}$.

It follows from taking $\ell \to \infty$ in (4.9) that $\varphi^\dagger(x) + \psi^\dagger(y) \leq c(x, y)$. Now integrability follows from

$$0 \geq \int_X \varphi^\dagger(x) - c_X(x)\,\mathrm{d}\mu(x) + \int_Y \psi^\dagger(y) - c_Y(y)\,\mathrm{d}\nu(y) \geq \sup_{\Phi_c} \mathbb{J} - M$$

where M is defined by (4.8). In particular, since $\varphi^\dagger - c_X \leq 0$, $\psi^\dagger - c_Y \leq 0$ then it follows that both integrals are finite and $\varphi^\dagger - c_X \in L^1(\mu)$, $\psi^\dagger - c_Y \in L^1(\nu)$. Hence $\varphi^\dagger \in L^1(\mu)$ and $\psi^\dagger \in L^1(\nu)$.

For the furthermore part of the theorem we use the double c-transform trick as in the proof of Lemma 4.12. For any $a \in \mathbb{R}$ we have, by Lemma 4.11,

$$\mathbb{J}(\varphi^\dagger, \psi^\dagger) \leq \mathbb{J}((\varphi^\dagger)^{cc} - a, (\varphi^\dagger)^c + a) = \mathbb{J}((\varphi^\dagger)^{cc}, (\varphi^\dagger)^c).$$

We only have to show $((\varphi^\dagger)^{cc}, (\varphi^\dagger)^c) \in L^1(\mu) \times L^1(\nu)$. Let $a = \inf_{y\in Y} \big(c_Y(y) - (\varphi^\dagger)^c(y)\big)$ then $a \in \mathbb{R}$ for the same reasons that $a_k \in \mathbb{R}$ in the proof of Lemma 4.12. Clearly $(\varphi^\dagger)^c(y) + a \leq c_Y(y)$ and

$$(\varphi^\dagger)^{cc}(x) = \inf_{y\in Y} \big(c(x, y) - (\varphi^\dagger)^c(y) - a\big) \leq \inf_{y\in Y} \big(c_X(x) + c_Y(y) - (\varphi^\dagger)^c(y) - a\big) \leq c_X(x).$$

Hence, $((\varphi^\dagger)^{cc} - a, (\varphi^\dagger)^c + a) \in L^1(\mu) \times L^1(\nu)$ by Lemma 4.11. Trivially $((\varphi^\dagger)^{cc}, (\varphi^\dagger)^c) \in L^1(\mu) \times L^1(\nu)$. **Applying the same trick again we have that $((\varphi^\dagger)^{cccc}, (\varphi^\dagger)^{ccc}) \in \Phi_c$ is a maximiser of $\mathbb{J}$ over Φ_c. It is left as an exercise to show that $(\varphi^\dagger)^{cccc} = (\varphi^\dagger)^{cc}$ and $(\varphi^\dagger)^{ccc} = (\varphi^\dagger)^c$ so that choosing $\eta = (\varphi^\dagger)^{cc}$ completes the proof.** $\square$

4.5 Kantorovich-Rubinstein Theorem

Section references: The statement and proof of the Kantorovich-Rubinstein Theorem (Theorem 4.13) comes form [23, Theorem 1.14].

When the cost function is a metric and $X = Y$ there is more structure to Kantorovich duality, in particular the optimal transport cost can be written as the supremum of $\int_X f \,\mathrm{d}(\mu - \nu)$ where f is a Lipschitz function with Lipschitz constant 1.

Theorem 4.13. Kantorovich-Rubinstein Theorem. *Let $X = Y$ be a Polish space and $c : X \times X \to [0, +\infty]$ be a lower semi-continuous metric on X. Define*

$$\|f\|_{\mathrm{Lip}} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{c(x, y)}$$

and $\mathbb{K}$ as in Definition 2.3. Let $\mu, \nu \in \mathcal{P}(X)$ and assume $c(\cdot, 0) \in L^1(\mu)$ and $c(0, \cdot) \in L^1(\nu)$. Then,

$$\min_{\pi \in \Pi(\mu,\nu)} \mathbb{K}(\pi) = \sup \left\{ \int_X f \,\mathrm{d}(\mu - \nu) \;:\; f \in L^1(|\mu - \nu|), \|f\|_{\mathrm{Lip}} \leq 1 \right\}.$$

Proof. We assume that c is bounded, i.e. $\sup_{x,y \in X} c(x, y) < +\infty$. This assumption can be lifted (but we omit the proof for this part) by considering a sequence of cost functions $c_n = \frac{c}{1+\frac{c}{n}}$ which are bounded and taking the limit $n \to \infty$. The argument for the latter follows approximately the same argument as in steps 2-3 of Lemma 4.8 (which was also omitted).

Let $\sup_{x,y \in X} c(x, y) \leq C$ then for any f with $\|f\|_{\mathrm{Lip}} \leq 1$ we have $|f(x) - f(y)| \leq C$ and in particular f is bounded. It follows that $f \in L^1(|\mu - \nu|)$. By Theorem 4.1 it is enough to check that

$$\sup_{(\varphi,\psi) \in \Phi_c} \mathbb{J}(\varphi, \psi) = \sup \left\{ \int_X f \,\mathrm{d}(\mu - \nu) \;:\; \|f\|_{\mathrm{Lip}} \leq 1 \right\}$$

(where $\mathbb{J}$, Φ_c are defined in Theorem 4.1).

We claim that for all $\eta \in L^1(\mu)$ (a) $\|\eta^c\|_{\mathrm{Lip}} \leq 1$ and (b) $\eta^{cc} = -\eta^c$, and for any f with $\|f\|_{\mathrm{Lip}} \leq 1$ (c) $(-f, f) \in \Phi_c$. Assuming (a-c), and by Theorem 4.10, we have

$$\sup_{(\varphi,\psi) \in \Phi_c} \mathbb{J}(\varphi, \psi) = \sup_{\eta \in L^1(\mu)} \mathbb{J}(\eta^{cc}, \eta^c) = \sup_{\eta \in L^1(\mu)} \mathbb{J}(-\eta^c, \eta^c) \leq \sup_{\|f\|_{\mathrm{Lip}} \leq 1} \mathbb{J}(f, -f) \leq \sup_{(\varphi,\psi) \in \Phi_c} \mathbb{J}(\varphi, \psi).$$

Hence $\sup_{(\varphi,\psi) \in \Phi_c} \mathbb{J}(\varphi, \psi) = \sup_{\|f\|_{\mathrm{Lip}} \leq 1} \mathbb{J}(f, -f)$. As $\mathbb{J}(f, -f) = \int_X f \,\mathrm{d}(\mu - \nu)$ we have proved the theorem assuming (a-c).

To check (a) we compute

$$\begin{aligned}
\eta^c(x) - \eta^c(y) &= \inf_{w \in X} \sup_{z \in X} \left(c(x, w) - c(y, z) + \eta(z) - \eta(w) \right) \\
&\leq \sup_{z \in X} \left(c(x, z) - c(y, z) \right) \\
&\leq c(x, y)
\end{aligned}$$

with the last line following from the triangle inequality. By symmetry $|\eta^c(x) - \eta^c(y)| \leq c(x, y)$ which proves $\|\eta^c\|_{\text{Lip}} \leq 1$.

By choosing $x = y$ we have

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

By the 1-Lipschitz property (part (a))

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

Hence $\eta^{cc} = -\eta^c$ which proves (b).

Finally, we have that for any f with $\|f\|_{\text{Lip}} \leq 1$ that $f \in L^1(\mu)$ and $f \in L^1(\nu)$ (since f is bounded). Now $f(x) - f(y) \leq f(y) + c(x, y) - f(y) = c(x, y)$ by the 1-Lipschitz property and therefore $(f, -f) \in \Phi_c$ as required to show (c). $\square$

Chapter 5

Semi-Discrete Optimal Transport

Section references: The proof of Lemma 5.2 is adapted from [15, Theorem 2] who consider more general costs.

In this chapter we consider another special case of the optimal transport (see also the spacial cases considered in Chapter 3). The case we consider here is when one of the measures is discrete, in particular we assume

$$\nu = \sum_{j=1}^n m_j \delta_{y_j}$$

for some set of points $\{y_j\}_{j=1}^n \subset \mathbb{R}^d$ and weights $\{m_j\}_{j=1}^n \subset [0,1]$ satisfying $\sum_{j=1}^n m_j = 1$. The cost function is $c(x,y) = |x-y|^2$ and $X = Y = \mathbb{R}^d$. We first define the Laguerre diagram (also known as the power diagram).

Definition 5.1. *Given a set of points $\{y_j\}_{j=1}^n$ and weights $\{w_j\}_{j=1}^n$ the Laguerre diagram is the collection of sets*

$$L_j = \left\{ x \in \mathbb{R}^d \,:\, |x-y_j|^2 - w_j < |x-y_i|^2 - w_i \,\forall i \right\}$$

for $j = 1, \dots, n$.

If $w_j = 0$ then the Laguerre diagram are the Voronoi cells. Note also that each L_j is open.

The objective of this chapter is to show that for any $\mu \in \mathcal{P}(\mathbb{R}^d)$ with a density (wrt the Lebesgue measure) there exists an optimal map $T : \mathbb{R}^d \to \mathbb{R}^d$ that defines a Laguerre diagram (upto the boundary of the cells L_j). Moreover we can find the weights $\{w_j\}_{j=1}^n$ as the solutions of the concave variational problem:

$$\text{maximise } g(W) \text{ over } W = (w_1, \dots, w_n) \in \mathbb{R}^n$$

where g is defined in (5.1) and ρ is the density of μ.

Lemma 5.2. *Let $\rho \in L^1(\mathbb{R}^d)$ be non-negative and integrate to unity, $\{m_j\}_{j=1}^n \subset [0,1]$ satisfy $\sum_{j=1}^n m_j = 1$, and $\{y_j\}_{j=1}^n \subset \mathbb{R}^d$. Then $g : \mathbb{R}^n \to \mathbb{R}$ defined by*

$$g(w_1, \dots, w_n) = \int_{\mathbb{R}^d} \inf_j \left[|x-y_j|^2 - w_j \right] \rho(x) \, \mathrm{d}x + \sum_{j=1}^n w_j m_j \tag{5.1}$$

is concave.

Proof. Let $\gamma : \mathbb{R}^d \to \{1, \dots, n\}$ be any partitioning of $\mathbb{R}^d$ and define

$$G(\gamma, W) = \int_{\mathbb{R}^d} \left(|x - y_{\gamma(x)}|^2 - w_{\gamma(x)} \right) \rho(x) \, \mathrm{d}x + \sum_{j=1}^n w_j m_j.$$

Let $\gamma^{-1}(j) = \{x \in \mathbb{R}^d \, : \, \gamma(x) = j\}$ then

$$\begin{aligned} G(\gamma, W) &= \sum_{j=1}^n \left(\int_{\gamma^{-1}(j)} \left(|x - y_j|^2 - w_j \right) \rho(x) \, \mathrm{d}x + w_j m_j \right) \\ &= \sum_{j=1}^n \int_{\gamma^{-1}(j)} |x - y_j|^2 \rho(x) \, \mathrm{d}x + \sum_{j=1}^n w_j \left(m - \int_{\gamma^{-1}(j)} \rho(x) \, \mathrm{d}x \right). \end{aligned}$$

For each γ, G is an affine function of W (and therefore $W \mapsto G(\gamma, W)$ is concave). Since

$$g(W) = \inf_\gamma G(\gamma, W)$$

it follows that g is also concave. □

The following result will be useful in the sequel.

Lemma 5.3. *Define $G : \mathbb{R}^n \to \mathbb{R}$ by* (5.1) *where $\rho \in L^1(\mathbb{R}^d)$, $\{y_j\}_{j=1}^n \subset \mathbb{R}^d$ and $\{m_j\}_{j=1}^n \subset \mathbb{R}$. Let $L_1(W), \dots, L_n(W)$ are the Laguerre diagram corresponding to $W = (w_1, \dots, w_n)$ and $\{y_i\}_{i=1}^n$. Then,*

$$\frac{\partial g}{\partial w_i}(W) = -\int_{L_i(W)} \rho(x) \, \mathrm{d}x + m_i.$$

Sketch Proof. Let $\alpha_j(x, W) = \chi_{L_j(W)}(x)(|x - y_j|^2 - w_j)\rho(x)$ so that

$$g(W) = \sum_{j=1}^n \left(\int_{\mathbb{R}^d} \alpha_j(x, W) \, \mathrm{d}x + w_j m_j \right).$$

For $x \in L_j(W)$ we haven $\chi_{L_j(W + te_i)}(x) = \chi_{L_j(W)}(x)$ for $t > 0$ sufficiently small where e_i is the i^{th} unit base vector in $\mathbb{R}^n$. Moreover,

$$\frac{1}{t} \left(\alpha_j(x, W + te_i) - \alpha_j(x, W) \right) = -\chi_{L_j(W)}(x) \delta_i(j) \rho(x)$$

for $t > 0$ sufficiently small. Hence,

$$\begin{aligned} \frac{\partial g}{\partial w_i}(W) &= \lim_{t \to 0^+} \frac{1}{t} \left(g(W + te_i) - g(W) \right) \\ &= \sum_{j=1}^n \lim_{t \to 0^+} \left(\int_{\mathbb{R}^d} \frac{1}{t} \left(\alpha_j(x, W + te_i) - \alpha_j(x, W) \right) \, \mathrm{d}x + (w_j + \delta_i(j)t) m_j - w_j m_j \right) \\ &= \sum_{j=1}^n \left(-\int_{\mathbb{R}^d} \chi_{L_j(W)}(x) \delta_i(j) \rho(x) \, \mathrm{d}x + \delta_i(j) m_j \right) \\ &= -\int_{\mathbb{R}^d} \chi_{L_i(W)}(x) \rho(x) \, \mathrm{d}x + m_i. \end{aligned}$$

To make the argument rigorous we would need to justify passing the limit $t \to 0^+$ through the integral. $\square$

We proceed to the main theorem of the section.

Theorem 5.4. *Let $\{y_j\}_{j=1}^n \subset \mathbb{R}^d$, $\{m_j\}_{j=1}^n \subset [0,1]$ with $\sum_{j=1}^n m_j = 1$, $\mu \in \mathcal{P}(\mathbb{R}^d)$ and define $\nu = \sum_{j=1}^n m_j \delta_{y_j}$. Assume $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ has density ρ with respect to the Lebesgue measure on $\mathbb{R}^d$. Define g by* (5.1) *and let $(w_1, \dots, w_n)$ maximise g. Let $\{L_j\}_{j=1}^n$ be the Laguerre diagram generated by $\{(y_j, w_j)\}_{j=1}^n$. Define $T(x) = y_j$ if $x \in L_j$ (which defines T μ-almost everywhere), $\psi(y_j) = w_j$ and $\varphi(x) = \inf_j (|x - y_j|^2 - w_j)$. Then*

1. *T is the solution of Monge's optimal transport problem with cost $c(x,y) = |x-y|^2$;*
2. *(φ, ψ) are an optimal pair for Kantorovich's dual problem with cost $c(x,y) = |x-y|^2$.*

Proof. We claim:

(a) $\varphi \in L^1(\mu)$;

(b) $T_\# \mu = \nu$; and

(c) $\int_{L_j} \rho(x) \, \mathrm{d}x = m_j$.

Assuming (a-c) we have

$$\varphi(x) + \psi(y_i) = \inf_j \left(|x - y_j|^2 - w_j + w_i \right) \le |x - y_i|^2$$

so $(\varphi, \psi) \in \Phi_c$ where $c(x,y) = |x-y|^2$ and Φ_c is defined by (4.2). By (2.1), Theorem 4.1, (a) and (b) we have

$$\mathbb{M}(T) \ge \inf_{\bar{T} : \bar{T}_\# \mu = \nu} \mathbb{M}(\bar{T}) \ge \min_{\pi \in \Pi(\mu,\nu)} \mathbb{K}(\pi) = \sup_{(\bar{\varphi}, \bar{\psi}) \in \Phi_c} \mathbb{J}(\bar{\varphi}, \bar{\psi}) \ge \mathbb{J}(\varphi, \psi). \tag{5.2}$$

Now, using (c),

$$\begin{aligned}
\mathbb{J}(\varphi, \psi) &= \int_{\mathbb{R}^d} \varphi(x) \rho(x) \, \mathrm{d}x + \sum_{j=1}^n m_j \psi(y_j) \\
&= \sum_{j=1}^n \left(\int_{L_j} \varphi(x) \rho(x) \, \mathrm{d}x + m_j \psi(y_j) \right) \\
&= \sum_{j=1}^n \left(\int_{L_j} \left(|x - y_j|^2 - w_j \right) \rho(x) \, \mathrm{d}x + m_j w_j \right) \\
&= \sum_{j=1}^n \int_{L_j} |x - y_j|^2 \rho(x) \, \mathrm{d}x \\
&= \int_{\mathbb{R}^d} |x - T(x)|^2 \rho(x) \, \mathrm{d}x \\
&= \mathbb{M}(T).
\end{aligned}$$

Hence all the inequalities in (5.2) must be equalities. In particular, (φ, ψ) are maximisers of $\mathbb{J}$ and T is a minimiser of $\mathbb{M}$ (with the appropriate constraints). We are left to show (a-c).

For (a) we have

$$-\sup_j w_j \leq \varphi(x) \leq |x - y_1|^2 - w_1$$

and therefore $|\varphi(x)| \leq 2|x|^2 + C$ where $C = 2|y_1|^2 - w_1 + \sup_j w_j$. Hence $\|\varphi\|_{L^1(\mu)} \leq 2\int_{\mathbb{R}^d} |x|^2 \, \mathrm{d}\mu(x) + C < +\infty$.

To show (b) we consider any $i = 1, \dots, n$ and by (c) we have

$$\mu(T^{-1}(y_i)) = \mu(\{x : T(x) = y_i\}) = \mu(L_i) = m_i.$$

Finally (c) follows from Lemma 5.3 since any maximiser of g satisfies $\frac{\partial g}{\partial w_j}(W) = 0$ for all $j = 1, \dots, n$. $\square$

Chapter 6

Existence and Characterisation of Transport Maps

Our aim is to characterise the optimal transport plans that arise as a minimiser to the Kantorovich optimal transportation problem and show sufficient conditions for the existence of optimal transport maps. In this chapter we will restrict ourselves to the cost function $c(x, y) = \frac{1}{2}|x - y|^2$. Generalisations to other cost functions are possible but with an additional notational burden.[1] The second restriction we make is to assume that X and Y are subsets of $\mathbb{R}^d$.

The chapter is structured as follows. We first state our objectives and in particular the results we will prove, then give motivating explanations. We will require some results and definitions from convex analysis which we give in Section 6.2 before finally proving the main theorems from the first section.

6.1 Knott-Smith Optimality and Brenier's Theorem

Section references: Theorem 6.1, Theorem 6.2 and Corollary 6.3 form [23, Theorem 2.12].

We will give (1) a characterisation of optimal transport plans and (2) sufficient conditions for the existence of optimal transport maps (and when $\min_{\pi\in\Pi(\mu,\nu)} \mathbb{K}(\pi) = \min_{T_\# \mu=\nu} \mathbb{M}(T)$).

We will restate the theorem in Section 6.3 with a change of notation (more precisely we look at the equivalent problems $\sup_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} x \cdot y \, \mathrm{d}\pi(x, y)$ and $\inf_{\tilde{\Phi}} \mathbb{J}$ where $\tilde{\Phi}$ is defined in below.

The subdifferential $\partial\varphi$ of a convex function φ is defined by

$$\partial\varphi(x) := \left\{ y \,:\, \varphi(z) \geq \varphi(x) + y \cdot (z - x) \,\forall z \in \mathbb{R}^d \right\}.$$

We will review some convex analysis in Section 6.2 which will inform the definition. For now it is enough to know that the subdifferential is a generalisation of the differential which always exists for lower semi-continuous convex functions. Note that the subdifferential is in general a set however if φ is differentiable at x then $\partial\varphi(x) = \{\nabla\varphi(x)\}$.

[1] For example it is possible to show the existence of optimal transport maps for strictly convex and superlinear cost functions, and characterise them in terms of c-superdifferentials of c-convex functions.

Theorem 6.1. Knott-Smith Optimality Criterion. *Let $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ with $X, Y \subset \mathbb{R}^d$ and assume that μ, ν both have finite second moments. Define $\mathbb{J}$ by* (4.1)*, $c(x,y) = \frac{1}{2}|x-y|^2$, and $\tilde{\Phi}$ by*

$$\tilde{\Phi} = \left\{ (\tilde{\varphi}, \tilde{\psi}) \in L^1(\mu) \times L^1(\nu) \, : \, \tilde{\varphi}(x) + \tilde{\psi}(y) \geq x \cdot y \right\}. \tag{6.1}$$

1. *Let $\pi^\dagger \in \Pi(\mu, \nu)$ be a minimizer of Kantorovich's optimal transport problem and let $(\tilde{\varphi}^\dagger, (\tilde{\varphi}^\dagger)^*) \in \tilde{\Phi}$ be a minimiser of the problem $\inf_{(\tilde{\varphi},\tilde{\psi})\in\tilde{\Phi}} \mathbb{J}(\tilde{\varphi}, \tilde{\psi})$. Then $\mathrm{supp}(\pi^\dagger) \subseteq \mathrm{Gra}(\partial\tilde{\varphi}^\dagger)$, or equivalently $y \in \partial\tilde{\varphi}^\dagger(x)$ for $\pi^\dagger$-almost every (x,y).*

2. *Let $\pi^\dagger \in \Pi(\mu, \nu)$ and suppose there exists an $L^1(\mu)$, convex, lower semi-continuous function $\tilde{\varphi}^\dagger$ such that $\mathrm{supp}(\pi^\dagger) \subseteq \mathrm{Gra}(\partial\tilde{\varphi}^\dagger)$. Then $\pi^\dagger$ is a minimizer of Kantorovich's optimal transport problem and $(\tilde{\varphi}^\dagger, (\tilde{\varphi}^\dagger)^*)$ is a minimiser of the problem $\inf_{(\tilde{\varphi},\tilde{\psi})\in\tilde{\Phi}} \mathbb{J}(\tilde{\varphi}, \tilde{\psi})$.*

Why do we expect the optimal plan to have support in the graph of the subdifferential of a convex function? Let us first consider the 1D case and a map $T_{\#}\mu = \nu$. We should expect any map that is 'optimal' to be order preserving, i.e. if $x_1 \leq x_2$ then $T(x_1) \leq T(x_2)$. This is equivalent to saying that T is non-decreasing.

Maps rule out splitting since each $x \mapsto T(x)$. However if we let T set valued (i.e. we are considering plans instead of maps) then the increasing property in some sense should still hold. Let

$$\Gamma = \{(x,y) \, : \, x \in X \text{ and } y \in T(x)\}.$$

We can write the increasing property as: for any $(x_1, y_1), (x_2, y_2) \in \Gamma$ with $x_1 \leq x_2$ then $y_1 \leq y_2$.

In convex analysis this property is called cyclical monotonicity. It can be shown that any cyclically monotone set can be written as the subgradient of a convex function (since any convex function has a non-decreasing derivative). Hence we expect any optimal plan to be supported in the subgradient of a convex function. This turns out to also be true in dimensions greater than one.

The next result specifically gives conditions sufficient for the existence of transport maps.

Theorem 6.2. Brenier's Theorem. *Let $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ with $X, Y \subset \mathbb{R}^d$ and assume that μ, ν both have finite second moments and that μ does not give mass to small sets. Define $\mathbb{J}$ by* (4.1)*, $c(x,y) = \frac{1}{2}|x-y|^2$, and $\tilde{\Phi}$ by* (6.1)*. Then,*

1. *there is a unique solution $\pi^\dagger \in \Pi(\mu, \nu)$ to Kantorovich's optimal transport problem;*

2. *there exists an $L^1(\mu)$, convex, lower semi-continuous function $\tilde{\varphi}^\dagger$ such that*

$$\pi^\dagger = (\mathrm{Id} \times \nabla\tilde{\varphi}^\dagger)_{\#}\mu \qquad \textit{or equivalently} \qquad \mathrm{d}\pi^\dagger(x,y) = \mathrm{d}\mu(x)\delta_{\nabla\tilde{\varphi}^\dagger(x)}(y);$$

3. *$(\tilde{\varphi}^\dagger, (\tilde{\varphi}^\dagger)^*)$ minimise $\mathbb{J}$ over $\tilde{\Phi}$;*

4. *if $\bar{\varphi}$ is any convex function with $(\nabla\bar{\varphi})_{\#}\mu = \nu$ then $\nabla\bar{\varphi} = \nabla\tilde{\varphi}^\dagger$.*

We show later (see Proposition 6.5 that any convex function is differentiable almost everywhere. The fact that the set of non-differentiability is a small set and π gives zero mass to small sets implies we can talk about the derivative of φ on the support of π as though it exists everywhere.

The following corollary is immediate from Brenier's theorem.

Corollary 6.3. *Under the assumptions of Theorem 6.2 $\nabla\tilde{\varphi}$ is the unique solution to the Monge transportation problem:*

$$\frac{1}{2}\int_X |x-\nabla\tilde{\varphi}(x)|^2 \,\mathrm{d}\mu(x) = \frac{1}{2}\inf_{T_{\#}\mu=\nu}\int_X |x-T(x)|^2 \,\mathrm{d}\mu(x).$$

We sketch the proof of the corollary. From the inequality

$$\min_{\pi\in\Pi(\mu,\nu)} \mathbb{K}(\pi) \leq \inf_{T\,:\,T_{\#}\mu=\nu} \mathbb{M}(T)$$

that was argued in Section 2.2 it is enough to show $T^\dagger = \nabla\tilde{\varphi}$ satisfies $\mathbb{M}(T^\dagger) \leq \min_{\Pi(\mu,\nu)}\mathbb{K}$ ($T^\dagger_{\#}\mu=\nu$ is given in Theorem 6.2). Now, let $\pi^\dagger\in\Pi(\mu,\nu)$ be as in Theorem 6.2,

$$\begin{aligned}
\mathbb{M}(T^\dagger) &= \frac{1}{2}\int_X |x-T^\dagger(x)|^2 \,\mathrm{d}\mu(x) \\
&= \frac{1}{2}\int_{X\times Y} |x-T^\dagger(x)|^2 \,\mathrm{d}\pi^\dagger(x,y) \\
&= \frac{1}{2}\int_{X\times Y} |x-y|^2 \,\mathrm{d}\pi^\dagger(x,y) \\
&= \min_{\pi\in\Pi(\mu,\nu)} \mathbb{K}(\pi)
\end{aligned}$$

since $T^\dagger(x)=y$ for π-almost every $(x,y)\in X\times Y$. Uniqueness of $T^\dagger$ follows from uniqueness of $\pi^\dagger$.

6.2 Preliminary Results from Convex Analysis

Section references: These results are a subset of the background in convex analsyis given in [23]. In particular Proposition 6.4 is [23, Proposition 2.4] and Proposition 6.6 is [23, Proposition 2.5].

In this section we drop the tilde notation, i.e. we write φ rather than $\tilde{\varphi}$.

In order to characterise subgradients we will use the convex conjugate defined below. This is essentially a special case of the Legendre-Fenchel transform we defined in Section 4.2. we recall that the Legendre-Fenchel transform (a.k.a. the convex conjugate) is defined as

$$\varphi^*(y) = \sup_{x\in\mathbb{R}^d}\left(x\cdot y - \varphi(x)\right).$$

The following proposition characterises the subdifferential.

Proposition 6.4. *Let φ be a proper, lower semi-continuous, convex function on $\mathbb{R}^d$. Then for all $x, y \in \mathbb{R}^d$*

$$x \cdot y = \varphi(x) + \varphi^*(y) \quad \Leftrightarrow \quad y \in \partial\varphi(x).$$

Proof. Since $\varphi^*(y) \geq x \cdot y - \varphi(x)$ for all x, y we have

$$\begin{aligned} x \cdot y = \varphi(x) + \varphi^*(y) \;&\Leftrightarrow\; x \cdot y \geq \varphi(x) + \varphi^*(y) \\ &\Leftrightarrow\; x \cdot y \geq \varphi(x) + y \cdot z - \varphi(z) \quad \forall z \in \mathbb{R}^d \\ &\Leftrightarrow\; \varphi(z) \geq \varphi(x) + y \cdot (z - x) \quad \forall z \in \mathbb{R}^d \\ &\Leftrightarrow\; y \in \partial\varphi(x) \end{aligned}$$

which proves the proposition. □

In fact if φ is convex then φ is differentiable almost everywhere, hence we have that $\partial\varphi(x) = \{\nabla\varphi(x)\}$ for almost every x.

Proposition 6.5. *If $\varphi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is convex then (1) φ is almost everywhere differentiable on the interior of it's domain, and (2) whenever φ is differentiable $\partial\varphi(x) = \{\nabla\varphi(x)\}$.*

Proof. Let $x \in \mathrm{int}(\mathrm{Dom}(\varphi))$ and δ^* be such that $\overline{B(x, \delta^*)} \subset \mathrm{int}(\mathrm{Dom}(\varphi))$. We show that φ is Lipschitz continuous on $B(x, \delta^*/(2\sqrt{d}))$. Then, by Rademacher's theorem[2], φ is differentiable almost everywhere on $B(x, \delta^*/(2\sqrt{d}))$, and therefore differentiable almost everywhere on $\mathrm{int}(\mathrm{Dom}(\varphi))$. This will complete the proof of (1).

We show φ is Lipschitz on $B(x, \delta^*/(2\sqrt{d}))$ by first showing that φ is uniformly bounded on $\overline{B(x, \delta^*/\sqrt{d})}$. Consider Q a d-dimensional closed cuboid with sides of length $\frac{2\delta^*}{\sqrt{d}}$ and centred at x. Note that $\overline{B(x, \delta^*/\sqrt{d})} \subset Q \subset \overline{B(x, \delta^*)}$ and the vertices $\{x_i\}_{i=1}^{2^d}$ satisfy $\{x_i\}_{i=1}^{2^d} \subset \partial B(x, \delta^*)$. Note also that the set of extreme points of Q is $\{x_i\}_{i=1}^{2^d}$. By the Minkowski-Carathéodory Theorem (Theorem 3.5) for each $y \in B(x, \delta^*/\sqrt{d})$ there exists $\{\lambda_i\}_{i=1}^{2^d} \subset [0, 1]$ with $\sum_{i=1}^{2^d} \lambda_i = 1$ and $y = \sum_{i=1}^{2^d} \lambda_i x_i$. So,

$$\varphi(y) = \varphi(\sum_{i=1}^{2^d} \lambda_i x_i) \leq \sum_{i=1}^{2^d} \lambda_i \varphi(x_i) \leq \max_{i=1,\dots,2^d} |\varphi(x_i)| =: C.$$

Now for $y \in B(x, \delta^*/\sqrt{d})$ we can define $y' = x - (y - x) = 2x - y$ and therefore $y' \in B(x, \delta^*/\sqrt{d})$ and $x = \frac{1}{2}y' + \frac{1}{2}y$. Hence, $\varphi(x) \leq \frac{1}{2}\varphi(y') + \frac{1}{2}\varphi(y)$. In particular,

$$\varphi(y) \geq 2\varphi(x) - \varphi(y') \geq 2\varphi(x) - C.$$

We have shown that

$$2\varphi(x) - C \leq \varphi(y) \leq C \qquad \forall y \in B(x, \delta^*/\sqrt{d}).$$

[2]Rademacher's theorem: if $U \subset \mathbb{R}^d$ is open and $f : U \to \mathbb{R}$ is Lipschitz continuous then f is differentiable almost everywhere on U.

Hence

$$\|\varphi\|_{L^\infty(\overline{B(x,\delta^*/\sqrt{d})})} \leq \max\{C - 2\varphi(x), C\} < \infty.$$

To show φ is Lipschitz on $B(x,\delta^*/(2\sqrt{d}))$ let $x_1, x_2 \in B(x,\delta^*/(2\sqrt{d}))$ ($x_1 \neq x_2$) and take x_3 to be the point of intersection of the line through x_1 and x_2 with $\partial B(x,\delta^*/\sqrt{d})$; there are two possibilities for x_3, we choose the option where x_2 lies between x_1 and x_3. Let $\lambda = \frac{|x_2-x_3|}{|x_1-x_3|} \in (0,1)$. Now,

$$\begin{aligned}
\lambda x_1 + (1-\lambda)x_3 &= \lambda x_2 + \lambda(x_1 - x_2) + (1-\lambda)x_2 + (1-\lambda)(x_3 - x_2) \\
&= x_2 + \frac{|x_3 - x_2|(x_1 - x_2)}{|x_3 - x_1|} + \frac{(|x_3 - x_1| - |x_3 - x_2|)(x_3 - x_2)}{|x_3 - x_1|} \\
&= x_2 + \frac{1}{|x_3 - x_1|}\left(|x_3 - x_2|(x_1 - x_2) + |x_2 - x_1|(x_3 - x_2)\right) \\
&= x_2
\end{aligned}$$

since $\frac{x_2-x_1}{|x_2-x_1|} = \frac{x_3-x_2}{|x_3-x_2|}$ (N.B. one can also define λ as the solution of $\lambda x_1 + (1-\lambda)x_3 = x_2$ then show $\lambda = \frac{|x_2-x_3|}{|x_1-x_3|}$). So by convexity of φ,

$$\begin{aligned}
\varphi(x_2) - \varphi(x_1) &\leq (1-\lambda)\left(\varphi(x_3) - \varphi(x_1)\right) \\
&= \frac{|x_1 - x_3| - |x_2 - x_3|}{|x_1 - x_3|}\left(\varphi(x_3) - \varphi(x_1)\right) \\
&\leq \frac{4\sqrt{d}M|x_1 - x_2|}{\delta^*}
\end{aligned}$$

where $M = \|\varphi\|_{L^\infty(\overline{B(x,\delta^*/\sqrt{d})})}$ and we use that $|x_1 - x_3| \geq \delta^*/(2\sqrt{d})$. Switching x_1 and x_2 implies that $|\varphi(x_2) - \varphi(x_1)| \leq \frac{4\sqrt{d}M|x_1-x_2|}{\delta^*}$ hence φ is Lipschitz continuous, with constant $L = \frac{4\sqrt{d}M}{\delta^*}$ in $B(x,\delta^*/(2\sqrt{d}))$.

For (2) let φ be differentiable at x. Then,

$$\begin{aligned}
\varphi(x) + \nabla\varphi(x)\cdot(z-x) &= \varphi(x) + \lim_{h\to 0^+}\frac{\varphi(x + (z-x)h) - \varphi(x)}{h} \\
&= \varphi(x) + \lim_{h\to 0^+}\frac{\varphi((1-h)x + hz) - \varphi(x)}{h} \\
&\leq \varphi(x) + \lim_{h\to 0^+}\frac{(1-h)\varphi(x) + h\varphi(z) - \varphi(x)}{h} \\
&= \varphi(z).
\end{aligned}$$

Hence $\nabla\varphi(x) \in \partial\varphi(x)$. Now if $y \in \partial\varphi(x)$ then

$$\varphi(x) + y\cdot(z-x) \leq \varphi(z)$$

for all $z \in \mathbb{R}^d$. Let $z = x + hw$ then we can infer that

$$y\cdot w \leq \frac{\varphi(x+hw) - \varphi(x)}{h}$$

for all $h > 0$ and $w \in \mathbb{R}^d$. Letting $h \to 0^+$ we have $y \cdot w \leq \nabla\varphi(x) \cdot w$ for all $w \in \mathbb{R}^d$. Substituting $w \mapsto -w$ we have $y \cdot w = \nabla\varphi(x) \cdot w$ for all $w \in \mathbb{R}^d$. Hence $y = \nabla\varphi(x)$. □

Our final preliminary result gives equivalent conditions for convexity and lower semi-continuity.

Proposition 6.6. *Let $\varphi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be proper. Then the following are equivalent:*

1. *φ is convex and lower semi-continuous;*
2. *$\varphi = \psi^*$ for some proper function ψ;*
3. *$\varphi^{**} = \varphi$.*

Proof. 3 clearly implies 2. We first show that 2 implies 1. Let $\varphi = \psi^*$ and we will show that φ is convex and lower semi-continuous. For convexity let $x_1, x_2 \in \mathbb{R}^d, t \in [0,1]$ then

$$\begin{aligned}
\varphi(tx_1 + (1-t)x_2) &= \psi^*(tx_1 + (1-t)x_2) \\
&= \sup_{y \in \mathbb{R}^d} \left((tx_1 + (1-t)x_2) \cdot y - \psi(y)\right) \\
&\leq \sup_{y \in \mathbb{R}^d} \left(tx_1 \cdot y - t\psi(y)\right) + \sup_{y \in \mathbb{R}^d} \left((1-t)x_2 \cdot y - (1-t)\psi(y)\right) \\
&= t\psi^*(x_1) + (1-t)\psi^*(x_2) \\
&= t\varphi(x_1) + (1-t)\varphi(x_2).
\end{aligned}$$

For lower semi-continuity let $x_m \to x$, then

$$\liminf_{m\to\infty} \varphi(x_m) = \liminf_{m\to\infty} \sup_{y \in \mathbb{R}^d} \left(x_m \cdot y - \psi(y)\right) \geq \lim_{m\to\infty} \left(x_m \cdot y - \psi(y)\right) = x \cdot y - \psi(y)$$

for any $y \in \mathbb{R}^d$. Taking the supremum over $y \in \mathbb{R}^d$ implies $\liminf_{m\to\infty} \varphi(x_n) \geq \varphi(x)$ as required.

Finally we show that 1 implies 3. Let φ be lower semi-continuous and convex. Fix $x \in \mathbb{R}^d$, we want to show that $\varphi(x) = \varphi^{**}(x)$. Since $\varphi^*(y) \geq x \cdot y - \varphi(x)$ for all $y \in \mathbb{R}^d$ then

$$\varphi(x) \geq \sup_{y \in \mathbb{R}^d} \left(x \cdot y - \varphi^*(y)\right) = \varphi^{**}(x).$$

We are left to show $\varphi(x) \leq \varphi^{**}(x)$.

Let $x \in \mathrm{int}(\mathrm{Dom}(\varphi))$, then since φ can be bounded below by an affine function passing through $\varphi(x)$ (since φ is convex) then $\partial\varphi(x) \neq \emptyset$. Let $y_0 \in \partial\varphi(x)$. By Proposition 6.4 $x \cdot y_0 = \varphi(x) + \varphi^*(y_0)$ then

$$\varphi(x) = x \cdot y_0 - \varphi^*(y_0) \leq \sup_{y \in \mathbb{R}^d} \left(x \cdot y - \varphi^*(y)\right) = \varphi^{**}(x).$$

Hence we have proved the proposition for any φ with $\mathrm{int}(\mathrm{Dom}(\varphi)) = \mathbb{R}^d$.

For all other φ we define $\psi_\varepsilon(x) = \frac{|x|^2}{\varepsilon}$ and

$$\varphi_\varepsilon(x) = \inf_{y \in \mathbb{R}^d} \left(\varphi(x-y) + \psi_\varepsilon(y)\right) = \inf_{y \in \mathbb{R}^d} \left(\varphi(y) + \psi_\varepsilon(x-y)\right).$$

In order to show $\varphi_\varepsilon = \varphi_\varepsilon^{**}$ on $\mathbb{R}^d$ it is enough to show that φ_ε is convex, lower semi-continuous and $\mathrm{int}(\mathrm{Dom}(\varphi_\varepsilon)) = \mathbb{R}^d$. For convexity we note,

$$\begin{aligned}
\varphi_\varepsilon(tx_1 + (1-t)x_2) &= \inf_{y\in\mathbb{R}} \left(\varphi(tx_1 + (1-t)x_2 - y) + \psi_\varepsilon(y)\right) \\
&= \inf_{y_1,y_2\in\mathbb{R}^d} \left(\varphi(t(x_1-y_1) + (1-t)(x_2-y_2)) + \psi_\varepsilon(ty_1 + (1-t)y_2)\right) \\
&\leq \inf_{y_1,y_2\in\mathbb{R}^d} \left(t\left(\varphi(x_1-y_1) + \psi_\varepsilon(y_1)\right) + (1-t)\left(\varphi(x_2-y_2) + \psi_\varepsilon(y_2)\right)\right) \\
&= t\varphi_\varepsilon(x_1) + (1-t)\varphi_\varepsilon(x_2).
\end{aligned}$$

The pointwise limit of an arbitrary collection of lower semi-continuous functions is lower semi-continuous, hence φ_ε is lower semi-continuous. Now let $x \in \mathbb{R}^d$ and since φ is proper there exists $y_0 \in \mathbb{R}^d$ such that $\varphi(y_0) < \infty$, hence $\varphi_\varepsilon(x) \leq \varphi(y_0) + \psi_\varepsilon(x - y_0)$. Since ψ is everywhere finite then it follows that $\varphi_\varepsilon(x)$ is finite hence $x \in \mathrm{Dom}(\varphi_\varepsilon)$. In particular $\mathrm{int}(\mathrm{Dom}(\varphi_\varepsilon)) = \mathbb{R}^d$.

We now show that $\liminf_{\varepsilon\to 0} \varphi_\varepsilon(x) \geq \varphi(x)$ (we actually show that $\lim_{\varepsilon\to 0} \varphi_\varepsilon(x) = \varphi(x)$ but the former statement is all we really need). Fix $x \in \mathbb{R}^d$ and note that $\varphi_\varepsilon(x) \leq A + \frac{B}{\varepsilon}$ Since φ is convex then it is bounded below by an affine function, say $\varphi(z) \geq a\cdot z + b$ for all $z \in \mathbb{R}^d$. Let y_ε be a minimising sequence, i.e. $\varphi_\varepsilon(x) \geq \varphi(x - y_\varepsilon) + \psi_\varepsilon(y_\varepsilon) - \varepsilon$. Then

$$A + \frac{B}{\varepsilon} \geq \varphi_\varepsilon(x) \geq a\cdot(x - y_\varepsilon) + \frac{|y_\varepsilon|^2}{\varepsilon} - \varepsilon \geq -\frac{(1+\varepsilon)|a|^2}{\varepsilon} - \frac{|x|^2}{2} + \frac{|y_\varepsilon|^2}{2\varepsilon} - \varepsilon.$$

This implies $|y_\varepsilon| = O(1)$.

Now let $\varepsilon_n \to 0$ be a subsequence with $\liminf_{\varepsilon\to 0} \varphi_\varepsilon(x) = \lim_{n\to\infty} \varphi_{\varepsilon_n}(x)$. Since y_{ε_n} is bounded then there exists a further subsequence (which we relabel) and some $y \in \mathbb{R}^d$ such that $y_{\varepsilon_n} \to y$. Furthermore,

$$\lim_{n\to\infty} \varphi_{\varepsilon_n}(x) = \lim_{n\to\infty} \left(\varphi(x - y_{\varepsilon_n}) + \psi_{\varepsilon_n}(y_{\varepsilon_n})\right) \geq \begin{cases} \varphi(x) & \text{if } y = 0 \\ +\infty & \text{else.} \end{cases}$$

In both cases the right hand side is greater than $\varphi(x)$ hence $\lim_{n\to\infty} \varphi_{\varepsilon_n}(x) \geq \varphi(x)$.

On the other hand, since $\varphi(x) \geq \varphi_\varepsilon(x)$ then

$$\varphi^{**}(x) = \sup_{y\in\mathbb{R}^d} \inf_{z\in\mathbb{R}^d} \left(y\cdot(x-z) + \varphi(z)\right) \geq \sup_{y\in\mathbb{R}^d} \inf_{z\in\mathbb{R}^d} \left(y\cdot(x-z) + \varphi_\varepsilon(z)\right) = \varphi^{**}(x).$$

Hence,

$$\varphi^{**}(x) \geq \liminf_{\varepsilon\to 0} \varphi_\varepsilon^{**}(x) = \liminf_{\varepsilon\to 0} \varphi_\varepsilon(x) \geq \varphi(x).$$

Which concludes the proof. $\square$

6.3 Proof of the Knott-Smith Optimality Criterion

Section references: The proof of the Knott-Smith optimality condition is based on Villani's proof in [23, Theorem 2.12].

Before proving the theorem we manipulate the Kantorovich optimal transport problem and the dual problem. Let $(\varphi,\psi)\in\Phi_c$ (where $c(x,y)=\frac{1}{2}|x-y|^2$) and define $\tilde{\varphi}(x)=\frac{1}{2}|x|^2-\varphi(x)$, $\tilde{\psi}(y)=\frac{1}{2}|y|^2-\psi(y)$. Clearly $\tilde{\varphi}\in L^1(\mu)$, $\tilde{\psi}\in L^1(\nu)$ whenever μ, ν have finite second moments. Furthermore

$$\tilde{\varphi}(x)+\tilde{\psi}(y)=\frac{1}{2}|x|^2+\frac{1}{2}|y|^2-\varphi(x)-\psi(y)\geq\frac{1}{2}|x|^2+\frac{1}{2}|y|^2-\frac{1}{2}|x-y|^2\geq x\cdot y.$$

In fact, $(\varphi,\psi)\in\Phi_c\Leftrightarrow(\tilde{\varphi},\tilde{\psi})\in\tilde{\Phi}$ where

$$\tilde{\Phi}=\left\{(\tilde{\varphi},\tilde{\psi})\in L^1(\mu)\times L^1(\nu)\,:\,\tilde{\varphi}(x)+\tilde{\psi}(y)=x\cdot y\right\}.$$

Furthermore $\mathbb{J}(\tilde{\varphi},\tilde{\psi})=M-\mathbb{J}(\varphi,\psi)$ where

$$M=\frac{1}{2}\int_X|x|^2\,\mathrm{d}\mu(x)+\frac{1}{2}\int_Y|y|^2\,\mathrm{d}\nu(y).$$

And for $\pi\in\Pi(\mu,\nu)$,

$$\mathbb{K}(\pi)=\frac{1}{2}\int_{X\times Y}|x-y|^2\,\mathrm{d}\pi(x,y)=M-\int_{X\times Y}x\cdot y\,\mathrm{d}\pi(x,y).$$

Hence,

$$M-\mathbb{J}(\tilde{\varphi},\tilde{\psi})=\mathbb{J}(\varphi,\psi)\leq\mathbb{K}(\pi)=M-\int_{X\times Y}x\cdot y\,\mathrm{d}\pi(x,y).$$

Or more conveniently,

$$\mathbb{J}(\tilde{\varphi},\tilde{\psi})\geq\int_{X\times Y}x\cdot y\,\mathrm{d}\pi(x,y).$$

Kantorovich duality (Theorem 4.1) implies that

$$\min_{(\tilde{\varphi},\tilde{\psi})\in\tilde{\Phi}}\mathbb{J}(\tilde{\varphi},\tilde{\psi})=\max_{\pi\in\Pi(\mu,\nu)}\int_{X\times Y}x\cdot y\,\mathrm{d}\pi(x,y).\tag{6.2}$$

We notice that if $\pi^\dagger\in\Pi(\mu,\nu)$ minimises $\mathbb{K}$ then it also maximises $\int_{X\times Y}x\cdot y\,\mathrm{d}\pi(x,y)$, and vice versa. On the other hand, if $(\varphi,\psi)\in\Phi_c$ maximises $\mathbb{J}$, then $(\tilde{\varphi},\tilde{\psi})=(\frac{1}{2}|\cdot|^2-\varphi,\frac{1}{2}|\cdot|^2-\psi)\in\tilde{\Phi}$ minimises $\mathbb{J}$, and vice versa.

Existence of maximisers **(Theorem 4.10)** of $\mathbb{J}$ over Φ_c imply there exists $\varphi\in L^1(\mu)$ **with** $\varphi=\varphi^{cc}$ such that $(\tilde{\varphi},\tilde{\psi})=(\frac{1}{2}|\cdot|^2-\varphi^{cc},\frac{1}{2}|\cdot|^2-\varphi^c)\in\tilde{\Phi}$ and $(\tilde{\varphi},\tilde{\psi})$ minimise $\mathbb{J}$ over $\tilde{\Phi}$. Furthermore,

$$\begin{aligned}\tilde{\psi}(y)&=\frac{1}{2}|y|^2-\varphi^c(y)\\&=\sup_{x\in X}\left(\frac{1}{2}|y|^2-\frac{1}{2}|x-y|^2+\varphi(x)\right)\\&=\sup_{x\in X}\left(x\cdot y-\tilde{\varphi}(x)\right)\\&=\tilde{\varphi}^*(y)\end{aligned}$$

where $\tilde{\varphi}^*$ is the Legendre-Fenchel transform of $\tilde{\varphi}$. We also have,

$$\begin{aligned}\tilde{\varphi}(x) &= \frac{1}{2}|x|^2 - \varphi^{cc}(x)\\ &= \sup_{y\in Y}\left(\frac{1}{2}|x|^2 - \frac{1}{2}|x-y|^2 + \varphi^c(y)\right)\\ &= \sup_{y\in Y}\left(\frac{1}{2}|x|^2 - \frac{1}{2}|x-y|^2 + \frac{1}{2}|y|^2 - \tilde{\varphi}^*(y)\right)\\ &= \sup_{y\in Y}\left(x\cdot y - \tilde{\varphi}^*(y)\right)\\ &= \tilde{\varphi}^{**}(x).\end{aligned}$$

Hence minimisers of $\mathbb{J}$ over $\tilde{\Phi}$ take the form $(\tilde{\varphi}^{**}, \tilde{\varphi}^*)$.

Let $\tilde{\eta} = \tilde{\varphi}^{**}$ then by Proposition 6.6 $\tilde{\eta}$ is convex and lower semi-continuous. Furthermore, again by Proposition 6.6 $\tilde{\eta}^* = \tilde{\varphi}^{***} = \tilde{\varphi}^*$. Hence there exists minimisers of $\mathbb{J}$ over $\tilde{\Phi}$ with the form $(\tilde{\eta}, \tilde{\eta}^*)$ where $\tilde{\eta}$ is a proper, convex and lower semi-continuous function.

Proof of Theorem 6.1. Let $\pi^\dagger \in \Pi(\mu,\nu)$ minimise $\mathbb{K}$ over $\Pi(\mu,\nu)$ and $\tilde{\varphi}$ be the proper lower semi-continuous function such that the pair $(\tilde{\varphi}, \tilde{\varphi}^*)$ minimise $\mathbb{J}$ over $\tilde{\Phi}$. By Kantorovich duality (in particular (6.2)) we have

$$\int_X \tilde{\varphi}(x)\,\mathrm{d}\mu(x) + \int_Y \tilde{\varphi}^*(y)\,\mathrm{d}\nu(y) = \int_{X\times Y} x\cdot y\,\mathrm{d}\pi^\dagger(x,y).$$

Equivalently,

$$\int_{X\times Y} \tilde{\varphi}(x) + \tilde{\varphi}^*(y) - x\cdot y\,\mathrm{d}\pi^\dagger(x,y) = 0.$$

By definition of the convex conjugate $\tilde{\varphi}(x)+\tilde{\varphi}^*(y) \geq x\cdot y$ and therefore the integrand is non-negative. We must have $\tilde{\varphi}(x) + \tilde{\varphi}^*(y) = x\cdot y$ for $\pi^\dagger$-almost every (x,y) and therefore by Proposition 6.4 $y \in \partial\tilde{\varphi}(x)$ for $\pi^\dagger$-almost every (x,y).

Conversely, suppose $\pi^\dagger \in \Pi(\mu,\nu)$, and $y \in \partial\tilde{\varphi}(x)$ for $\pi^\dagger$-almost every (x,y) where $\tilde{\varphi}$ is a $L^1(\mu)$ proper, lower semi-continuous and convex function. We show that $\pi^\dagger$ and $(\tilde{\varphi}, \tilde{\varphi}^*)$ are optimal for their respective problems. Then by Proposition 6.4,

$$\int_{X\times Y} \tilde{\varphi}(x) + \tilde{\varphi}^*(y) - x\cdot y\,\mathrm{d}\pi^\dagger(x,y) = 0.$$

Notice that by definition of the Legendre-Fenchel transform we have that $\tilde{\varphi}(x) + \tilde{\varphi}^*(y) \geq x\cdot y$. We will show integrability of $\tilde{\varphi}^*$ shortly, for now it is assumed then $(\tilde{\varphi}, \tilde{\varphi}^*) \in \tilde{\Phi}$. Hence,

$$\min_{\tilde{\Phi}} \mathbb{J} \leq \mathbb{J}(\tilde{\varphi}, \tilde{\varphi}^*) = \int_{X\times Y} x\cdot y\,\mathrm{d}\pi^\dagger(x,y) \leq \max_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} x\cdot y\,\mathrm{d}\pi(x,y).$$

By duality (i.e. (6.2)) it follows that $(\tilde{\varphi}, \tilde{\varphi}^*)$ in $\tilde{\Phi}$ achieves the minimum of $\mathbb{J}$ and $\pi^\dagger$ achieves the maximum of $\int_{X\times Y} x\cdot y\,\mathrm{d}\pi(x,y)$ is $\Pi(\mu,\nu)$. Hence $\pi^\dagger$ is an optimal plan in the Kantorovich sense.

The last detail we have to show is $\tilde{\varphi}^* \in L^1(\nu)$. Since $\tilde{\varphi}$ is convex then $\tilde{\varphi}^*$ can be bounded below by an affine function, in particular there exists $x_0 \in X$ such that $\tilde{\varphi}^*(y) \geq x_0 \cdot y - \tilde{\varphi}(x_0) \geq x_0 \cdot y - b_0 =: f(y)$. So the integral,

$$\|\tilde{\varphi}^* - f\|_{L^1(\nu)} = \int_Y \tilde{\varphi}^*(y) - f(y)\,\mathrm{d}\nu(y) \leq \mathbb{J}(\tilde{\varphi}, \tilde{\varphi}^*) + \|\tilde{\varphi}\|_{L^1(\mu)} + \frac{1}{2}|x_0|^2 + \frac{1}{2}\int_Y |y|^2\,\mathrm{d}\nu(y) + b_0$$

is finite. Hence $\tilde{\varphi}^* - f \in L^1(\nu)$, and since $f \in L^1(\nu)$ then $\tilde{\varphi}^* \in L^1(\nu)$ as required. $\square$

6.4 Proof of Brenier's Theorem

Section references: The proof of the Brenier's theorem is based on Villani's proof in [23, Theorem 2.12].

Proof of Theorem 6.2. Let $\pi^\dagger$ be a minimiser of Kantorovich's optimal transport problem. If we write (by disintegration of measures)

$$\pi^\dagger(A \times B) = \int_A \pi^\dagger(B|x)\,\mathrm{d}\mu(x),$$

for some family $\{\pi^\dagger(\cdot|x)\}_{x\in X} \subset \mathcal{P}(Y)$, then $\mathrm{supp}(\pi^\dagger(\cdot|x)) \subseteq \partial\tilde{\varphi}^\dagger(x)$ for μ-a.e. $x \in X$ by Theorem 6.1. By Proposition 6.5, $\partial\tilde{\varphi}^\dagger(x) = \{\nabla\tilde{\varphi}^\dagger(x)\}$ for $\mathcal{L}$-a.e. $x \in X$ (and therefore μ-a.e. $x \in X$). Hence $\mathrm{supp}(\pi^\dagger(\cdot|x)) \subset \{\nabla\tilde{\varphi}^\dagger(x)\}$ for μ-a.e. $x \in X$. This implies $\pi^\dagger(\cdot|x) = \partial_{\nabla\tilde{\varphi}^\dagger(x)}$ for μ-a.e. $x \in X$. We have shown that any optimal $\pi^\dagger$ that can be written as

$$\pi^\dagger = (\mathrm{Id} \times \nabla\tilde{\varphi}^\dagger)_{\#}\mu$$

and since

$$\begin{aligned}
\nu(B) &= \pi^\dagger(\mathbb{R}^d \times B) \\
&= (\mathrm{Id} \times \nabla\tilde{\varphi}^\dagger)_{\#}\mu(\mathbb{R}^d \times B) \\
&= \mu\left(\left(\mathrm{Id} \times \nabla\tilde{\varphi}^\dagger\right)^{-1}(\mathbb{R}^d \times B)\right) \\
&= \mu\left(\left\{x \,:\, (\mathrm{Id} \times \nabla\tilde{\varphi}^\dagger)(x) \in \mathbb{R}^d \times B\right\}\right) \\
&= \mu\left(\left\{x \,:\, \nabla\tilde{\varphi}^\dagger(x) \in B\right\}\right) \\
&= (\nabla\tilde{\varphi}^\dagger)_{\#}\mu(B)
\end{aligned}$$

then we also have that $(\nabla\tilde{\varphi}^\dagger)_{\#}\mu = \nu$. Since the choice of optimal $\pi^\dagger$ was made independently of $\tilde{\varphi}^\dagger$ then we have that the minimiser to Kantorovich's problem is unique.

We are left to show uniqueness (up to additive constants) of convex functions that derivatives push μ onto ν. Assume $\bar{\varphi}$ is another convex function with $(\nabla\bar{\varphi})_{\#}\mu = \nu$. We will show that $\nabla\tilde{\varphi}^\dagger = \nabla\bar{\varphi}$ upto μ null sets.

By Theorem 6.1 we know that $(\mathrm{Id} \times \nabla\bar{\varphi})_{\#}\mu$ is an optimal transport plan and the pair $(\bar{\varphi}, \bar{\varphi}^*)$ minimize $\mathbb{J}$ over $\tilde{\Phi}$. So,

$$\int_X \bar{\varphi}\,\mathrm{d}\mu + \int_Y \bar{\varphi}^*\,\mathrm{d}\nu = \int_X \tilde{\varphi}^\dagger\,\mathrm{d}\mu + \int_Y (\tilde{\varphi}^\dagger)^*\,\mathrm{d}\nu.$$

The above implies that

$$\begin{aligned}
\int_{X\times Y} \bar{\varphi}(x) + \bar{\varphi}^*(y)\,\mathrm{d}\pi^\dagger(x,y) &= \int_{X\times Y} \tilde{\varphi}^\dagger(x) + (\tilde{\varphi}^\dagger)^*(y)\,\mathrm{d}\pi^\dagger(x,y) \\
&= \int_{X\times Y} x\cdot y\,\mathrm{d}\pi^\dagger(x,y) \\
&= \int_{X\times Y} x\cdot y\,\mathrm{d}(\mathrm{Id}\times\nabla\tilde{\varphi}^\dagger)_{\#}\mu(x,y) \\
&= \int_X x\cdot\nabla\tilde{\varphi}^\dagger(x)\,\mathrm{d}\mu(x)
\end{aligned}$$

where the second line follows as $y \in \partial\tilde{\varphi}^\dagger(x)$ for $\pi^\dagger$-a.e. (x,y) and by Proposition 6.4. Also,

$$\int_{X\times Y} \bar{\varphi}(x) + \bar{\varphi}^*(y)\,\mathrm{d}\pi^\dagger(x,y) = \int_X \bar{\varphi}(x) + \bar{\varphi}^*(\nabla\tilde{\varphi}^\dagger(x))\,\mathrm{d}\mu(x).$$

Hence

$$\int_X \left(\bar{\varphi}(x) + \bar{\varphi}^*(\nabla\tilde{\varphi}^\dagger(x)) - x\cdot\nabla\tilde{\varphi}^\dagger(x)\right)\,\mathrm{d}\mu(x) = 0.$$

In particular, $\bar{\varphi}(x) + \bar{\varphi}^*(\nabla\tilde{\varphi}^\dagger(x)) - x\cdot\nabla\tilde{\varphi}^\dagger(x) = 0$ for μ-almost every x. By Proposition 6.4 this implies that $\nabla\tilde{\varphi}^\dagger(x) \in \partial\bar{\varphi}(x)$ for μ-almost every x and therefore $\nabla\tilde{\varphi}^\dagger(x) = \nabla\bar{\varphi}(x)$ for μ-almost every x. □

Chapter 7

Wasserstein Distances

Eulerian based costs, such as L^p, define a metric based on "pointwise differences". This has some notable disadvantages, for example consider in 1D two indicator functions $f(x) = \chi_{[0,1]}(x)$ and $f_\delta(x) = \chi_{[\delta,\delta+1]}(x)$. Notice that in L^p,

$$\|f - f_\delta\|_{L^p}^p = \begin{cases} 2\delta & \text{if } |\delta| < 1 \\ 2 & \text{else.} \end{cases}$$

In particular, we notice that once $|\delta| \geq 1$ the L^p distance is constant. In more general examples, where f and f_δ are not necessarily indicator functions, the L^p distance will be the sum of the L^p norms whenever the supports of f and f_δ are disjoint.

Why do we care? Say we are trying to fit a parametrised curve f_δ to f. Then say we start from a bad initialisation where the support of f_δ is disjoint from the support of f. In this regime the derivative $\frac{\mathrm{d}}{\mathrm{d}\delta}\|f_\delta - f\|_{L^p} = 0$, this is a problem for gradient based optimisation.

On the other hand, we would hope that a transport based distance would do a better job. In particular, in the elementary example $f(x) = \chi_{[0,1]}(x)$ and $f_\delta(x) = \chi_{[\delta,\delta+1]}(x)$ the OT cost would be

$$\min_{T_\# f = f_\delta} \int_0^1 |x - T(x)|^p \,\mathrm{d}x = |\delta|^p$$

where the cost is $c(x, y) = |x - y|^p$ and with an abuse of notation we associate f and f_δ with the measures with density f and f_δ respectively. Note that the OT cost now strictly increases as a function of $|\delta|$.

The objective of this section is to understand how the optimal transport can be used to define a metric and some of the metric properties. In particular, we will define the p-Wasserstein distance (also sometimes known as the earth movers distance when $p = 1$) in the next section. In Section 7.2 we look at the topology of 2-Wasserstein spaces and show that the p-Wasserstein distance metrizes the weak* convergence. Finally we will look at geodesics.

Throughout this chapter we will assume that $c(x, y) = |x - y|^p$ for $p \in [1, +\infty)$ and X, Y are subsets of $\mathbb{R}^d$.

Before proceeding to the p-Wasserstein distance, let us note one other important example that can be posed as an optimal transport problem. Let $c(x, y) = \mathbb{1}_{x \neq y}$, i.e. $c(x, y) = 0$ if $x = y$

and $c(x,y) = 1$ otherwise. Then the optimal transport problem coincides with the total variation distance between measures.

Proposition 7.1. *Let $\mu, \nu \in \mathcal{P}(X)$ where $X \subseteq \mathbb{R}^d$ and $c(x,y) = \mathbb{1}_{x \neq y}$ then*

$$\inf_{\pi \in \Pi(\mu,\nu)} \mathbb{K}(\pi) = \frac{1}{2} \|\mu - \nu\|_{\mathrm{TV}}$$

where

$$\|\mu\|_{\mathrm{TV}} := 2 \sup_A |\mu(A)|.$$

Proof. By the Kantorovich-Rubinstein Theorem (see Theorem 4.13),

$$\begin{aligned} \min_{\pi \in \Pi(\mu,\nu)} \mathbb{K}(\pi) &= \sup \left\{ \int_X f \, \mathrm{d}(\mu - \nu) \; : \; f \in L^1(|\mu - \nu|), \, |f(x) - f(y)| \leq 1 \, \forall x, y \in X \right\} \\ &= \sup \left\{ \int_X f \, \mathrm{d}(\mu - \nu) \; : \; f \in L^1(|\mu - \nu|), \, 0 \leq f(x) \leq 1 \, \forall x \in X \right\} \end{aligned}$$

since $\|f\|_{\mathrm{Lip}} = \sup_{x \neq y} |f(x) - f(y)|$. Now let $\nu - \mu = (\nu - \mu)_+ - (\nu - \mu)_-$ be the decomposition of $\nu - \mu$ where $(\nu - \mu)_\pm \in \mathcal{M}_+(X)$ are singular. It follows that

$$\|\mu - \nu\|_{\mathrm{TV}} = 2(\nu - \mu)_+(X).$$

And,

$$\sup_{0 \leq f \leq 1} \int_X f \, \mathrm{d}(\mu - \nu) = (\nu - \mu)_+(X)$$

by choosing $f(x) = 1$ on the support of $(\nu - \mu)_+$ and $f(x) = 0$ on the support of $(\nu - \mu)_-$. Hence $\min_{\pi \in \Pi(\mu,\nu)} \mathbb{K}(\pi) = \sup_{0 \leq f \leq 1} \mathbb{J}(-f, f) = \frac{1}{2} \|\mu - \nu\|_{\mathrm{TV}}$. □

7.1 Wasserstein Distances

Section references: The proof of that the Wasserstein distance is a metric, Proposition 7.3, comes from [20, Proposition 5.1 and Lemma 5.4], with the preliminary result, Lemma 7.4, coming from [20, Lemma 5.5]. The equivalence of Wasserstein norms is from [20, Section 5.1].

We will work on the space of probability measures on $X \subseteq \mathbb{R}^d$ with bounded p^{th} moment, i.e.

$$\mathcal{P}_p(X) := \left\{ \mu \in \mathcal{P}(X) \; : \; \int_X |x|^p \, \mathrm{d}\mu(x) < +\infty \right\}.$$

Of course, if X is bounded then $\mathcal{P}_p(X) = \mathcal{P}(X)$. We now define the p-Wasserstein distance, it will be the objective of this section to prove that the p-Wasserstein distance is a metric.

Definition 7.2. *Let $\mu, \nu \in \mathcal{P}_p(X)$, then the p-Wasserstein distance is defined as*

$$d_{W^p}(\mu, \nu) = \min_{\pi \in \Pi(\mu,\nu)} \left(\int_{X \times X} |x - y|^p \, \mathrm{d}\pi(x,y) \right)^{\frac{1}{p}}.$$

The p-Wasserstein distance is the p^{th} root of the minimum of the Kantorovich optimal transport problem for cost function $c(x,y) = |x-y|^p$. The motivation is that this cost resembles an L^p distance (in fact we use properties of L^p distances to prove the triangle inequality). One could also consider an analogous distance for cost function $c(x,y) = d(x,y)$ where d is a metric on X. This type of distance is known as the earth movers distance. Notice that when $p = 1$ the p-Wasserstein distance is also a earth movers distance. We will not focus on earth movers distances here.

Let us note here that $\mu, \nu \in \mathcal{P}_p(X)$ is enough to guarantee $d_{W^p}(\mu,\nu) < +\infty$. In particular,

$$d_{W^p}^p(\mu,\nu) \leq p \inf_{\pi \in \Pi(\mu,\nu)} \int_{X\times X} |x|^p + |y|^p \,\mathrm{d}\pi(x,y) = p \int_X |x|^p \,\mathrm{d}\mu(x) + p \int_X |y|^p \,\mathrm{d}\nu(y).$$

We now state the result that d_{W^p} is a metric. The proof, minus the triangle inequality, is given below.

Proposition 7.3. *Let $X \subseteq \mathbb{R}^d$. The distance $d_{W^p} : \mathcal{P}_p(X) \times \mathcal{P}_p(X) \to [0,\infty)$ is a metric on $\mathcal{P}_p(X)$.*

Proof. We give the proof of all the required criteria with the exception of the triangle inequality which will require some preliminary results. Firstly, it is clear that $d_{W^p}(\mu,\nu) \geq 0$ for all $\mu, \nu \in \mathcal{P}(X)$ and by symmetry of the cost function $c(x,y) = |x-y|^p$ and $\pi \in \Pi(\mu,\nu) \Leftrightarrow S_\# \pi \in \Pi(\nu,\mu)$ where $S(x,y) = (y,x)$ we have symmetry of d_{W^p}. Now if $\mu = \nu$ then we can take $\pi(x,y) = \delta_x(y)\mu(x)$ so that

$$d_{W^p}^p(\mu,\nu) \leq \int_{X\times X} |x-y|^p \,\mathrm{d}\pi(x,y) = 0$$

as $x = y$ π-almost everywhere. Now if $d_{W^p}(\mu,\nu) = 0$ then there exists $\pi \in \Pi(\mu,\nu)$ such that $x = y$ π-almost everywhere. Hence for any test function $f : X \to \mathbb{R}$,

$$\int_X f(x) \,\mathrm{d}\mu(x) = \int_{X\times X} f(x) \,\mathrm{d}\pi(x,y) = \int_{X\times X} f(y) \,\mathrm{d}\pi(x,y) = \int_X f(y) \,\mathrm{d}\nu(y).$$

As this holds for all test functions f then $\mu = \nu$. $\square$

The following lemma is known as the gluing lemma and we will use it to "glue" two transport plans $\pi_1 \in \Pi(\mu,\nu)$ and $\pi_2 \in \Pi(\nu,\omega)$. The triangle inequality then follows from the triangle inequality for L^p distances.

Lemma 7.4. *Let $X, Y, Z \subseteq \mathbb{R}^d$. Given measures $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y), \omega \in \mathcal{P}(Z)$ and transport plans $\pi_1 \in \Pi(\mu,\nu)$ and $\pi_2 \in \Pi(\nu,\omega)$ there exists a measure $\gamma \in \mathcal{P}(X \times Y \times Z)$ such that $P_\#^{X,Y}\gamma = \pi_1$ and $P_\#^{Y,Z}\gamma = \pi_2$ where $P^{X,Y}(x,y,z) = (x,y)$ and $P^{Y,Z}(x,y,z) = (y,z)$ are the projections onto the two first and two last variables respectively.*

Proof. By the disintegration of measures we can write

$$\pi_1(A \times B) = \int_B \pi_1(A|y) \,\mathrm{d}\nu(y)$$

for some family of probability measures $\pi_1(\cdot|y) \in \mathcal{P}(X)$, and similarly for π_2,

$$\pi_2(B \times C) = \int_B \pi_2(C|y)\,\mathrm{d}\nu(y).$$

Define $\gamma \in \mathcal{M}(X \times Y \times Z)$ by

$$\gamma(A \times B \times C) = \int_B \pi_1(A|y)\pi_2(C|y)\,\mathrm{d}\nu(y).$$

Then,

$$\gamma(A \times B \times Z) = \int_B \pi_1(A|y)\pi_2(Z|y)\,\mathrm{d}\nu(y) = \int_B \pi_1(A|y)\,\mathrm{d}\nu(y) = \pi_1(A \times B).$$

Similarly, $\gamma(X \times B \times C) = \pi_2(B \times C)$. Therefore, $P_\#^{X,Y}\gamma = \pi_1$ and $P_\#^{Y,Z}\gamma = \pi_2$ as required. $\square$

We are now in a position to complete the proof of Proposition 7.3.

Proof of Proposition 7.3 (triangle inequality). Let $\mu, \nu, \omega \in \mathcal{P}_p(X)$ and assume $\pi_{XY} \in \Pi(\mu, \nu)$, $\pi_{YZ} \in \Pi(\nu, \omega)$ are optimal, i.e.

$$\begin{aligned} d_{W^p}^p(\mu, \nu) &= \int_{X \times X} |x - y|^p \,\mathrm{d}\pi_{XY}(x, y) \\ d_{W^p}^p(\nu, \omega) &= \int_{X \times X} |y - z|^p \,\mathrm{d}\pi_{YZ}(y, z). \end{aligned}$$

Let $\gamma \in \mathcal{P}(X \times X \times X)$ be such that $P_\#^{X,Y}\gamma = \pi_{XY}$ and $P_\#^{Y,Z}\gamma = \pi_{YZ}$ (such γ exists by Lemma 7.4). Let $\pi_{XZ} = P_\#^{X,Z}\gamma$. Then,

$$\begin{aligned} \pi_{XZ}(A \times X) &= P_\#^{X,Z}\gamma(A \times X) \\ &= \gamma\left(\left\{(x, y, z) \,:\, P^{X,Z}(x, y, z) = (x, z) \in A \times X\right\}\right) \\ &= \gamma\left(\{(x, y, z) \,:\, x \in A\}\right) \\ &= \gamma(A \times X \times X) \\ &= \pi_{XZ}(A \times X) \\ &= \mu(A). \end{aligned}$$

Similarly $\pi_{XZ}(X \times B) = \omega(B)$. So, $\pi_{XZ} \in \Pi(\mu, \omega)$.

Now,

$$
\begin{aligned}
d_{W^p}(\mu,\omega) &\leq \left(\int_{X\times X} |x-z|^p \,\mathrm{d}\pi_{XZ}(x,z)\right)^{\frac{1}{p}} \\
&= \left(\int_{X\times X\times X} |x-z|^p \,\mathrm{d}\gamma(x,y,z)\right)^{\frac{1}{p}} \\
&\leq \left(\int_{X\times X\times X} |x-y|^p \,\mathrm{d}\gamma(x,y,z)\right)^{\frac{1}{p}} + \left(\int_{X\times X\times X} |y-z|^p \,\mathrm{d}\gamma(x,y,z)\right)^{\frac{1}{p}} \\
&= \left(\int_{X\times X} |x-y|^p \,\mathrm{d}\pi_{XY}(x,y)\right)^{\frac{1}{p}} + \left(\int_{X\times X} |y-z|^p \,\mathrm{d}\pi_{YZ}(y,z)\right)^{\frac{1}{p}} \\
&= d_{W^p}(\mu,\nu) + d_{W^p}(\nu,\omega).
\end{aligned}
$$

This proves the triangle inequality. □

One can also prove the triangle inequality using an transport maps and an approximation argument. Slightly more precisely, if μ, ν and ω all have densities with respect to Lebesgue then we know there exits transport maps T and S with $T_{\#}\mu = \nu$ and $S_{\#}\nu = \omega$. The map $S\circ T$ then pushes μ onto ω. One can argue, similarly to our proof, that $d_{W^p}(\mu,\nu) + d_{W^p}(\nu,\omega) \geq d_{W^p}(\mu,\omega)$. To extend the argument to arbitrary probability measures μ, ν and ω one uses mollifiers to define $\tilde{\mu}_\varepsilon = \mu * J_\varepsilon$, analogously for $\tilde{\nu}_\varepsilon, \tilde{\omega}_\varepsilon$, where $J_\varepsilon = \frac{1}{\varepsilon^d} J(\cdot/\varepsilon)$ and J is a standard mollifier. The measures $\tilde{\mu}, \tilde{\nu}, \tilde{\omega}$ have densities with respect to the Lebesgue measure and one can show $d_{W^p}(\tilde{\mu}_\varepsilon, \tilde{\nu}_\varepsilon) \to d_{W^p}(\mu,\nu)$ as $\varepsilon \to 0$. We refer to [20, Lemma 5.2 and Lemma 5.3] for full details.

Our final result of the section gives sufficient conditions for equivalence (in the sense of a homeomorphism) of p-Wasserstein distances.

Proposition 7.5. *Let $X \subseteq \mathbb{R}^d$. For every $p \in [1,+\infty)$ and any $\mu,\nu \in \mathcal{P}_p(X)$ we have $d_{W^p}(\mu,\nu) \geq d_{W^1}(\mu,\nu)$. Furthermore, if X is bounded then $d_{W^p}^p(\mu,\nu) \leq \mathrm{diam}(X)^{p-1} d_{W^1}(\mu,\nu)$.*

Proof. By Jensen's inequality, for $\pi \in \Pi(\mu,\nu)$, we have

$$
\left(\int_{X\times X} |x-y|^p \,\mathrm{d}\pi(x,y)\right)^{\frac{1}{p}} \geq \int_{X\times X} |x-y| \,\mathrm{d}\pi(x,y).
$$

Hence, $d_{W^p}(\mu,\nu) \geq d_{W^1}(\mu,\nu)$.

Now if X is bounded, then $\forall x,y \in X$,

$$
|x-y|^p \leq (\max_{w,z\in X} |w-z|^{p-1})|x-y| = (\mathrm{diam}(X))^{p-1}|x-y|.
$$

Hence,

$$
\int_{X\times X} |x-y|^p \,\mathrm{d}\pi(x,y) \leq (\mathrm{diam}(X))^{p-1} \int_{X\times X} |x-y| \,\mathrm{d}\pi(x,y).
$$

From which it follows $d_{W^p}^p(\mu,\nu) \leq \mathrm{diam}(X)^{p-1} d_{W^1}(\mu,\nu)$. □

In fact the above is also true for $p=+\infty$, however we do not consider (or even define) d_{W^∞} here and instead refer to [20, Section 5.5.1] for more information on the ∞-Wasserstein distance.

7.2 The Wasserstein Topology

Section references: The two results regarding the relationship between convergence in Wasserstein distance and weak convergence can be found in [20, Theorem 5.10 and Theorem 5.11].*

In this section we prove the relationship of convergence in p-Wasserstein distance with weak* convergence. We start with when $X \subset \mathbb{R}^d$ is compact.

Theorem 7.6. *Let $X \subset \mathbb{R}^d$ be compact, and $\mu_m, \mu \in \mathcal{P}(X)$. Then $\mu_m \overset{*}{\rightharpoonup} \mu$ if and only if $d_{W^p}(\mu_m, \mu) \to 0$.*

Proof. By Proposition 7.5 it is enough to show the result for $p = 1$. Assume $d_{W^1}(\mu_m, \mu) \to 0$. By the Kantorovich-Rubinstein theorem, see Theorem 4.13, we can write

$$d_{W^1}(\mu, \nu) = \sup \left\{ \int_X \varphi \, \mathrm{d}(\mu - \nu) \, : \, \varphi \in L^1(|\mu - \nu|), |\varphi(x) - \varphi(y)| \leq |x - y| \right\}.$$

Let φ be a Lipschitz function with $\mathrm{Lip}(\varphi) > 0$ then $\tilde{\varphi} = \frac{1}{\mathrm{Lip}(\varphi)} \varphi$ is a 1-Lipschitz function and therefore

$$\frac{1}{\mathrm{Lip}(\varphi)} \int_X \varphi \, \mathrm{d}(\mu_m - \mu) = \int_X \tilde{\varphi} \, \mathrm{d}(\mu_m - \mu) \leq d_{W^1}(\mu_m, \mu) \to 0.$$

By substituting $\varphi \mapsto -\varphi$ we have that

$$\int_X \varphi \, \mathrm{d}\mu_m \to \int_X \varphi \, \mathrm{d}\mu$$

for any Lipschitz function φ. By the Portmanteau theorem $\mu_m \overset{*}{\rightharpoonup} \mu$.

For the converse statement we assume that $\mu_m \overset{*}{\rightharpoonup} \mu$ and let m_k be the subsequence such that

$$\lim_{k \to \infty} d_{W^1}(\mu_{m_k}, \mu) = \limsup_{m \to \infty} d_{W^1}(\mu_m, \mu).$$

Let $\tilde{\varphi}_{m_k}$ be 1-Lipschitz and such that

$$d_{W^1}(\mu_{m_k}, \mu) \leq \int_X \tilde{\varphi}_{m_k} \, \mathrm{d}(\mu_{m_k} - \mu) + \frac{1}{k}.$$

Pick $x_0 \in \mathrm{supp}(\mu)$. Note that, for any $\varphi \in L^1(\nu)$ where $\nu \in \mathcal{M}(X)$, $c \in \mathbb{R}$ that

$$\int_X (\varphi + c) \, \mathrm{d}\nu = \int_X \varphi \, \mathrm{d}\nu$$

if $\nu(X) = 0$. Hence if we let $\varphi_{m_k}(x) = \tilde{\varphi}_{m_k}(x) - \tilde{\varphi}_{m_k}(x_0)$ then

$$d_{W^1}(\mu_{m_k}, \mu) \leq \int_X \varphi_{m_k} \, \mathrm{d}(\mu_{m_k} - \mu) + \frac{1}{k},$$

φ_{m_k} are 1-Lipschitz (in particular equicontinuous) and bounded. By the Arzelà-Ascoli theorem there exists a further subsequence (relabelled) such that $\varphi_{m_k} \to \varphi$ uniformly. In particular, φ is 1-Lipschitz. Hence,

$$\begin{aligned}
\limsup_{m\to\infty} d_{W^1}(\mu_m,\mu) &\leq \limsup_{k\to\infty}\left(\int_X \varphi_{m_k}\,\mathrm{d}(\mu_{m_k}-\mu) + \frac{1}{k}\right)\\
&\leq \limsup_{k\to\infty}\left(\int_X (\varphi_{m_k}-\varphi)\,\mathrm{d}(\mu_{m_k}-\mu) + \int_X \varphi\,\mathrm{d}(\mu_{m_k}-\mu)\right)\\
&\leq \limsup_{k\to\infty}\|\varphi_{m_k}-\varphi\|_{L^\infty} + \limsup_{k\to\infty}\int_X \varphi\,\mathrm{d}(\mu_{m_k}-\mu)\\
&= 0.
\end{aligned}$$

Hence, $d_{W^1}(\mu_m,\mu)\to 0$ as $m\to\infty$. □

We now generalise to unbounded domains.

Theorem 7.7. *Let $\mu_m,\mu \in \mathcal{P}_p(\mathbb{R}^d)$. Then $d_{W^p}(\mu_m,\mu) \to 0$ if and only if $\int_{\mathbb{R}^d}|x|^p\,\mathrm{d}\mu_m \to \int_{\mathbb{R}^d}|x|^p\,\mathrm{d}\mu$ and $\mu_m \overset{*}{\rightharpoonup} \mu$.*

Proof. Let $d_{W^p}(\mu_m,\mu)\to 0$. Then by Proposition 7.5 we have $d_{W^1}(\mu_m,\mu)\to 0$. Analogously to the proof of Theorem 7.6 we have $\int_X \varphi\,\mathrm{d}(\mu_m-\mu)\to 0$ for all Lipschitz functions φ. Hence, by the Portmanteau theorem $\mu_m \overset{*}{\rightharpoonup} \mu$.

To show $\int_{\mathbb{R}^d}|x|^p\,\mathrm{d}\mu_m \to \int_{\mathbb{R}^d}|x|^p\,\mathrm{d}\mu$ we note that

$$\int_{\mathbb{R}^d}|x|^p\,\mathrm{d}\mu_m = d^p_{W^p}(\mu_m,\delta_0) \quad\text{and}\quad \int_{\mathbb{R}^d}|x|^p\,\mathrm{d}\mu = d^p_{W^p}(\mu,\delta_0).$$

Now,

$$d_{W^p}(\mu_m,\delta_0) \leq d_{W^p}(\mu_m,\mu) + d_{W^p}(\mu,\delta_0) \to d_{W^p}(\mu,\delta_0)$$

and

$$d_{W^p}(\mu_m,\delta_0) \geq d_{W^p}(\mu,\delta_0) - d_{W^p}(\mu_m,\mu) \to d_{W^p}(\mu,\delta_0)$$

Hence $\int_{\mathbb{R}^d}|x|^p\,\mathrm{d}\mu_m \to \int_{\mathbb{R}^d}|x|^p\,\mathrm{d}\mu$.

For the converse statement let $\mu_m \overset{*}{\rightharpoonup} \mu$ and $\int |x|^p\,\mathrm{d}\mu \to \int |x|^p\,\mathrm{d}\mu$. For any $R>0$ let $\phi_R(x) = (|x|\wedge R)^p = (\min\{|x|,R\})^p$ which is continuous and bounded. We have

$$\int_{\mathbb{R}^d}\left(|x|^p - \phi_R(x)\right)\,\mathrm{d}\mu_m \to \int_{\mathbb{R}^d}\left(|x|^p-\phi_R(x)\right)\,\mathrm{d}\mu \tag{7.1}$$

by weak* convergence and convergence of p^{th} moments. Now

$$\int_{\mathbb{R}^d}\left(|x|^p-\phi_R(x)\right)\,\mathrm{d}\mu(x) = \int_{|x|>R}|x|^p - R^p\,\mathrm{d}\mu \leq \int_{|x|>R}|x|^p\,\mathrm{d}\mu < \infty.$$

In particular, we let $\varepsilon > 0$ and choose $R > 0$ such that

$$\int_{\mathbb{R}^d} (|x|^p - \phi_R(x)) \, \mathrm{d}\mu(x) < \frac{\varepsilon}{2}.$$

By (7.1) we also have $\int_{\mathbb{R}^d} (|x|^p - \phi_R(x)) \, \mathrm{d}\mu_m(x) < \varepsilon$ for m sufficiently large.

For $a > b > 0$ and $p \geq 1$ we have $(a+b)^p = a^p + pb\xi^{p-1}$ for some $\xi \in [a, a+b]$. Hence, $(a+b)^p \geq a^p + pa^{p-1}b \geq a^p + b^p$.

Using the above, for $|x| > R$ we have $(|x| - R)^p \leq |x|^p - R^p = |x|^p - \phi_R(x)$. So for m sufficiently large,

$$\int_{|x|>R} (|x| - R)^p \, \mathrm{d}\mu_m < \varepsilon \quad \text{and} \quad \int_{|x|>R} (|x| - R)^p \, \mathrm{d}\mu < \varepsilon.$$

Let $P_R : \mathbb{R}^d \to \overline{B(0,R)}$ be the projection onto the ball $\overline{B(0,R)}$, i.e.

$$P_R(x) = \begin{cases} x & \text{if } x \in B(0,R) \\ \operatorname{argmin}_{y \in \partial B(0,R)} |y - x| & \text{else.} \end{cases}$$

The map P_R is continuous and equal to the identity on $\overline{B(0,R)}$. For all $x \notin \overline{B(0,R)}$ we have $|x - P_R(x)| = |x| - R$. Hence,

$$\begin{aligned} d_{W^p}(\mu, (P_R)_\# \mu) &\leq \left(\int_{\mathbb{R}^d} |x - P_R(x)|^p \, \mathrm{d}\mu(x) \right)^{\frac{1}{p}} \\ &= \left(\int_{|x|>R} |x - P_R(x)|^p \, \mathrm{d}\mu(x) \right)^{\frac{1}{p}} \\ &= \left(\int_{|x|>R} (|x| - R)^p \, \mathrm{d}\mu(x) \right)^{\frac{1}{p}} \\ &\leq \varepsilon^{\frac{1}{p}}, \end{aligned}$$

and similarly,

$$d_{W^p}(\mu_m, (P_R)_\# \mu_m) \leq \varepsilon^{\frac{1}{p}}.$$

For any $\varphi \in C_b^0(\mathbb{R}^d)$ we have

$$\int \varphi \, \mathrm{d}(P_R)_\# \mu_m = \int \varphi(P_R(x)) \, \mathrm{d}\mu_m \to \int \varphi(P_R(x)) \, \mathrm{d}\mu = \int \varphi \, \mathrm{d}(P_R)_\# \mu$$

since $\varphi \circ P_R$ is continuous and bounded. Hence, $(P_R)_\# \mu_m \overset{*}{\rightharpoonup} (P_R)_\# \mu$.

Now, $(P_R)_\# \mu_m$, $(P_R)_\# \mu$ have support in $\overline{B(0,R)}$ (a compact set) so by Theorem 7.6 we have $d_{W^p}((P_R)_\# \mu_m, (P_R)_\# \mu) \to 0$. Hence,

$$\begin{aligned} \limsup_{m \to \infty} d_{W^p}(\mu_m, \mu) &\leq \limsup_{m \to \infty} \Big(d_{W^p}(\mu_m, (P_R)_\# \mu_m) + d_{W^p}((P_R)_\# \mu_m, (P_R)_\# \mu) \\ &\qquad + d_{W^p}((P_R)_\# \mu, \mu) \Big) \\ &\leq 2\varepsilon^{\frac{1}{p}}. \end{aligned}$$

Letting $\varepsilon \to 0$ implies $\lim_{m\to\infty} d_{W^p}(\mu_m, \mu) = 0$ as required. □

7.3 Geodesics in the Wasserstein Space

Section references: The result that the Wasserstein space is a geodesic space (Theorem 7.11) can be found in [20, Theorem 5.27].

The aim of this section is to show that the p-Wasserstein space $(\mathcal{P}_p(X), d_{W^p})$ is a geodesic space. We start with some definitions. The definitions are given in terms of a metric space $(\mathcal{Z}, d)$, of course we have in mind that this will later be the Wasserstein space.

Definition 7.8. *Let $p \in [1, +\infty]$, $(\mathcal{Z}, d)$ be a metric space, and $\omega : (a, b) \to \mathcal{Z}$ a curve in $\mathcal{Z}$. We say $\omega \in \mathrm{AC}^p((a, b), \mathcal{Z})$ if there exists $g \in L^p((a, b))$ such that $d(\omega(t_0), \omega(t_1)) \leq \int_{t_0}^{t_1} g(s)\,\mathrm{d}s$ for any $a < t_0 < t_1 < b$. If $p = 1$ we say ω is an absolutely continuous curve. If $g \in L^p_{\mathrm{loc}}((a, b))$ then we say $\omega \in \mathrm{AC}^p_{\mathrm{loc}}((a, b), \mathcal{Z})$ and curves $\omega \in \mathrm{AC}^1_{\mathrm{loc}}((a, b), \mathcal{Z})$ are called locally absolutely continuous.*

Definition 7.9. *Let $(\mathcal{Z}, d)$ be a metric space and $\omega : [0, 1] \to \mathcal{Z}$ a curve in $\mathcal{Z}$. We define the length of ω by*

$$\mathrm{Len}(\omega) := \sup\left\{\sum_{k=0}^{n-1} d(\omega(t_k), \omega(t_{k+1})) \;:\; n \geq 1, 0 = t_0 < t_1 < \cdots < t_n = 1\right\}.$$

A curve $\omega : [0, 1] \to \mathcal{Z}$ is said to be a geodesic between $z_0 \in \mathcal{Z}$ and $z_1 \in \mathcal{Z}$ if

$$\omega \in \operatorname{argmin}\left\{\mathrm{Len}(\tilde{\omega}) \,\middle|\, \tilde{\omega} : [0, 1] \to \mathcal{Z},\, \tilde{\omega}(0) = z_0,\, \tilde{\omega}(1) = z_1\right\}.$$

A curve $\omega : [0, 1] \to \mathcal{Z}$ is said to be a constant speed geodesic between $z_0 \in \mathcal{Z}$ and $z_1 \in \mathcal{Z}$ if

$$d(\omega(t), \omega(s)) = |t - s| d(\omega(0), \omega(1)).$$

Note that if $\omega : [0, 1] \to \mathcal{Z}$ is a constant speed geodesic then it is a geodesic. Indeed, assume that $\omega : [0, 1] \to \mathcal{Z}$ and $\tilde{\omega} : [0, 1] \to \mathcal{Z}$ satisfy $\omega(0) = z_0 = \tilde{\omega}(0)$, $\omega(1) = z_1 = \tilde{\omega}(1)$,

$$d(\omega(t), \omega(s)) = |t - s| d(z_0, z_1) \quad \forall 0 \leq t < s \leq 1, \qquad \text{and} \qquad \mathrm{Len}(\tilde{\omega}) < \mathrm{Len}(\omega).$$

I.e. we assume that ω is a constant speed geodesic but not a geodesic. Then there exists $n \in \mathbb{N}$ and $0 \leq t_0 < t_1 < \cdots < t_n = 1$ such that

$$\mathrm{Len}(\tilde{\omega}) < \sum_{k=0}^{n-1} d(\omega(t_k), \omega(t_{k+1})) = d(z_0, z_1) \sum_{k=0}^{n-1} (t_{k+1} - t_k) = d(z_0, z_1).$$

This implies $\mathrm{Len}(\tilde{\omega}) < d(z_0, z_1)$. Clearly this is a contradiction (choosing $n = 1$ in the definition of $\mathrm{Len}(\tilde{\omega})$ implies $\mathrm{Len}(\tilde{\omega}) \geq d(z_0, z_1)$).

Note also that if $d(\omega(t), \omega(s)) = |t - s| d(z_0, z_1)$ then $\omega \in \mathrm{AC}^1((0, 1), \mathcal{Z})$ with $g(s) = d(z_0, z_1)$.

Definition 7.10. *Let $(\mathcal{Z}, d)$ be a metric space. We say $(\mathcal{Z}, d)$ is a length space if*

$$d(x,y) = \inf\left\{\mathrm{Len}(\omega) \,:\, \omega \in \mathrm{AC}^1((0,1),\mathcal{Z}),\, \omega(0) = x,\, \omega(1) = y\right\}.$$

We say $(\mathcal{Z}, d)$ is a geodesic space if

$$d(x,y) = \min\left\{\mathrm{Len}(\omega) \,:\, \omega \in \mathrm{AC}^1((0,1),\mathcal{Z}),\, \omega(0) = x,\, \omega(1) = y\right\}.$$

We now show that the Wasserstein space $(\mathcal{P}_p(X), d_{W^p})$ is a geodesic space.

Theorem 7.11. *Let $p \geq 1$, $X \subseteq \mathbb{R}^d$ be convex and define $P_t : X \times X \to X$ by $P_t(x,y) = (1-t)x + ty$. Let $\mu, \nu \in \mathcal{P}_p(X)$ and assume $\pi \in \Pi(\mu,\nu)$ minimises $\mathbb{K}$ over $\Pi(\mu,\nu)$ with cost $c(x,y) = |x-y|^p$. Then the curve $\mu_t = (P_t)_{\#}\pi$ is a constant speed geodesic in $(P_p(X), d_{W^p})$ connecting μ and ν. In particular, if $\pi = (\mathrm{Id} \times T)_{\#}\mu$ for some transport map $T : X \to X$ that pushes forwards μ to ν, i.e. $T_{\#}\mu = \nu$ (that is T is a solution to the Monge optimal transport problem), then $\mu_t = ((1-t)\mathrm{Id} + tT)_{\#}\mu$.*

Proof. Note that $P_0 = P^X$ and $P_1 = P^Y$. Therefore, $\mu_0 = (P_0)_{\#}\pi = \mu$, $\mu_1 = (P_1)_{\#}\pi = \nu$, so μ_t connects μ and ν. To show $d_{W^p}(\mu_s,\mu_t) = |t-s| d_{W^p}(\mu,\nu)$ it is enough to prove that $d_{W^p}(\mu_s,\mu_t) \leq |t-s| d_{W^p}(\mu,\nu)$. Indeed assuming this is true, then if $d_{W^p}(\mu_s,\mu_t) < |t-s| d_{W^p}(\mu,\nu)$ for any $0 \leq s < t \leq 1$ we have

$$\begin{aligned}
d_{W^p}(\mu,\nu) &\leq d_{W^p}(\mu,\mu_s) + d_{W^p}(\mu_s,\mu_t) + d_{W^p}(\mu_t,\nu) \\
&< (s + (t-s) + (1-t)) d_{W^p}(\mu,\nu) \\
&= d_{W^p}(\mu,\nu)
\end{aligned}$$

a contradiction.

To show $d_{W^p}(\mu_s,\mu_t) \leq |t-s| d_{W^p}(\mu,\nu)$ let $\pi_{s,t} = (P_s, P_t)_{\#}\pi$. Then for any (measurable) $A \subseteq X$,

$$\begin{aligned}
\pi_{s,t}(A \times X) &= \pi\left(\{(x,y) \,:\, (1-s)x + sy \in A,\, (1-t)x + ty \in X\}\right) \\
&= \pi\left(\{(x,y) \,:\, (1-s)x + sy \in A\}\right) \\
&= (P_s)_{\#}\pi(A) \\
&= \mu_s(A).
\end{aligned}$$

Hence $P^X_{\#}\pi_{s,t} = \mu_s$. Similarly, $P^Y_{\#}\pi_{s,t} = \mu_t$ so $\pi_{s,t} \in \Pi(\mu_s,\mu_t)$. Now,

$$\begin{aligned}
d_{W^p}(\mu_s,\mu_t) &\leq \left(\int_{X\times X} |x-y|^p \, \mathrm{d}\pi_{s,t}(x,y)\right)^{\frac{1}{p}} \\
&= \left(\int_{X\times X} |P_s(x,y) - P_t(x,y)|^p \, \mathrm{d}\pi(x,y)\right)^{\frac{1}{p}} \\
&= \left(\int_{X\times X} |(t-s)x - (t-s)y|^p \, \mathrm{d}\pi(x,y)\right)^{\frac{1}{p}} \\
&= |t-s| \left(\int_{X\times X} |x-y|^p \, \mathrm{d}\pi(x,y)\right)^{\frac{1}{p}} \\
&= |t-s| d_{W^p}(\mu,\nu)
\end{aligned}$$

as required.

If $\pi = (\mathrm{Id} \times T)_{\#}\mu$ where T is as in the statement of the theorem, then for $A \subset X$ (measurable) we have

$$\begin{aligned}
\mu_t(A) &= (P_t)_{\#}\pi(A) \\
&= \pi\left(\{(x,y) \,:\, (1-t)x + ty \in A\}\right) \\
&= (\mathrm{Id} \times T)_{\#}\mu\left(\{(x,y) \,:\, (1-t)x + ty \in A\}\right) \\
&= \mu\left(\{x \,:\, (1-t)x + tT(x) \in A\}\right) \\
&= ((1-t)\mathrm{Id} + tT)_{\#}\mu(A)
\end{aligned}$$

which shows $\mu_t = ((1-t)\mathrm{Id} + tT)_{\#}\mu$. □

The fact that the Wasserstein space $(\mathcal{P}_p(X), d_{W^p})$ is a geodesic space follows from the following argument. Pick $\mu, \nu \in \mathcal{P}_p(X)$ and let μ_t be the constant speed geodesic between μ and ν. First we see that $(\mu_t)_{t\in[0,1]}$ is a geodesic, and therefore

$$(\mu_t)_{t\in[0,1]} \in \operatorname{argmin}\left\{\mathrm{Len}(\omega) \,\middle|\, \omega : [0,1] \to \mathcal{P}_p(X),\, \omega(0) = \mu,\, \omega(1) = \nu\right\}.$$

Secondly, since $d_{W^p}(\mu_{t_k}, \mu_{t_{k+1}}) = |t_{k+1} - t_k| d_{W^p}(\mu, \nu)$ it follows immediately from the definition that $\mathrm{Len}((\mu_t)_{t\in[0,1]}) = d_{W^p}(\nu, \nu)$. Since the choice of μ and ν was arbitrary we have that $(\mathcal{P}_p(X), d_{W^p})$ satisfies the definition of a geodesic space given by Definition 7.10.

Chapter 8

Gradient Flows in Wasserstein Spaces

Section references: These notes follow the exposition given by Daneri and Savaré in [7].

The aim of this section is to study gradient flows of functions defined on measures. In particular, we wish to consider $\phi : \mathcal{P}_2(X) \to \mathbb{R}$ where $X \subset \mathbb{R}^d$. Before we get to gradient flows in Wasserstein spaces we first review gradient flows in Euclidean spaces. In Euclidean spaces the gradient flow is given by

$$\frac{\mathrm{d}u}{\mathrm{d}t}(t) = -\nabla\phi(u(t))$$

where $u(t) \in \mathbb{R}^d$ for each $t > 0$. In Euclidean spaces we can find equivalent characterisations of the gradient flow which we use to motivate definitions in Section 8.2 that are applicable in generic metric spaces where the gradient $\nabla\phi$ cannot in general be defined. We apply these definitions to the Wasserstein space in Section 8.3 and in particular we consider the application to the Fokker Plank equation.

8.1 Gradient Flows for Convex Functions in $\mathbb{R}^d$

Section references: In Euclidean spaces there are many references for gradient flows, I base these notes on [7, Section 1]. In particular, Proposition 8.5 is from [7, part of Theorem 1.3], Theorem 8.7 is from [7, Theorem 1.5], Proposition 8.8 is from [7, Theorem 1.3 and Proposition 1.6], and Proposition 8.9 is from [7, Theorem 1.3 and Proposition 1.7].

In this section we will let $\langle\cdot,\cdot\rangle$ be the Euclidean inner product and $|\cdot|$ be the associated norm. We start of by recalling gradient flows in Euclidean spaces followed by the method of minimising movements. We then find equivalent characterisations of gradient flows in Euclidean spaces which we use to motivate definitions of gradient flows in metric spaces.

8.1.1 Formulation

We assume $\phi : \mathbb{R}^d \to \mathbb{R}$ is C^2 and λ-convex, where λ-convexity is defined as follows.

Definition 8.1. *We say $\phi : \mathbb{R}^d \to \mathbb{R}$ is λ-convex if for all $t \in [0,1]$, $x_0, x_1 \in \mathbb{R}^d$ we have*

$$\phi(x_t) \leq (1-t)\phi(x_0) + t\phi(x_1) - \frac{\lambda}{2}t(1-t)|x_0 - x_1|^2$$

where $x_t = (1-t)x_0 + tx_1$.

In fact, there are several equivalent ways to characterise λ-convexity when ϕ is C^2.

Proposition 8.2. *If $\phi : \mathbb{R}^d \to \mathbb{R}$ is C^2 then the following are equivalent:*

1. (Hessian inequality) *for all $x \in \mathbb{R}^d$,*

$$D^2\phi(x) \geq \lambda \mathrm{Id}, \qquad \textit{i.e.}\ \langle D^2\phi(x)\xi, \xi\rangle \geq \lambda|\xi|^2,\ \forall \xi \in \mathbb{R}^d;$$

2. (λ-monotonicity of $\nabla\phi$) *for all $t \in [0,1]$, $x_0, x_1 \in \mathbb{R}^d$,*

$$\langle \nabla\phi(x_0) - \nabla\phi(x_1), x_1 - x_0\rangle \geq \lambda|x_0 - x_1|^2;$$

3. (λ-convexity inequality) *for all $t \in [0,1]$, $x_0, x_1 \in \mathbb{R}^d$,*

$$\phi(x_t) \leq (1-t)\phi(x_0) + t\phi(x_1) - \frac{\lambda}{2}t(1-t)|x_0 - x_1|^2$$

where $x_t = (1-t)x_0 + tx_1$;

4. (subgradient inequality) *for all $x_0, x_1 \in \mathbb{R}^d$*

$$\langle \nabla\phi(x_1), x_1 - x_0\rangle - \frac{\lambda}{2}|x_1 - x_0|^2 \geq \phi(x_1) - \phi(x_0) \geq \langle \nabla\phi(x_0), x_1 - x_0\rangle + \frac{\lambda}{2}|x_1 - x_0|^2.$$

Remark 8.3. Since ϕ is λ-convex if and only if $\tilde{\phi} = \phi - \frac{\lambda}{2}|\cdot|^2$ is convex then any λ convex function is bounded below by a quadratic function. In particular, for any λ-convex function ϕ there exists $a \in \mathbb{R}$ and $b \in \mathbb{R}^d$ such that

$$\phi(x) \geq a + b \cdot x + \frac{\lambda}{2}|x|^2.$$

Let us now define the gradient flow in $\mathbb{R}^d$.

Definition 8.4. *A gradient flow of ϕ starting from the initial condition $u_0 \in \mathbb{R}^d$ is any curve $(0, +\infty) \ni t \mapsto u(t) \in \mathbb{R}^d$ that solves (uniquely)*

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}u(t) = -\nabla\phi(u(t)) & \textit{in } (0,+\infty) \\ \lim_{t\to 0^+} u(t) = u_0. \end{cases} \tag{8.1}$$

By ODE theory (8.1) admits a unique global solution.

To motivate the definition of a gradient flow we consider the problem of minimising a function ϕ. In particular, given an estimate $u(t)$ at time t of the minimum of ϕ we want to know how to update $u(t)$ so that it moves in the direction of the minima. Elementary calculus implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(u(t)) = \langle \nabla\phi(u(t)), \frac{\mathrm{d}u}{\mathrm{d}t}(t)\rangle.$$

Hence $\frac{\mathrm{d}u}{\mathrm{d}t}(t) = -\nabla\phi(u(t))$ is the direction of steepest descent.

We recall some basic properties of gradient flows.

Proposition 8.5. *Let $\phi \in C^2(\mathbb{R}^d)$ be λ-convex and $u_0 \in \mathbb{R}^d$. Assume $(0,+\infty) \ni t \mapsto u(t) \in \mathbb{R}^d$ solves* (8.1) *then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(u(t)) = -|\nabla\phi(u(t))|^2.$$

Moreover, if v satisfies

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}v(t) = -\nabla\phi(v(t)) \quad \textit{in } (0,+\infty) \\ \lim_{t\to 0^+} v(t) = v_0. \end{cases}$$

then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{2\lambda t}|u(t)-v(t)|^2\right) \leq 0. \tag{8.2}$$

Note that if $\lambda > 0$ then the minimiser is unique and moreover (8.2) implies that if u^* is the unique minimisers then

$$|u(t)-u^*| \leq e^{-\lambda t}|u(0)-u^*|.$$

8.1.2 Minimising Movement Scheme

We want to construct a discrete approximation to the scheme

$$\begin{cases} \frac{\mathrm{d}u}{\mathrm{d}t}(t) & = -\nabla\phi(u(t)) \\ u(0) & = u_0. \end{cases}$$

Given a time step τ we let $t_\tau^{(n)} = n\tau$ for $n = 0, 1, \dots$ be a set of discrete times and we look for a sequence $U_\tau^{(n)}$ such that $U_\tau^{(n)} \approx u(t_\tau^{(n)})$. A natural method to construct $U_\tau^{(n)}$ is via the implicit Euler scheme:

$$\frac{U_\tau^{(n)} - U_\tau^{(n-1)}}{\tau} = -\nabla\phi(U_\tau^{(n)}), \quad n = 1, 2, \dots. \tag{8.3}$$

Remark 8.6. The sequence $\{U_\tau^{(n)}\}_{n=1}^\infty$ solve (8.3) if and only if each $U_\tau^{(n)}$ minimises $\Phi(\cdot; U_\tau^{(n)})$ where

$$\Phi(U; U_\tau^{(n)}) = \frac{1}{2\tau}|U - U_\tau^{(n-1)}|^2 + \phi(U).$$

Since Φ admits a unique minimiser whenever $\frac{1}{\tau} + \lambda > 0$ then the $U_\tau^{(n)}$ is uniquely defined (this follows as Φ is $(\frac{1}{\tau} + \lambda$-convex).

We let U_τ be the piecewise linear interpolation of $\{U_\tau^{(n)}\}_{n\in\mathbb{N}}$, i.e.

$$U_\tau(t) := \frac{t - t_\tau^{(n-1)}}{\tau} U_\tau^{(n)} + \frac{t_\tau^{(n)} - t}{\tau} U_\tau^{(n-1)}, \quad \text{for } t \in [t_\tau^{(n-1)}, t_\tau^{(n)}].$$

Theorem 8.7. *Let $\phi \in C^2(\mathbb{R}^d)$ be λ-convex, and assume that $t_\tau^{(n)} = n\tau$ and $U_\tau^{(n)}$ satisfies* (8.3) *with $\lim_{\tau\to 0^+} U_\tau^{(0)} = u_0$. Then U_τ converges uniformly on every compact set $[0,T]$ to u where u is the solution of* (8.1)*. Moreover, for each $T > 0$ there exists $C = C(\lambda, T)$ such that*

$$\sup_{t\in[0,T]} |u(t) - U_\tau(t)| \le |u_0 - U_\tau^{(0)}| + C|\nabla\phi(u_0)|\tau.$$

For the proof of the theorem see the fourth example sheet.

8.1.3 Metric Characterisation of Gradient Flows

We first state two equivalent characterisations of gradient flows in Euclidean spaces.

Proposition 8.8. *Assume $\phi \in C^1(\mathbb{R}^d)$. A C^1 curve $u : [0,\infty) \to \mathbb{R}^d$ with $u(0) = u_0$ satisfies* (8.1) *if and only if it satisfies the energy dissipation equality:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(u(t)) = -\frac{1}{2}\left|\frac{\mathrm{d}u}{\mathrm{d}t}(t)\right|^2 - \frac{1}{2}\left|\nabla\phi(u(t))\right|^2 \quad \forall t \in (0,\infty), \tag{8.4}$$

or in integral form:

$$\phi(u(t)) + \frac{1}{2}\int_0^t \left|\frac{\mathrm{d}u}{\mathrm{d}t}(r)\right|^2 + \left|\nabla\phi(u(r))\right|^2 \,\mathrm{d}r = \phi(u_0) \quad \forall t \in (0,\infty).$$

Proof. Let $u \in C^1([0,+\infty);\mathbb{R}^d)$ satisfy (8.1) then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\phi(u(t)) &= \left\langle \nabla\phi(u(t)), \frac{\mathrm{d}u}{\mathrm{d}t}(t) \right\rangle \\
&= \frac{1}{2}\left\langle \nabla\phi(u(t)), \frac{\mathrm{d}u}{\mathrm{d}t}(t) \right\rangle + \frac{1}{2}\left\langle \nabla\phi(u(t)), \frac{\mathrm{d}u}{\mathrm{d}t}(t) \right\rangle \\
&= -\frac{1}{2}\left|\frac{\mathrm{d}u}{\mathrm{d}t}(t)\right|^2 - \frac{1}{2}|\nabla\phi(u(t))|^2
\end{aligned}$$

where we substitute $\frac{\mathrm{d}u}{\mathrm{d}t}(t) = \nabla\phi(u(t))$ in the final equality.

On the other hand if u solves (8.4) then

$$\begin{aligned}
\frac{1}{2}\left|\frac{\mathrm{d}u}{\mathrm{d}t}(t) + \nabla\phi(u(t))\right|^2 &= \frac{1}{2}\left|\frac{\mathrm{d}u}{\mathrm{d}t}(t)\right|^2 + \frac{1}{2}|\nabla\phi(u(t))|^2 + \left\langle \frac{\mathrm{d}u}{\mathrm{d}t}(t), \nabla\phi(u(t)) \right\rangle \\
&= \frac{1}{2}\left|\frac{\mathrm{d}u}{\mathrm{d}t}(t)\right|^2 + \frac{1}{2}|\nabla\phi(u(t))|^2 + \frac{\mathrm{d}}{\mathrm{d}t}\phi(u(t)) \\
&= 0.
\end{aligned}$$

Hence $\frac{\mathrm{d}u}{\mathrm{d}t}(t) = \nabla\phi(u(t))$. $\square$

Proposition 8.9. *Assume $\phi \in C^2(\mathbb{R}^d)$ is λ-convex. A C^1 curve $u : [0,\infty) \to \mathbb{R}^d$ with $u(0) = u_0$ satisfies* (8.1) *if and only if it satisfies the evolution variational inequality:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}|u(t)-v|^2 + \frac{\lambda}{2}|u(t)-v|^2 \leq \phi(v) - \phi(u(t)) \quad \forall t \in (0,\infty) \textit{ and } \forall v \in \mathbb{R}^d. \tag{8.5}$$

Proof. Assume u satisfies (8.1) then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}|u(t)-v|^2 &= \left\langle u(t)-v, \frac{\mathrm{d}u}{\mathrm{d}t}(t)\right\rangle \\ &= -\left\langle u(t)-v, \nabla\phi(u(t))\right\rangle \\ &\leq \phi(v) - \phi(u(t)) - \frac{\lambda}{2}|v-u(t)|^2\end{aligned}$$

by Proposition 8.2(4).

On the other hand, if u satisfies (8.5) then

$$\left\langle u(t)-v, \frac{\mathrm{d}u}{\mathrm{d}t}(t)\right\rangle + \frac{\lambda}{2}|u(t)-v|^2 \leq \phi(v) - \phi(u(t)) \quad \forall t \in (0,\infty) \text{ and } \forall v \in \mathbb{R}^d.$$

Letting $v = u(t) + \varepsilon\xi$ where $\varepsilon > 0$ and $\xi \in \mathbb{R}^d$ we have

$$-\left\langle \xi, \frac{\mathrm{d}u}{\mathrm{d}t}(t)\right\rangle + \frac{\lambda}{2}\varepsilon|\xi|^2 \leq \frac{\phi(u(t)+\varepsilon\xi) - \phi(u(t))}{\varepsilon} \quad \forall t \in (0,\infty), \forall \varepsilon > 0 \text{ and } \forall \xi \in \mathbb{R}^d.$$

Letting $\varepsilon \to 0^+$ implies

$$-\left\langle \xi, \frac{\mathrm{d}u}{\mathrm{d}t}(t)\right\rangle \leq \langle \nabla\phi(u(t)), \xi\rangle \quad \forall t \in (0,\infty) \text{ and } \forall \xi \in \mathbb{R}^d.$$

Hence, $\frac{\mathrm{d}u}{\mathrm{d}t}(t) = -\nabla\phi(u(t))$ for all $t \in (0,\infty)$ as required. □

In the next section the two propositions define two gradient flows (the two gradient flows may not be in general equivalent).

8.2 Gradient Flows in Metric Spaces

Section References: I base these notes on [7, Section 2]. In particular, Theorem 8.10 is from [2, Theorem 1.1.2], Proposition 8.13 is from [7, part of Theorem 2.6], Proposition 8.15 is from [7, Theorem 2.5], Theorem 8.17 is from [7, Theorem 2.11], Theorem 8.19 is from [8, Theorem 1.1], Theorem 8.21 is from [17, Theorem 2.10], and Theorem 8.22 is from [17, Theorem 5.5].

We recall from Definition 7.8 that $\omega : (a,b) \to \mathcal{Z}$ is an absolutely continuous curve in a metric space $(\mathcal{Z}, d)$ if there exists $g \in L^1((a,b))$ such that

$$d(\omega(t_0), \omega(t_1)) \leq \int_{t_0}^{t_1} g(s)\,\mathrm{d}s \qquad \forall a < t_0 < t_1 < b. \tag{8.6}$$

When $g \in L^p((a,b))$ then we write $\omega \in \mathrm{AC}^p((a,b),\mathcal{Z})$.

For absolutely continuous curves we can define the metric derivative by

$$|\omega'|(t) = \lim_{s\to t} \frac{d(\omega(s),\omega(t))}{|t-s|} \tag{8.7}$$

which by the following theorem exists (proof not included but can be found in [2, Theorem 1.1.2]).

Theorem 8.10. *Let $(\mathcal{Z},d)$ be a complete and separable metric space. If $\omega : (a,b) \to \mathcal{Z}$ is absolutely continuous then the limit in* (8.7) *exists for $\mathcal{L}$-a.e. $t \in (a,b)$. Moreover the function $t \mapsto |\omega'|(t)$ is in $L^1((a,b))$, is admissible in* (8.6)*, i.e. $d(\omega(t_0),\omega(t_1)) \leq \int_{t_0}^{t_1} |\omega'|(s)\,\mathrm{d}s$ for all $t_0,t_1 \in (a,b)$, and if $g \in L^1((a,b))$ is any other function satisfying* (8.6) *then $|\omega'|(t) \leq g(t)$ for $\mathcal{L}$-a.e. $t \in (a,b)$.*

Let us also define the metric slope.

Definition 8.11. *Let $(\mathcal{Z},d)$ be a metric space. The metric slope of $\phi : \mathcal{Z} \to \mathbb{R}$ at $v \in \mathcal{Z}$ is defined to be*

$$|\partial\phi|(v) := \limsup_{w\to v} \frac{(\phi(v)-\phi(w))_+}{d(w,v)}.$$

8.2.1 Evolution Variational Inequality Gradient Flows

Our first definition of a gradient flow in a metric space comes from solutions to (8.5).

Definition 8.12. *Given a metric space $(\mathcal{Z},d)$ and a function $\phi : \mathcal{Z} \to \mathbb{R}\cup\{+\infty\}$ a evolution variational inequality (EVI$_\lambda$) gradient flow is a locally absolutely continuous curve $(0,+\infty) \ni t \mapsto u(t) \in \mathrm{Dom}(\phi)$ satisfying*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} d^2(u(t),v) + \frac{\lambda}{2} d^2(u(t),v) \leq \phi(v) - \phi(u(t)) \qquad \textit{for a.e. } t \in (0,\infty) \textit{ and } \forall v \in \mathrm{Dom}(\phi).$$

Similarly to Proposition 8.5 we have the following exponential convergence result.

Proposition 8.13. *Let $(\mathcal{Z},d)$ be a complete and separable metric space and $\phi : \mathcal{Z} \to \mathbb{R}$ a proper function. If u and v are two EVI$_\lambda$ gradient flows with initial conditions $u(0) = u_0$ and $v(0) = v_0$ then*

$$d(u(t),v(t)) \leq e^{-\lambda t} d(u_0,v_0).$$

Proof. Consider two EVI$_\lambda$ gradient flows u and v, i.e.

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} d^2(u(t),w) + \frac{\lambda}{2} d^2(u(t),w) &\leq \phi(w) - \phi(u(t)) \\
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s} d^2(v(s),w) + \frac{\lambda}{2} d^2(v(s),w) &\leq \phi(w) - \phi(v(s))
\end{aligned}$$

for a.e. $s,t \in (0,\infty)$ and for all $w \in \mathrm{Dom}(\phi)$. If we choose $w = v(s)$ in the first equation, $w = u(t)$ in the second, and sum we arrive at

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} d^2(u(t),v(s)) + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s} d^2(v(s),u(t)) \leq -\lambda d^2(u(t),v(s)).$$

Since
$$\frac{\mathrm{d}}{\mathrm{d}t}d^2(u(t),v(t)) = \frac{\mathrm{d}}{\mathrm{d}t}d^2(u(t),v(s))\lfloor_{s=t} + \frac{\mathrm{d}}{\mathrm{d}t}d^2(u(s),v(t))\lfloor_{s=t}$$
then we have
$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}d^2(u(t),v(t)) \leq -\lambda d^2(u(t),v(t)).$$
Multiplying by $e^{2\lambda t} > 0$ we have
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{2\lambda t}d^2(u(t),v(t))\right) = e^{2\lambda t}\frac{\mathrm{d}}{\mathrm{d}t}d^2(u(t),v(t)) + 2\lambda e^{2\lambda t}d^2(u(t),v(t)) \leq 0.$$
Integrating the above equation implies the result. □

Remark 8.14. If $\lambda > 0$ then the above proposition implies that we have at most one EVI$_\lambda$ gradient flow for a given initial condition, and exponential convergence to the minimiser, i.e. if u^* minimises ϕ then
$$d(u(t),u^*) \leq e^{-\lambda t}d(u_0,u^*).$$

It might be surprising that we did not assume that ϕ is λ-convex. In Euclidean spaces one can show that if the EVI$_\lambda$ gradient flow exists for all initial conditions $u_0 \in \mathbb{R}^d$ then ϕ is λ-convex. Conversely if ϕ is λ-convex then the EVI$_\lambda$ gradient flow exists for all initial conditions. Therefore there is an equivalence between the existence of EVI$_\lambda$ gradient flows and λ-convexity. In the results in the metric space setting we assume that the EVI$_\lambda$ gradient flow exists.

We can reformulate the above EVI$_\lambda$ gradient flow without using derivatives as follows.

Proposition 8.15. *Let $(\mathcal{Z},d)$ be a complete and separable metric space and $\phi : \mathcal{Z} \to \mathbb{R}$ a proper and lower semi-continuous function. Then $u : (0,+\infty) \to \mathrm{Dom}(\phi) \subseteq \mathcal{Z}$ is an EVI$_\lambda$ gradient flows if and only if for all $0 < s < t < +\infty$ and $v \in \mathrm{Dom}(\phi)$ we have*
$$\frac{1}{2}e^{\lambda(t-s)}d^2(u(t),v) - \frac{1}{2}d^2(u(s),v) \leq E_\lambda(t-s)\left(\phi(v) - \phi(u(t))\right)$$
where $E_\lambda(t) = \frac{e^{\lambda t}-1}{\lambda}$.

The proof is an exercise.

8.2.2 Energy Dissipation Equality Gradient Flows

We now come to our second form of gradient flows in a metric space.

Definition 8.16. *Let $(\mathcal{Z},d)$ be a metric space and assume $\phi : \mathcal{Z} \to \mathbb{R}$ has a metric slope. We say that $u \in \mathrm{AC}^2_{\mathrm{loc}}((0,+\infty),\mathcal{Z})$ is an energy dissipation equality (EDE) gradient flow if u satisfies*
$$\frac{1}{2}\int_s^t |u'|^2(r)\,\mathrm{d}r + \frac{1}{2}\int_s^t |\partial\phi|^2(u(r)) = \phi(u(s)) - \phi(u(t))$$
for all $0 < s \leq t < +\infty$.

The next theorem asserts that when the EVI$_\lambda$ and EDE gradient flows both exist then they are equal.

Theorem 8.17. *Let $(\mathcal{Z}, d)$ be a complete and separable metric space and assume $\phi : \mathcal{Z} \to \mathbb{R}$ is proper and lower semi-continuous function with a metric slope. Let $u \in \mathrm{AC}^2_{\mathrm{loc}}((0, +\infty), \mathcal{Z})$ be a EDE gradient flow of ϕ and v be the EVI$_\lambda$ gradient flow. If $u(0) = v(0)$ then $u \equiv v$.*

The proof is omitted.

8.2.3 Minimising Movements Gradient Flows

Our third and final definition of gradient flows is via the minimising movements variational scheme. The idea is to use the discrete approximation in Section 8.1.2 to define a gradient flow. We start by defining the minimising movement scheme in metric spaces.

Definition 8.18. *Given $\tau > 0$, a proper lower semi-continuous function ϕ and $U_\tau^{(0)} \in \mathrm{Dom}(\phi)$ the τ-discrete minimising movement starting from $U_\tau^{(0)}$ is any sequence $\{U_\tau^{(n)}\}_{n\in\mathbb{N}}$ in $\mathrm{Dom}(\phi)$ satisfying*

$$U_\tau^{(n)} \in \operatorname*{argmin}_V \Phi(V; U_\tau^{(n-1)}), \quad n = 1, 2, \dots$$

where $\Phi(U; U_\tau^{(n-1)}) = \frac{1}{2\tau} d^2(U, U_\tau^{(n-1)}) + \phi(U)$.

A τ-discrete solution $\bar{U}_\tau$ is a piecewise constant interpolation of a τ-discrete minimising movement defined by

$$\bar{U}_\tau(0) = U_\tau^{(0)}, \qquad \bar{U}_\tau(t) = U_\tau^{(n)} \text{ if } t \in (t_\tau^{(n-1)}, t_\tau^{(n)}], n \geq 1$$

where $t_\tau^{(n)} = n\tau$.

To define the minimising movement scheme we need to know that minimisers of Φ exist. This is often done by the direct method from the calculus of variations ("compactness plus lower semi-continuity implies the existence of minimisers"). Another approach which uses fewer assumptions is through Ekeland's variational principle which we state now.

Theorem 8.19. Ekeland's Variational Principle. *Let $(\mathcal{Z}, d)$ be a complete and separable metric space and $\Phi : \mathcal{Z} \to (-\infty, +\infty]$ a lower semi-continuous, proper function that is bounded from below. Then for all $\varepsilon > 0$, $\eta > 0$ and $x \in \mathrm{Dom}(\phi)$ satisfying $\Phi(x) \leq \inf \Phi + \varepsilon$ there exists $y \in \mathrm{Dom}(\Phi)$ such that*

$$\begin{aligned}
&\Phi(y) \leq \Phi(x) \\
&d(x, y) \leq \eta \\
&\Phi(y) < \Phi(z) + \eta d(z, y) \qquad\qquad \forall z \in \mathcal{Z} \setminus \{y\}.
\end{aligned}$$

Now the idea is to approximate the minimising movement scheme with another scheme that is easier to analyse. To this end we define the relaxed minimising movement scheme as follows.

Definition 8.20. *Given $\tau > 0$ $\eta \geq 0$, a proper lower semi-continuous function ϕ and $U^{(0)}_{\tau,\eta} \in \mathrm{Dom}(\phi)$ the (τ,η)-discrete minimising movement starting from $U^{(0)}_{\tau,\eta}$ is any sequence $\{U^{(n)}_{\tau,eta}\}_{n\in\mathbb{N}}$ in $\mathrm{Dom}(\phi)$ satisfying*

$$\Phi(U^{(n)}_{\tau,\eta}; U^{(n-1)}_{\tau,\eta}) \leq \Phi(V; U^{(n-1)}_{\tau}) + \frac{\eta}{2} d(U^{(n)}_{\tau,\eta}, U^{(n-1)}_{\tau,\eta}) d(V, U^{(n)}_{\tau,\eta}), \quad \forall V \in \mathrm{Dom}(\phi),$$
$$\text{and } \Phi(U^{(n)}_{\tau,\eta}; U^{(n-1)}_{\tau,\eta}) \leq \phi(U^{(n-1)}_{\tau,\eta})$$

for all $n = 1, 2, \ldots$ and where $\Phi(U; U^{(n-1)}_{\tau}) = \frac{1}{2\tau} d^2(U, U^{(n-1)}_{\tau}) + \phi(U)$.

A (τ,η)-discrete solution $\bar{U}_{\tau,\eta}$ is a piecewise constant interpolation of a (τ,η)-discrete minimising movement defined by

$$\bar{U}_{\tau,\eta}(0) = U^{(0)}_{\tau,\eta}, \qquad \bar{U}_{\tau,\eta}(t) = U^{(n)}_{\tau,\eta} \text{ if } t \in (t^{(n-1)}_{\tau}, t^{(n)}_{\tau}], n \geq 1$$

where $t^{(n)}_{\tau} = n\tau$.

Note that a $(\tau, 0)$-discrete solution (respectively a $(\tau, 0)$-discrete minimising movement) is a τ-discrete solution (respectively a τ-discrete minimising movement). For the relaxed problem we can infer the existence of minimisers.

Theorem 8.21. *Let $(\mathcal{Z}, d)$ be a complete and separable metric space and ϕ a lower semi-continuous, proper function that is quadratically bounded from below, i.e. there exists $\kappa_0, \phi_0 \in \mathbb{R}$ and $x_0 \in \mathcal{Z}$ such that*

$$\phi(x) + \frac{\kappa_0}{2} d^2(x, x_0) \geq \phi_0 \quad \forall x \in \mathcal{Z}.$$

Then for every $\eta > 0$ and every $\tau > 0$ with $\tau\kappa_0 < 1$ the relaxed minimising movement scheme admits at least a (τ,η)-discrete minimising movement $\{U^{(n)}_{\tau,\eta}\}_{n\in\mathbb{N}}$.

Proof. First we claim that $\Phi(\cdot; U^{(n-1)}_{\tau,\eta})$ is bounded from below. Indeed,

$$\begin{aligned}
\Phi(V; U^{(n-1)}_{\tau,\eta}) &= \frac{1}{2\tau} d^2(V, U^{(n-1)}_{\tau,\eta}) + \phi(V) \\
&\geq \frac{1}{2\tau} d^2(V, U^{(n-1)}_{\tau,\eta}) + \phi_0 - \frac{\kappa_0}{2} d^2(V, x_0) \\
&\geq \frac{1}{2\tau} d^2(V, U^{(n-1)}_{\tau,\eta}) + \phi_0 - \frac{\kappa_0}{2} \left(d(V, U^{(n-1)}_{\tau,\eta}) + d(U^{(n-1)}_{\tau,\eta}, x_0) \right)^2 \\
&\geq \frac{1}{2\tau} d^2(V, U^{(n-1)}_{\tau,\eta}) + \phi_0 - \frac{\kappa_0}{2} \left(1 + \frac{\varepsilon}{2} \right) d^2(V, U^{(n-1)}_{\tau,\eta}) \\
&\qquad\qquad - \frac{\kappa_0}{2} \left(1 + \frac{1}{2\varepsilon} \right) d^2(U^{(n-1)}_{\tau,\eta}, x_0) \\
&= \frac{1}{2} \left(\frac{1}{\tau} - \kappa_0 \left(1 + \frac{\varepsilon}{2} \right) \right) d^2(V, U^{(n-1)}_{\tau,\eta}) + \phi_0 - \frac{\kappa_0}{2} \left(1 + \frac{1}{2\varepsilon} \right) d^2(U^{(n-1)}_{\tau,\eta}, x_0)
\end{aligned}$$

where on the penultimate line we use Young's inequality which holds for any $\varepsilon > 0$. If $\tau\kappa_0 \leq 0$ then

$$\Phi(V; U^{(n-1)}_{\tau,\eta}) \geq \phi_0 - \frac{\kappa_0}{2} \left(1 + \frac{1}{2\varepsilon} \right) d^2(U^{(n-1)}_{\tau,\eta}, x_0)$$

for any $\varepsilon > 0$. If $\tau\kappa_0 \in (0,1)$ then let $\delta = 1 - \tau\kappa_0$ and note that $\delta \in (0,1)$. If we choose $\varepsilon = \frac{2\delta}{1-\delta} > 0$ then elementary algebra reveals

$$\frac{1}{\tau} - \kappa_0 \left(1 + \frac{\varepsilon}{2}\right) = 0$$

and therefore

$$\Phi(V; U_{\tau,\eta}^{(n-1)}) \geq \phi_0 - \frac{\kappa_0}{2}\left(1 + \frac{1}{2\varepsilon}\right) d^2(U_{\tau,\eta}^{(n-1)}, x_0).$$

In both cases ($\tau\kappa_0 \leq 0$ and $\tau\kappa_0 \in (0,1)$) we have that $\Phi(\cdot; U_{\tau,\eta}^{(n-1)})$ is bounded from below.

Assuming that $U_{\tau,\eta}^{(n-1)} \in \mathrm{Dom}(\phi)$ we have that $\Phi(\cdot; U_{\tau,\eta}^{(n-1)})$ is proper (since $\Phi(0; U_{\tau,\eta}^{(n-1)}) = \phi(U_{\tau,\eta}^{(n-1)}) < +\infty$) and by assumptions on ϕ we have that $\Phi(\cdot; U_{\tau,\eta}^{(n-1)})$ is lower semi-continuous. Hence we may apply Ekeland's variational principle.

If $U_{\tau,\eta}^{(n-1)}$ minimises $\Phi(\cdot; U_{\tau,\eta}^{(n-1)})$ then we pick $U_{\tau,\eta}^{(n)} = U_{\tau,\eta}^{(n-1)}$ and clearly $(U_{\tau,\eta}^{(n)}, U_{\tau,\eta}^{(n-1)})$ satisfy the conditions in Definition 8.20.

On the other hand, if $U_{\tau,\eta}^{(n-1)}$ does not minimise $\Phi(\cdot; U_{\tau,\eta}^{(n-1)})$ then

$$\varepsilon = 2(\phi(U_{\tau,\eta}^{(n-1)}) - \inf_V \Phi(V; U_{\tau,\eta}^{(n-1)})) > 0.$$

Let $\eta_k \to 0^+$ then by Ekeland's variational principle there exists a sequence $V_k \in \mathrm{Dom}(\Phi(\cdot; U_{\tau,\eta}^{(n-1)}))$ such that

$$\begin{aligned} &\Phi(V_k; U_{\tau,\eta}^{(n-1)}) \leq \Phi(U_{\tau,\eta}^{(n-1)}; U_{\tau,\eta}^{(n-1)}) = \phi(U_{\tau,\eta}^{(n-1)}), \qquad \text{and} \\ &\Phi(V_k; U_{\tau,\eta}^{(n-1)}) < \Phi(W; U_{\tau,\eta}^{(n-1)}) + \eta_k d(W, V_k) \quad \forall W \in \mathcal{Z} \setminus \{V_k\}. \end{aligned}$$

We claim that there exists K_1 and $\delta > 0$ such that $d(V_k, U_{\tau,\eta}^{(n-1)}) \geq \delta$ for all $k \geq K_1$. Indeed, if there exists a subsequence $k_m \to \infty$ such that $d(V_{k_m}, U_{\tau,\eta}^{(n-1)}) \to 0$ then

$$\phi(U_{\tau,\eta}^{(n-1)}) \leftarrow \phi(V_{k_m}) + \frac{\eta}{2} d^2(V_{k_m}, U_{\tau,\eta}^{(n-1)}) \leq \Phi(W) + \eta_{k_m} d(W, V_{k_m}) \to \Phi(W; U_{\tau,\eta}^{(n-1)})$$

for all $W \in \mathrm{Dom}(\Phi)$ and therefore $U_{\tau,\eta}^{(n-1)}$ minimises $\Phi(\cdot; U_{\tau,\eta}^{(n-1)})$, a contradiction. Hence there can be no subsequence $k_m \to \infty$ such that $d(V_{k_m}, U_{\tau,\eta}^{(n-1)}) \to 0$. In particular, there exists some K_1 and $\delta > 0$ such that $d(V_k, U_{\tau,\eta}^{(n-1)}) \geq \delta$ for all $k \geq K_1$.

Let K_2 be such that $2\eta_k \leq \delta\eta$ for all $k \geq K_2$ and choose $U_{\tau,\eta}^{(n)} = V_K$ where $K = \max\{K_1, K_2\}$. We have

$$\Phi(U_{\tau,\eta}^{(n)}; U_{\tau,\eta}^{(n-1)}) \leq \phi(U_{\tau,\eta}^{(n-1)})$$

and, for all $W \in \mathcal{Z}$,

$$\begin{aligned} \Phi(U_{\tau,\eta}^{(n)}; U_{\tau,\eta}^{(n-1)}) &\leq \Phi(W; U_{\tau,\eta}^{(n-1)}) + \eta_K d(W, U_{\tau,\eta}^{(n)}) \\ &\leq \Phi(W; U_{\tau,\eta}^{(n-1)}) + \frac{\delta\eta}{2} d(W, U_{\tau,\eta}^{(n)}) \\ &\leq \Phi(W; U_{\tau,\eta}^{(n-1)}) + \frac{\eta}{2} d(U_{\tau,\eta}^{(n-1)}, U_{\tau,\eta}^{(n)}) d(W, U_{\tau,\eta}^{(n)}). \end{aligned}$$

Hence, $(U_{\tau,\eta}^{(n)}, U_{\tau,\eta}^{(n-1)})$ satisfy the conditions in Definition 8.20. $\square$

Let us now compare the limit, when $\tau \to 0^+$, of the (τ,η)-discrete solutions with the EVI$_\lambda$ gradient flow. In the following theorem geodesically convex is defined as follows: we say $D \subset \mathcal{Z}$ is geodesically convex if for all $x_1, x_2 \in D$ there exists a geodesics between x_1 and x_2 that is contained in D.

Theorem 8.22. *Let $(\mathcal{Z}, d)$ be a complete and separable metric space and ϕ be a lower semi-continuous, proper function such that $\overline{\mathrm{Dom}(\phi)}$ is geodesically convex. Let $\tau > 0$, $\eta \geq 0$ and $\lambda > 0$ satisfy $\eta - \lambda < \frac{1}{2\tau}$. Assume that the EVI$_\lambda$ gradient flow $u : (0, +\infty) \to \mathcal{Z}$ exists and $\bar{U}_{\tau,\eta}$ be a (τ,η)-discrete solution starting from $\bar{U}_{\tau,\eta}(0) = u(0)$. Then $\bar{U}_{\tau,\eta}$ converges uniformly to u on compact intervals.*

We omit the proof of the theorem.

8.3 Gradient Flows in the Wasserstein Space

Section References: the majority of this section is based on [7, Section 4] although we often refer to the book [2] which contains similar results with proofs. In particular, Theorem 8.23 is from [2, Theorem 8.3.1], Propositions 8.25 and 8.26 are from [2, Proposition 8.4.5 and Proposition 8.4.6], Theorem 8.28 is from [7, Theorem 4.11], Theorem 8.29 can be found in either [4, Proposition 1.1] or [23, Theorem 8.1], and Theorem 8.30 can be found in [7, Theorem 4.20].

We now look at applying the gradient flow formulation of the previous section to the Wasserstein space $(\mathcal{P}_p(X), d_{W^p})$ which we recall from Chapter 7 is a metric space. This section is broken into four parts. In the first part we define the Wasserstein tangent space. The second part is really a diversion, albeit to a very important result. Motivated by the continuity equation we derive the Benamou and Brenier formulation of optimal transport. This recasts the optimal transport problem into a fluid mechanics framework which can lead to numerical approaches (not pursued in these notes). The third part contains the important application of the gradient flows in Wasserstein space of the Fokker-Plank equation (other examples are also readily derived, see for example [7, Section 4.4]). The final part uses the method of minimising movements to derive the same evolutionary PDE's as in the second part.

In this chapter we will use the continuity equation. In terms of mass the continuity equation can be derived as follows. Assume we have a density $\rho(x,t)$ at time t. Then each $A \subset \mathbb{R}^d$ has mass $\int_A \rho(x,t)\,\mathrm{d}x$ at time t. Moreover, if we assume that mass is only lost through the boundary of A, and that mass is moving with speed $v(x,t) \cdot n(x,t)$ where n is the unit normal to the boundary of A then we have

$$\int_A \frac{\partial \rho}{\partial t}(x,t)\,\mathrm{d}x = -\int_{\partial A} v(x,t) \cdot n(x,t)\rho(x,t)\,\mathrm{d}x.$$

By the divergence theorem

$$\int_A \frac{\partial \rho}{\partial t}(x,t)\,\mathrm{d}x = -\int_A \nabla \cdot (v(x,t)\rho(x,t))\,\mathrm{d}x.$$

So if we assume this is true for all $A \subset \mathbb{R}^d$ we have that

$$\frac{\partial \rho}{\partial t}(x,t) + \nabla \cdot (v(x,t)\rho(x,t)) = 0.$$

The above equation is called the *continuity equation*. As we will consider trajectories in the space of probability measures (where mass is conserved) the continuity equation will play an important role.

In the rest of this chapter we change notation slightly. Previously we denoted our trajectories by $u(t) \in \mathcal{Z}$, now that $\mathcal{Z} = \mathcal{P}_2(\mathbb{R}^d)$ and we wish to avoid writing $u(t) \in \mathcal{P}_2(\mathbb{R}^d)$ and in particular to avoid notation such as $u(t)(A)$ where $A \subset \mathbb{R}^d$ we will use a subscript to denote the t dependence, i.e. $u_t = u(t)$. Secondly, to be consistent with the rest of these notes we go back to using Greek letters for probability measures, therefore we use μ_t where in Sections 8.1 and 8.2 we had written $u(t)$.

8.3.1 Wasserstein Tangent Spaces

If we consider a curve in the space $\mathcal{P}_2(\mathbb{R}^d)$, i.e. $(0,+\infty) \ni t \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$, then it is natural to think of this in terms of particles rearranging in $\mathbb{R}^d$ under a conservation of mass assumption, hence the continuity equation should hold.

Theorem 8.23. Absolutely continuous curves and the continuity equation. *The following holds.*

(1) Let $(0,+\infty) \ni t \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve and let $|\mu'| \in L^1((0,+\infty))$ be its metric derivative. Then there exists a vector field $v_t \in L^2(\mu_t,\mathbb{R}^d)$ such that

$$\|v_t\|_{L^2(\mu_t;\mathbb{R}^d)} \leq |\mu'|(t) \quad \text{for } \mathcal{L}^1\text{-a.e. } t \in (0,+\infty)$$

and

$$\frac{\partial}{\partial t}\mu_t + \nabla \cdot (v_t \mu_t) = 0 \quad \text{in } \mathbb{R}^d \times (0,+\infty) \tag{8.8}$$

where the above holds in the sense of distributions (see below).

(2) Let $(0,+\infty) \ni t \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ be continuous with respect to the weak topology on $\mathcal{P}(\mathbb{R}^d)$ and satisfies* (8.8) *in the sense of distributions for some vector field v_t with*

$$\int_0^\infty \|v_t\|_{L^2(\mu_t;\mathbb{R}^d)} \,\mathrm{d}t < +\infty$$

then μ_t is an absolutely continuous curve and $|\mu'|(t) \leq \|v_t\|_{L^2(\mu_t;\mathbb{R}^d)}$ for $\mathcal{L}^1$-a.e. $t \in (0,+\infty)$.

The proof is omitted from the notes but can be found in [2, Theorem 8.3.1]. When we say (8.8) holds in the sense of distributions we mean that for any test function $f \in C_c^\infty(\mathbb{R}^d \times (0,\infty))$ it holds that

$$\int_0^\infty \int_{\mathbb{R}^d} \left(\frac{\partial f}{\partial t}(x,t) + v_t(x) \cdot \nabla f(x,t) \right) \mathrm{d}\mu_t(x)\,\mathrm{d}t = 0.$$

Remark 8.24. Let μ_t be an absolutely continuous curve and M be the minimal norm of v_t satisfying (8.8), i.e.

$$M = \inf_{v_t \text{ satisfying (8.8)}} \|v_t\|_{L^2(\mu_t;\mathbb{R}^d)}.$$

By part (1) of the above theorem there exists a vector field $\bar{v}_t$ such that

$$M \leq \|\bar{v}_t\|_{L^2(\mu_t;\mathbb{R}^d)} \leq |\mu'|(t)$$

almost everywhere. On the other hand, if v_t is any vector field satisfying (8.8) then by part (2) of the above theorem $|\mu'|(t) \leq \|v_t\|_{L^2(\mu_t;\mathbb{R}^d)}$, hence $|\mu'|(t) \leq M$. It follows that $M = |\mu'|(t)$. Moreover we have that if v_t satisfies (8.8) then $\|v_t\|_{L^2(\mu_t;\mathbb{R}^d)} \leq |\mu'|(t)$ if and only if $\|v_t\|_{L^2(\mu_t;\mathbb{R}^d)} = |\mu'|(t)$

The next proposition allows us to characterise the vector fields v_t satisfying $\|v_t\|_{L^2(\mu_t;\mathbb{R}^d)} = |\mu'|(t)$. The proof is omitted but can be found in [2, Proposition 8.4.5].

Proposition 8.25. *Let $(0,+\infty) \ni t \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve, $|\mu'| \in L^1((0,+\infty))$ be its metric derivative, and assume $v_t \in L^2(\mu_t;\mathbb{R}^d)$ satisfies* (8.8). *Then*

$$\|v_t\|_{L^2(\mu_t;\mathbb{R}^d)} \leq |\mu'|(t)$$

for $\mathcal{L}^1$ almost every $t \in (0,+\infty)$ if and only if $v_t \in \overline{\{\nabla\varphi \,:\, \varphi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu_t;\mathbb{R}^d)}$. Moreover the vector v_t is determined uniquely by (8.8) *and $\|v_t\|_{L^2(\mu_t;\mathbb{R}^d)} \leq |\mu'|(t)$.*

The next proposition will be used to motivate the definition of the tangent space. Again the proof is omitted but can be found in [2, Proposition 8.4.6].

Proposition 8.26. *Let $(0,+\infty) \ni t \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve and assume $v_t \in \overline{\{\nabla\varphi \,:\, \varphi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu_t;\mathbb{R}^d)}$ satisfies* (8.8). *Further assume that there exists an optimal transport map $T^{(t,s)}$ between μ_t and μ_s for all s,t (where $T^{(t,s)}_{\#}\mu_t = \mu_s$). Then,*

$$\lim_{h\to 0} \frac{T^{(t,t+h)} - \mathrm{Id}}{h} = v_t$$

where the limit is taken in $L^2(\mu_t;\mathbb{R}^d)$.

Hence we can see the set $\overline{\{\nabla\varphi \,:\, \varphi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu_t;\mathbb{R}^d)}$ as the set of admissible directions in which we can perturb a measure and still be in the space of probability measures.

Definition 8.27. *Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. We define the tangent space to $\mathcal{P}_2(\mathbb{R}^d)$ at the point μ by*

$$\mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) := \overline{\{\nabla\varphi \,:\, \varphi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu_t;\mathbb{R}^d)}.$$

Let us also notice that the above proposition allows us to differentiate $t \mapsto \frac{1}{2}d_W^2(\mu_t,\sigma)$ along absolutely continuous curves μ_t.

Theorem 8.28. *Let $\sigma \in \mathcal{P}_2(\mathbb{R}^d)$, $(0,+\infty) \ni t \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve, and $v_t \in \mathrm{Tan}_{\mu_t}\mathcal{P}_2(\mathbb{R}^d)$ satisfy* (8.8). *Assume that there exists an optimal transport map $T^{(t)}$ between μ_t and σ for all t (where $T^{(t)}_{\#}\mu_t = \sigma$). Then,*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} d^2_{W^2}(\mu_t,\sigma) = \int_{\mathbb{R}^d} (x - T^{(t)}(x)) \cdot v_t(x) \,\mathrm{d}\mu_t(x).$$

Sketch Proof. The missing detail in the proof is to show that the optimal transport maps $T^{(t)}$ and $T^{(t,t+h)}$ (where $T^{(t,t+h)}$ is the optimal transport map between μ_t and μ_s as in Proposition 8.26) can be related by

$$T^{(t)} = T^{(t+h)} \circ T^{(t,t+h)}. \tag{8.9}$$

The reason why the above identity should be natural (at least for small h) is that one can show $[T^{(t+h)} \circ T^{(t,t+h)}]_{\#}\mu_t = \sigma$ and therefore the map $T^{(t+h)} \circ T^{(t,t+h)}$ is a candidate for a minimiser of the Monge problem between μ_t and σ. The fact that this map is optimal follows from the optimality of $T^{(t)}$ and $T^{(t,t+h)}$, and requires h to be sufficiently small.

Given (8.9) holds we can compute $\frac{\mathrm{d}}{\mathrm{d}t} d^2_{W^2}(\mu_t, \sigma)$ as follows

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} d^2_{W^2}(\mu_t, \sigma) &= \lim_{h\to 0^+} \frac{1}{h}\left(\int_{\mathbb{R}^d} \left|T^{(t+h)}(x) - x\right|^2 \,\mathrm{d}\mu_{t+h}(x) - \int_{\mathbb{R}^d} \left|T^{(t)}(x) - x\right|^2 \,\mathrm{d}\mu_t(x)\right) \\
&= \lim_{h\to 0^+} \frac{1}{h} \int_{\mathbb{R}^d} \left(\left|T^{(t+h)}(T^{(t,t+h)}(x)) - T^{(t,t+h)}(x)\right|^2 - \left|T^{(t)}(x) - x\right|^2 \right) \mathrm{d}\mu_t(x) \\
&= \lim_{h\to 0^+} \frac{1}{h} \int_{\mathbb{R}^d} \left(\left|T^{(t)}(x) - T^{(t,t+h)}(x)\right|^2 - \left|T^{(t)}(x) - x\right|^2 \right) \mathrm{d}\mu_t(x) \\
&= \lim_{h\to 0^+} \frac{1}{h} \int_{\mathbb{R}^d} \left(\left|T^{(t)}(x) - x + x - T^{(t,t+h)}(x)\right|^2 - \left|T^{(t)}(x) - x\right|^2 \right) \mathrm{d}\mu_t(x) \\
&= \lim_{h\to 0^+} \frac{1}{h} \int_{\mathbb{R}^d} \left(x - T^{(t,t+h)}(x)\right) \cdot \left(2T^{(t)}(x) - x - T^{(t,t+h)}(x)\right) \mathrm{d}\mu_t(x) \\
&= 2\int_{\mathbb{R}^d} \left(x - T^{(t)}(x)\right) \cdot v_t(x) \,\mathrm{d}\mu_t(x)
\end{aligned}$$

where on the penultimate line we use the elementary inequality $|a+b|^2 - |a|^2 = b \cdot (2a+b)$, and on the last line we use Proposition 8.26) to infer

$$\frac{x - T^{(t,t+h)}(x)}{h} \to -v_t(x) \qquad \text{and} \qquad T^{(t,t+h)}(x) \to x$$

as $h \to 0$. □

8.3.2 The Benamou and Brenier Fluid Mechanics Interpretation

A result closely related result to Theorem 8.23 is the Benamou and Brenier formulation of optimal transport. We recall that the metric derivative of an absolutely continuous curve μ_t satisfies $|\mu_t'| = \|v_t\|_{L^2(\mu_t)}$ where v_t satisfies (8.8). A constant speed geodesic between μ_0 and μ_1 can be found by minimising $\int_0^1 |\mu_t'|^2 \,\mathrm{d}t$, hence we can equivalently minimise $\int_0^1 \|v_t\|^2_{L^2(\mu_t)} \,\mathrm{d}t$ over v_t satisfying (8.8). The Benamou and Brenier result states that the latter minimisation problem is equal to the minimum of the Kantorovich optimal transport problem (for quadratic costs).

Theorem 8.29. Benamou and Brenier Formula. *Let $X \subset \mathbb{R}^d$ be compact and $\mu_i \in \mathcal{P}(X)$ has density ρ_i with respect to the Lebesgue measure for $i = 0, 1$. We define $V(\rho_0, \rho_1)$ to be the set of all $(\rho, v) = (\rho_t, v_t)_{t\in[0,1]}$ satisfying*

$\mu_t \in \mathcal{P}(X)$ *is the probability measure with density* ρ_t
$[0,1] \ni t \mapsto \mu_t$ *is continuous where $\mathcal{P}(X)$ is equipped with the weak* topology*
$v_t \in L^2(\mu_t)$
$\dfrac{\partial \rho_t}{\partial t} + \nabla \cdot (\rho_t v_t) = 0$ *weakly*
$\rho_{t=0} = \rho_0 \qquad \rho_{t=1} = \rho_1.$

Define $\mathbb{K}$ to be the Kantorovich cost in Definition 2.3 with $c(x, y) = |x - y|^2$. Then,

$$\min_{\pi\in\Pi(\mu_0,\mu_1}\mathbb{K}(\pi) = \inf_{(\rho,v)\in V(\rho_0,\rho_1)} \int_0^1 \int_X |v_t(x)|^2 \rho_t(x)\,\mathrm{d}x\,\mathrm{d}t.$$

Sketch Proof. Assume $(\rho, v) \in V(\rho_0, \rho_1)$. Define the trajectory $T_t(x)$ by $\frac{\partial}{\partial t}T_t(x) = v_t(T_t(x))$ and $T_0 = \mathrm{Id}$. We claim that $\mu_t = [T_t]_\#\mu_0$ where μ_t is the measure with density ρ_t. From the first exercise sheet it is enough to show that

$$\rho_0(x) = \rho_t(T_t(x)) \det(\nabla T_t(x)) \tag{8.10}$$

where we have assumed that T_t is C^1, bijective and $\det(\nabla T_t(x)) > 0$. Clearly (8.10) holds when $t = 0$. Now,

$$\begin{aligned}
\frac{\partial}{\partial t}\rho_t(T_t(x)) \det(\nabla T_t(x)) &= \left(\frac{\partial \rho_t}{\partial t}(T_t(x)) + \nabla\rho_t(T_t(x)) \cdot \frac{\partial T_t}{\partial t}(x)\right) \det(\nabla T_t(x)) \\
&\quad + \rho_t(T_t(x)) \det(\nabla T_t(x)) \mathrm{tr}\left([\nabla T_t(x)]^{-1} \frac{\partial}{\partial t}\nabla T_t(x)\right) \\
&= \det(\nabla T_t(x))\Big(\frac{\partial \rho_t}{\partial t}(T_t(x)) + \nabla\rho_t(T_t(x)) \cdot v_t(T_t(x)) \\
&\quad + \rho_t(T_t(x))\mathrm{tr}\left([\nabla T_t(x)]^{-1}\nabla v_t(T_t(x))\nabla T_t(x)\right)\Big) \\
&= \det(\nabla T_t(x))\Big(\frac{\partial \rho_t}{\partial t}(T_t(x)) + \nabla\rho_t(T_t(x)) \cdot v_t(T_t(x)) \\
&\quad + \rho_t(T_t(x))\nabla \cdot v_t(T_t(x))\Big) \\
&= \det(\nabla T_t(x))\Big(\frac{\partial \rho_t}{\partial t} + \nabla \cdot (\rho v_t)\Big)(T_t(x)) \\
&= 0
\end{aligned}$$

where we use $\frac{\partial}{\partial t}\nabla T_t(x) = \nabla \frac{\partial}{\partial t} T_t(x) = \nabla[v_t(T_t(x))] = [\nabla v_t](T_t(x))\nabla T_t(x)$ and $\mathrm{tr}(ABC) = \mathrm{tr}(ACB)$ and therefore if $A = C^{-1}$ then $\mathrm{tr}(C^{-1}BC) = \mathrm{tr}(B)$. Hence (8.10) holds for all $t \geq 0$.

We know that T_1 is admissible in the Monge optimal transport problem. Hence,

$$
\begin{aligned}
\int_0^1 \int_{\mathbb{R}^d} |v_t(x)|^2 \rho_t(x)\,\mathrm{d}x\,\mathrm{d}t &= \int_0^1 \int_{\mathbb{R}^d} |v_t(x)|^2\,\mathrm{d}[T_t]_{\#}\mu_0(x)\,\mathrm{d}t \\
&= \int_0^1 \int_{\mathbb{R}^d} |v_t(T_t(x))|^2\,\mathrm{d}\mu_0(x)\,\mathrm{d}t \\
&= \int_{\mathbb{R}^d} \int_0^1 \left|\frac{\partial T_t}{\partial t}(x)\right|^2 \mathrm{d}t\,\mathrm{d}\mu_0(x) \\
&\geq \int_{\mathbb{R}^d} \left|\int_0^1 \frac{\partial T_t}{\partial t}(x)\,\mathrm{d}t\right|^2 \mathrm{d}\mu_0(x) \quad \text{by Jensen's inequality} \\
&= \int_{\mathbb{R}^d} |T_1(x)-x|^2\,\mathrm{d}\mu_0(x) \\
&= \mathbb{M}(T_1) \\
&\geq \inf_{T\,:\,T_{\#}\mu_0=\mu_1} \mathbb{M}(T).
\end{aligned}
$$

On the other hand we define $T_t^\dagger = tT_t^\dagger + (1-t)\mathrm{Id}$, where $T^\dagger$ is the solution to Monge's optimal transport problem, $\mu_t^\dagger = [T_t^\dagger]_{\#}\mu_0$ and $v_t^\dagger$ to satisfy $v_t^\dagger(T_t^\dagger(x)) = \frac{\partial T_t^\dagger}{\partial t}(x)$. One can show $(\rho_t^\dagger, v_t^\dagger)$ satisfies $\frac{\partial \rho_t^\dagger}{\partial t} + \nabla\cdot(\rho_t^\dagger v_t^\dagger) = 0$ where $\rho_t^\dagger$ is the density of $\mu_t^\dagger$. Hence $(\rho_t^\dagger, v_t^\dagger)$ are admissible and moreover since $\frac{\partial T_t^\dagger}{\partial t}(x) = T^\dagger - \mathrm{Id}$ is independent of t then the previous calculation implies

$$
\int_0^1 \int_{\mathbb{R}^d} |v_t^\dagger(x)|^2 \rho_t^\dagger(x)\,\mathrm{d}x\,\mathrm{d}t = \mathbb{M}(T^\dagger) = \min_{T\,:\,T_{\#}\mu_0=\mu_1} \mathbb{M}(T).
$$

This implies

$$
\inf_{(\rho,v)} \int_0^1 \int_{\mathbb{R}^d} |v_t(x)|^2 \rho_t(x)\,\mathrm{d}x\,\mathrm{d}t \leq \min_{T\,:\,T_{\#}\mu_0=\mu_1} \mathbb{M}(T)
$$

which completes the proof. □

8.3.3 Gradient Flows in $(\mathcal{P}_2(\mathbb{R}^d), d_{W^2})$ and Evolutionary PDEs

The idea is to consider the gradient flow of a function $\phi : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ with respect to the Wasserstein metric to arrive at a evolutionary partial differential equation of the form

$$
\frac{\partial}{\partial t}\mu - \nabla\cdot\left(\mu\nabla\frac{\delta\phi}{\delta\mu}\right) = 0 \tag{8.11}
$$

where $\frac{\delta\phi}{\delta\mu}$ is the first variation of ϕ, i.e.

$$
\left\langle \frac{\delta\phi}{\delta\mu}, \chi \right\rangle = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\phi(\mu+\varepsilon\chi)|_{\varepsilon=0}
$$

for any χ is some suitable space. At this point it is perhaps not clear how all the terms in (8.11) are defined; this should become clear as we consider a special case. We define $\phi(\mu) = \mathcal{U}(\mu) + \mathcal{V}(\mu)$ where

$$\mathcal{U}(\mu) = \begin{cases} \int_{\mathbb{R}^d} U(\rho(x))\,\mathrm{d}x & \text{if } \mu \ll \mathcal{L}^d \text{ and where } \rho = \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d} \\ +\infty & \text{else} \end{cases} \tag{8.12}$$

and

$$\mathcal{V}(\mu) = \int_{\mathbb{R}^d} V(x)\,\mathrm{d}\mu(x). \tag{8.13}$$

So formally $\frac{\delta\phi}{\delta\mu}(\mu) = U'(\rho) + V$ when $\mu \ll \mathcal{L}^d$ with $\rho = \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d}$ (exercise). In particular, if $U(r) = r\log r$ then we have $\frac{\delta\phi}{\delta\mu}(\mu) = \log\rho + 1 + V$. Hence, if we rewrite (8.11) in terms of densities then in this special case we have

$$\frac{\partial\rho}{\partial t} - \nabla\cdot(\nabla\rho + \rho\nabla V) = \frac{\partial\rho}{\partial t} - \nabla\cdot(\rho\nabla(\log\rho + V)) = 0.$$

Using measures rather than densities we can rewrite the above as

$$\frac{\partial}{\partial t}\mu_t - \nabla\cdot(\nabla\mu_t + \mu_t\nabla V) = 0. \tag{8.14}$$

The following theorem states the existence of gradient flows of $\phi : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ in the Wasserstein space (see also [7, Theorem 4.20]).

Theorem 8.30. *Let V and U satisfy*

$$V : \mathbb{R}^d \to \mathbb{R} \quad \lambda\text{-convex function for } \lambda \in \mathbb{R},$$

$$U : [0,+\infty) \to \mathbb{R} \quad \text{convex and such that } U(0) = 0, \liminf_{s\to 0^+}\frac{U(s)}{s^\alpha} > -\infty \text{ for some } \alpha > \frac{d}{d+2},$$

$$\lim_{s\to\infty}\frac{U(s)}{s} = +\infty, s \mapsto s^d U(s^{-d}) \text{ is convex and non-increasing on } (0,+\infty).$$

Let $\phi = \mathcal{U} + \mathcal{V}$ where $\mathcal{U}$ and $\mathcal{V}$ are defined by (8.12) *and* (8.13) *respectively. Then for all $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ there exists a unique $\mu_t \in \mathrm{Lip}_{\mathrm{loc}}((0,+\infty);\mathcal{P}(\mathbb{R}^d))$ and $v_t \in \mathrm{Tan}_{\mu_t}(\mathcal{P}_2(\mathbb{R}^d))$ such that $\lim_{t\to 0^+}\mu_t = \mu_0$ in $\mathcal{P}_2(\mathbb{R}^d)$, $t \mapsto \int_{\mathbb{R}^d}|v_t|^2\,\mathrm{d}\mu_t = |\mu_t'|^2 \in L^\infty_{\mathrm{loc}}((0,+\infty))$,*

$$\frac{\partial}{\partial t}\mu_t + \nabla\cdot(\mu_t v_t) = 0 \quad \text{in } (0,+\infty)\times\mathbb{R}^d, \tag{8.15}$$

and for all $\sigma \in \mathrm{Dom}(\phi)$

$$\begin{gathered} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}d^2_{W^2}(\mu_t,\sigma) + \frac{\lambda}{2}d^2_{W^2}(\mu_t,\sigma) \leq \phi(\sigma) - \phi(\mu_t) \text{ and} \\ \int_{\mathbb{R}^d\times\mathbb{R}^d}\left(\langle v_t(x), x-y\rangle + \frac{\lambda}{2}|y-x|^2\right)\mathrm{d}\pi_t(x,y) \leq \phi(\sigma) - \phi(\mu_t) \end{gathered} \tag{8.16}$$

for any $\pi_t \in \Pi_{\mathrm{opt}}(\mu_t,\sigma)$ where $\Pi_{\mathrm{opt}}(\mu_t,\sigma)$ is the set of optimal transport plans for Kantorovich's optimal transport problem with cost $c(x,y) = |x-y|^2$ between μ_t and σ.

Let us consider an application of the above theorem to when $U(\rho) = \rho \log \rho$. Let $\zeta \in C_c^\infty(\mathbb{R}^d)$ and define $T_\varepsilon = \mathrm{Id} + \varepsilon \nabla \zeta$. It is left as an exercise to show that if $\varepsilon \max_{\mathbb{R}^d} \|D^2\zeta\| < 1$ then $\pi_\varepsilon = (\mathrm{Id} \times T_\varepsilon)_{\#}\mu_t$ is an optimal transport plan (with quadratic cost) between μ_t and $[T_\varepsilon]_{\#}\mu_t$. Choosing $\sigma = [T_\varepsilon]_{\#}\mu_t$ in (8.16) we have

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\langle v_t(x), x - y \rangle + \frac{\lambda}{2}|y-x|^2 \right) \, \mathrm{d}(\mathrm{Id} \times T_\varepsilon)_{\#}\mu_t(x,y) \leq \phi([T_\varepsilon]_{\#}\mu_t) - \phi(\mu_t).$$

Now the left hand side can be manipulated as follows

$$\begin{aligned}
&\int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\langle v_t(x), x - y \rangle + \frac{\lambda}{2}|y-x|^2 \right) \, \mathrm{d}(\mathrm{Id} \times T_\varepsilon)_{\#}\mu_t(x,y) \\
&\qquad = \int_{\mathbb{R}^d} \left(\langle v_t(x), x - T_\varepsilon(x) \rangle + \frac{\lambda}{2}|T_\varepsilon(x)-x|^2 \right) \, \mathrm{d}\mu_t(x) \\
&\qquad = \int_{\mathbb{R}^d} \left(-\varepsilon \langle v_t(x), \nabla\zeta(x) \rangle + \frac{\varepsilon^2\lambda}{2}|\nabla\zeta(x)|^2 \right) \, \mathrm{d}\mu_t(x) \\
&\qquad \geq -\varepsilon \int_{\mathbb{R}^d} \langle v_t(x), \nabla\zeta(x) \rangle \, \mathrm{d}\mu_t(x).
\end{aligned}$$

Now we assume that $\rho_t = \frac{\mathrm{d}\mu_t}{\mathrm{d}\mathcal{L}^d}$ and $\rho_t^{(\varepsilon)} = \frac{\mathrm{d}\mu_t^{(\varepsilon)}}{\mathrm{d}\mathcal{L}^d}$ where $\mu_t^{(\varepsilon)} = [T_\varepsilon]_{\#}\mu_t$. By the change of variables formula we have

$$\rho_t(x) = \rho_t^{(\varepsilon)}(T_\varepsilon(x))|\det(\nabla T_\varepsilon(x))| = \rho_t^{(\varepsilon)}(T_\varepsilon(x))|\det(\mathrm{Id} + \varepsilon D^2\zeta(x))|.$$

So,

$$\begin{aligned}
\rho(\mu_t^{(\varepsilon)}) - \phi(\mu_t) &= \int_{\mathbb{R}^d} \left(\log \rho_t^{(\varepsilon)}(x) + V(x) \right) \, \mathrm{d}\mu_t^{(\varepsilon)}(x) - \int_{\mathbb{R}^d} \left(\log \rho_t(x) + V(x) \right) \, \mathrm{d}\mu_t(x) \\
&= \int_{\mathbb{R}^d} \left(\log \rho_t^{(\varepsilon)}(T_\varepsilon(x)) - \log \rho_t(x) \right) \, \mathrm{d}\mu_t(x) + \int_{\mathbb{R}^d} \left(V(T_\varepsilon(x)) - V(x) \right) \, \mathrm{d}\mu_t(x) \\
&= -\int_{\mathbb{R}^d} \log|\det(\mathrm{Id} + \varepsilon D^2\zeta(x))| \, \mathrm{d}\mu_t(x) + \int_{\mathbb{R}^d} \left(V(T_\varepsilon(x)) - V(x) \right) \, \mathrm{d}\mu_t(x).
\end{aligned}$$

Hence,

$$\begin{aligned}
-\varepsilon \int_{\mathbb{R}^d} \langle v_t(x), \nabla\zeta(x) \rangle \, \mathrm{d}\mu_t(x) \leq &-\int_{\mathbb{R}^d} \log|\det(\mathrm{Id} + \varepsilon D^2\zeta(x))| \, \mathrm{d}\mu_t(x) \\
&+ \int_{\mathbb{R}^d} \left(V(T_\varepsilon(x)) - V(x) \right) \, \mathrm{d}\mu_t(x).
\end{aligned}$$

Dividing by $\varepsilon > 0$ and letting $\varepsilon \to 0^+$ we obtain

$$\begin{aligned}
-\int_{\mathbb{R}^d} \langle v_t(x), \nabla\zeta(x) \rangle \, \mathrm{d}\mu_t(x) \leq &-\lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^d} \frac{1}{\varepsilon} \log|\det(\mathrm{Id} + \varepsilon D^2\zeta(x))| \, \mathrm{d}\mu_t(x) \\
&+ \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^d} \frac{V(T_\varepsilon(x)) - V(x)}{\varepsilon} \, \mathrm{d}\mu_t(x).
\end{aligned}$$

Now, assume V is sufficiently differentiable we have

$$\begin{aligned}\frac{V(T_\varepsilon(x)) - V(x)}{\varepsilon} &= \frac{1}{\varepsilon}\left(\nabla V(x)\cdot(T_\varepsilon(x) - x) + \text{h.o.t.'s}\right)\\ &= \nabla V(x)\cdot\nabla\zeta(x) + \text{h.o.t.'s},\\ &\to \nabla V(x)\cdot\nabla\zeta(x)\end{aligned}$$

as $\varepsilon \to 0^+$. Let $F_D(\varepsilon) = \log|\det(\mathrm{Id} + \varepsilon D)|$ then we have $\frac{\mathrm{d}F_D}{\mathrm{d}\varepsilon}(\varepsilon) = \mathrm{tr}((\mathrm{Id} + \varepsilon D)^{-1}D)$ (see Exercise Sheet 4). Hence,

$$\begin{aligned}\frac{1}{\varepsilon}\left(F_{D^2\zeta(x)}(\varepsilon) - F_{D^2\zeta(x)}(0)\right) &= \frac{\partial F_{D^2\zeta(x)}}{\partial\varepsilon}(0) + \text{h.o.t.'s}\\ &= \mathrm{tr}(D^2\zeta(x)) + \text{h.o.t.'s}\\ &= \Delta\zeta(x) + \text{h.o.t.'s}\\ &\to \Delta\zeta(x)\end{aligned}$$

as $\varepsilon \to 0^+$. The previous calculations show

$$-\int_{\mathbb{R}^d}\langle v_t(x), \nabla\zeta(x)\rangle\,\mathrm{d}\mu_t(x) \leq \int_{\mathbb{R}^d}\left(\nabla V(x)\cdot\nabla\zeta(x) - \Delta\zeta(x)\right)\,\mathrm{d}\mu_t(x) \quad \forall\zeta\in C_c^\infty(\mathbb{R}^d).$$

Substituting $\zeta \mapsto -\zeta$ implies the inverse inequality, hence

$$-\int_{\mathbb{R}^d}\langle v_t(x), \nabla\zeta(x)\rangle\,\mathrm{d}\mu_t(x) = \int_{\mathbb{R}^d}\left(\nabla V(x)\cdot\nabla\zeta(x) - \Delta\zeta(x)\right)\,\mathrm{d}\mu_t(x) \quad \forall\zeta\in C_c^\infty(\mathbb{R}^d).$$

In weak form we can write

$$-v_t\mu_t = \mu_t\nabla V + \nabla\mu_t.$$

Substituting this into (8.15) implies

$$\frac{\partial}{\partial t}\mu_t - \nabla\cdot(\mu_t\nabla V + \nabla\mu_t) = 0$$

which we note agrees with (8.14).

8.3.4 The Minimising Movement Scheme in the Wasserstein Space

The Wasserstein minimising movement scheme is defined to be the sequence satisfying

$$\mu_\tau^{(n+1)} \in \operatorname*{argmin}_{\mu}\left\{\phi(\mu) + \frac{1}{2\tau}d_{W^2}^2(\mu, \mu_\tau^{(n)})\right\}$$

for some $\tau > 0$. Therefore, assuming differentiability of ϕ and d_{W^2}, we have that

$$\frac{\partial}{\partial\mu}\phi(\mu_\tau^{(n+1)}) + \frac{1}{2\tau}\frac{\partial}{\partial\mu}d_{W^2}^2(\mu_\tau^{(n+1)}, \mu_\tau^{(n)}) = 0. \tag{8.17}$$

The following theorem gives conditions sufficient for the differentiability of the Wasserstein distance.

Theorem 8.31. *Let $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, $c(x,y) = |x-y|^2$, define Φ_c by* (4.2) *and $\mathbb{J}$ by* (4.1)*. Assume the Kantorovich dual problem, $\max_{\Phi_c} \mathbb{J}$, has a unique maximiser $(\varphi, \varphi^c) \in \Phi_c$. Then $\mu \mapsto d^2_{W^2}(\mu, \nu)$ is differentiable at μ and*

$$\frac{1}{2}\frac{\partial}{\partial \mu} d^2_{W^2}(\mu, \nu) = \varphi.$$

Sketch Proof. We let $F(\mu) = d^2_{W^2}(\mu, \nu)$ and therefore the first variation of F in the direction $\omega \in \mathcal{P}_2(X)$ is (by definition)

$$\partial F(\mu; \omega) = \lim_{h \to 0^+} \frac{F(\mu + h\omega) - F(\mu)}{h} = \lim_{h \to 0^+} \frac{d^2_{W^2}(\mu + h\omega, \nu) - d^2_{W^2}(\mu, \nu)}{h}.$$

Note that the dual constraint set Φ_c depends on the measures μ and ν, however if we assume that $\omega \ll \mu$ then any statement that holds μ almost everywhere also holds $\mu + h\omega$ almost everywhere (and vice versa). Letting $\Phi_c(\mu, \nu)$ be the constraint set where the dependence of the measures is included we have that, by our simplification, $(\varphi, \psi) \in \Phi_c(\mu, \nu) \Leftrightarrow (\varphi, \psi) \in \Phi_c(\mu + h\omega, \nu)$. Similarly, we note that the functional $\mathbb{J}$ also depends on the measures μ and ν hence we write $\mathbb{J}(\cdot, \cdot; \mu, \nu)$ in order to include this dependence. Hence,

$$\mathbb{J}(\varphi, \psi; \mu + h\omega, \nu) = \mathbb{J}(\varphi, \psi; \mu, \nu) + h \int_X \varphi \, \mathrm{d}\omega.$$

Letting (φ, ψ) and (φ_h, ψ_h) (which exist by duality) solve

$$\begin{aligned} d^2_{W^2}(\mu, \nu) &= \mathbb{J}(\varphi, \psi; \mu, \nu) \\ d^2_{W^2}(\mu + h\omega, \nu) &= \mathbb{J}(\varphi_h, \psi_h; \mu + h\omega, \nu) \end{aligned}$$

we have (by duality again)

$$d^2_{W^2}(\mu + h\omega, \nu) = \mathbb{J}(\varphi_h, \psi_h; \mu, \nu) + h \int_X \varphi_h \, \mathrm{d}\omega \le d^2_{W^2}(\mu, \nu) + h \int_X \varphi_h \, \mathrm{d}\omega$$

and

$$d^2_{W^2}(\mu + h\omega, \nu) \ge \mathbb{J}(\varphi, \psi; \mu + h\omega, \nu) = \mathbb{J}(\varphi, \psi; \mu, \nu) + h \int_X \varphi \, \mathrm{d}\omega = d^2_{W^2}(\mu, \nu) + h \int_X \varphi \, \mathrm{d}\omega.$$

Combining these two inequalities we have

$$\int_X \varphi \, \mathrm{d}\omega \le \frac{d^2_{W^2}(\mu + h\omega, \nu) - d^2_{W^2}(\mu, \nu)}{h} \le \int_X \varphi_h \, \mathrm{d}\omega.$$

Now assuming that $\varphi_h \to \varphi$ in $L^1(\omega)$ we have that

$$\partial F(\mu; \omega) = \int_X \varphi \, \mathrm{d}\omega$$

as claimed. □

By the above theorem and (8.17) we have

$$\frac{\partial}{\partial \mu}\phi(\mu_\tau^{(n+1)}) + \frac{\varphi_\tau^{(n)}}{\tau} = 0$$

where $\varphi_\tau^{(n)}$ is the maximiser of the Kantorovich dual problem between $\mu_\tau^{(n)}$ and $\mu_\tau^{(n+1)}$. Now if we recall that $T_\tau^{(n)}$ (where $T_\tau^{(n)}$ is the solution of the Monge problem between $\mu_\tau^{(n)}$ and $\mu_\tau^{(n+1)}$) and $\varphi_\tau^{(n)}$ are related by

$$T_\tau^{(n)}(x) = \nabla \tilde{\varphi}_\tau^{(n)}(x) \qquad \tilde{\varphi}_\tau^{(n)}(x) = \frac{1}{2}|x|^2 - \varphi_\tau^{(n)}(x) \qquad \Rightarrow \qquad T_\tau^{(n)}(x) = x - \nabla \varphi_\tau^{(n)}(x),$$

and $T_\tau^{(n)}$ and $v_\tau^{(n)}$ (the displacement field) are related by $v_\tau^{(n)} = \frac{x - T_\tau^{(n)}(x)}{\tau}$ Hence the gradient flow satisfies

$$0 = \nabla \frac{\partial}{\partial \mu}\phi(\mu_\tau^{(n+1)}) + \frac{\nabla \varphi_\tau^{(n)}}{\tau} = \nabla \frac{\partial}{\partial \mu}\phi(\mu_\tau^{(n+1)}) + v_\tau^{(n)}.$$

Letting $\tau \to 0^+$ then the gradient flow is expected to satisfy

$$\nabla \frac{\partial}{\partial \mu}\phi(\mu(t)) + v_t = 0.$$

Hence $v_t = -\nabla \frac{\partial}{\partial \mu}\phi(\mu_t)$. Substituting v_t into (8.8) we have

$$\frac{\partial}{\partial t}\mu_t - \nabla \cdot \left(\mu_t \nabla \frac{\partial \phi}{\partial \mu}(\mu_t)\right) = 0$$

which we note coincides with (8.11).

Chapter 9

Numerical Approaches to Computing Optimal Transport Distances

In this chapter we discuss some methods for computing optimal transport distances. In particular, we review the following methods: linear programming (very briefly), entropy regularised optimal transport, and flow minimisation. Although we do not discuss it further here one should also note that for semi-discrete optimal transport (see Chapter 5) Theorem 5.4 provides an obvious avenue for developing a numerical method; in particular one can design a numerical scheme, e.g. gradient ascent for maximising g and recover optimal transport maps via the partitioning defined through the Laguerre diagram. The objective is to understand how the linear programming, entropy regularised and flow minimisation approaches work, and although we will often describe the algorithm this will be done in pseudo-code rather than any computational language.

9.1 A Linear Programming Approach

Consider the case when

$$\mu = \sum_{i=1}^m p_i \delta_{x_i} \quad \text{and} \quad \nu = \sum_{j=1}^n q_j \delta_{y_j} \quad \text{where} \quad \sum_{i=1}^m p_i = 1 = \sum_{j=1}^n q_j. \tag{9.1}$$

Then optimal transport problem in the Kantorovich formulation can be written as

$$\mathbb{K}(\mu, \nu) = \min_{\pi} \sum_{i=1}^m \sum_{j=1}^n c(x_i, y_j) \pi_{ij}$$

where the minimum is taken over matrices $\pi \in \mathbb{R}_+^{n \times m}$ such that the row sums are $(p_1, \dots, p_m)$ and the column sums are $(q_1, \dots, q_n)$. This is a linear programme and can be solved, for example, by using the simplex or interior point methods (not covered in this course). In full generality a linear programme is usually written, given $\alpha \in \mathbb{R}^d$, $A \in \mathbb{R}^{d \times d}$ and $b \in \mathbb{R}^d$,

$$\text{maximise } \alpha^\top z \text{ subject to } Az \le b \text{ and } z \ge 0.$$

Although this works well for small problems (i.e. when n and m are small) linear programming algorithms quickly become infeasible for large problems. There are ad-hoc multiscale methods such as [18] that allow linear programming methods to larger problems but these do not have theoretical guarantees (for example, there are no guarantees regarding finding the global minimiser).

9.2 An Entropy Regularisation Approach

Section references: The proof of the optimal form for minimising the Kullback-Liebler divergence, Proposition 9.1, comes from [19, Proposition 4.3].

We assume two pairs μ, ν are discrete probability measures that can be written in the form (9.1). It was proposed in [6] to consider the entropy regularised problem

$$S_\varepsilon(\mu,\nu) = \min_{\pi \in \Pi(\mu,\nu)} \left(\sum_{i=1}^m \sum_{j=1}^n c(x_i, y_j)\pi_{ij} - \varepsilon H(\pi) \right) \tag{9.2}$$

where $\varepsilon > 0$ is a positive parameter that controls the amount of regularisation, H is entropy and defined by

$$H(\pi) = -\sum_{i=1}^n \sum_{j=1}^m \pi_{ij} \log \pi_{ij}$$

and the set $\Pi(\mu,\nu)$ can be thought of as the set of matrices $\pi \in \mathbb{R}_+^{n\times m}$ such that the row sums are $\boldsymbol{p} = (p_1, \dots, p_m)$ and the column sums are $\boldsymbol{q} = (q_1, \dots, q_n)$. ***In the remainder of this chapter we will, with an abuse of notation, assume that π are matrices.*** It is beyond the scope of the course but one can show that as $\varepsilon \to 0^+$ that $S_\varepsilon(\mu,\nu) \to \mathbb{K}(\mu,\nu)$ under appropriate conditions on c. The measure S_ε is referred to as the Sinkhorn distance.

Let C be the $n \times m$ matrix with $(i,j)^{\text{th}}$ entry $C_{ij} = c(x_i, y_j)$. Then,

$$\begin{aligned} \sum_{i=1}^m \sum_{j=1}^n C_{ij}\pi_{ij} - \varepsilon H(\pi) &= \sum_{i=1}^m \sum_{j=1}^n \pi_{ij} \left(C_{ij} + \varepsilon \log \pi_{ij} \right) \\ &= \varepsilon \sum_{i=1}^m \sum_{j=1}^n \pi_{ij} \log \left(\pi_{ij} \mathrm{e}^{\frac{C_{ij}}{\varepsilon}} \right) \\ &= \varepsilon \mathrm{KL}(\pi | K) \end{aligned}$$

where KL is the Kullback-Leibler divergence defined by

$$\mathrm{KL}(\pi|K) = -\sum_{i=1}^m \sum_{j=1}^n \pi_{ij} \log \frac{K_{ij}}{\pi_{ij}}$$

and K is the $n \times m$ matrix with $(i,j)^{\text{th}}$ entry $K_{ij} = \mathrm{e}^{-\frac{C_{ij}}{\varepsilon}}$. Hence

$$S_\varepsilon(\mu,\nu) = \varepsilon \min_{\pi \in \Pi(\mu,\nu)} \mathrm{KL}(\pi|K).$$

Let us define the sets:

$$\mathcal{C}_1 = \left\{\pi \in \mathbb{R}_+^{n\times m} \,:\, \pi \mathbb{1}_n = \boldsymbol{p}\right\} \tag{9.3}$$

$$\mathcal{C}_2 = \left\{\pi \in \mathbb{R}_+^{n\times m} \,:\, \pi^\top \mathbb{1}_m = \boldsymbol{q}\right\} \tag{9.4}$$

where $\mathbb{1}_n \in \mathbb{R}^n$ and $\mathbb{1}_m \in \mathbb{R}^m$ are the vectors of ones. If $\pi \in \mathcal{C}_1$ then the row sums are $\boldsymbol{p}$, and if $\pi \in \mathcal{C}_2$ then the columns sums are $\boldsymbol{q}$, hence we can write (with an abuse of notation) $\pi(\mu,\nu) = \mathcal{C}_1 \cap \mathcal{C}_2$.

Proposition 9.1. *Let $K \in \mathbb{R}^{n\times m}$ and $\boldsymbol{p} \in \mathbb{R}_+^m$, $\boldsymbol{q} \in \mathbb{R}_+^n$ with $\sum_{i=1}^m p_i = 1 = \sum_{j=1}^n q_j$ and define $\mathcal{C}_1$, $\mathcal{C}_2$ by (9.3-9.4). Then the solution $\pi^\dagger$ of*

$$\min_{\pi \in \mathcal{C}_1 \cap \mathcal{C}_2} \mathrm{KL}(\pi | K) \tag{9.5}$$

has the form $\pi_{ij} = u_i K_{ij} v_j$ for some variables $\boldsymbol{u} \in \mathbb{R}_+^n$ and $\boldsymbol{v} \in \mathbb{R}_+^m$.

Sketch Proof. Let $\boldsymbol{f} \in \mathbb{R}^m$ and $\boldsymbol{g} \in \mathbb{R}^n$ be the dual variables, then the Lagrangian form of (9.5) can be written

$$\min_{\pi \in \mathcal{C}_1 \cap \mathcal{C}_2} \mathrm{KL}(\pi|K) = \min_{\pi \in \mathbb{R}^{n\times m}} \sup_{\boldsymbol{f}\in\mathbb{R}^n, \boldsymbol{g}\in\mathbb{R}^m} \left(\mathrm{KL}(\pi|K) + \boldsymbol{f}\cdot(\pi\mathbb{1} - \boldsymbol{p}) + \boldsymbol{g}\cdot(\pi^\top \mathbb{1} - \boldsymbol{q})\right). \tag{9.6}$$

Define

$$\mathcal{E}(\pi, \boldsymbol{f}, \boldsymbol{g}) = \mathrm{KL}(\pi|K) + \boldsymbol{f}\cdot(\pi\mathbb{1} - \boldsymbol{p}) + \boldsymbol{g}\cdot(\pi^\top\mathbb{1} - \boldsymbol{q}).$$

For fixed $\boldsymbol{f}$ and $\boldsymbol{g}$ the optimal π, i.e. the solution to $\min_{\pi\in\mathbb{R}^{n\times m}} \mathcal{E}(\pi,\boldsymbol{f},\boldsymbol{g})$ satisfies

$$\frac{\partial \mathcal{E}}{\partial \pi_{ij}}(\pi, \boldsymbol{f}, \boldsymbol{g}) = 0 \quad \forall i, j.$$

A simple computation implies

$$\begin{aligned}\frac{\partial \mathcal{E}}{\partial \pi_{ij}}(\pi,\boldsymbol{f},\boldsymbol{g}) &= \frac{\partial}{\partial \pi_{ij}}\left(\sum_{k,\ell} \pi_{k\ell}\log\frac{\pi_{k\ell}}{K_{k\ell}} + \sum_k f_k\left(\sum_\ell \pi_{k\ell} - p_k\right) + \sum_\ell g_\ell\left(\sum_k \pi_{k\ell} - q_\ell\right)\right)\\ &= \log\frac{\pi_{ij}}{K_{ij}} + 1 + f_i + g_j.\end{aligned}$$

Hence $\pi_{ij} = e^{-f_i}K_{ij}e^{-g_j}e^{-1}$. Setting $\boldsymbol{u} = e^{-\boldsymbol{f}}$ and $\boldsymbol{v} = e^{-\boldsymbol{g}-1}$, and assuming a min-max principle (that we can switch the min and sup in (9.6)) completes the argument. □

By the above proposition the optimal solution to (9.5) can be written in the form $\pi^\dagger = \mathrm{diag}(\boldsymbol{u})K\mathrm{diag}(\boldsymbol{v})$. Since $\pi^{\mathrm{diag}} \in \mathcal{C}_1 \cap \mathcal{C}_2$ we must also have

$$\mathrm{diag}(\boldsymbol{u})K\mathrm{diag}(\boldsymbol{v})\mathbb{1}_n = \boldsymbol{p} \qquad \mathrm{diag}(\boldsymbol{v})K^\top\mathrm{diag}(\boldsymbol{u})\mathbb{1}_m = \boldsymbol{q}.$$

Since $\mathrm{diag}(\boldsymbol{v})\mathbb{1}_n = \boldsymbol{v}$ and $\mathrm{diag}(\boldsymbol{u})\mathbb{1}_m = \boldsymbol{u}$ the above implies

$$\boldsymbol{u}\odot(K\boldsymbol{v}) = \boldsymbol{p} \qquad \boldsymbol{v}\odot(K^\top\boldsymbol{u}) = \boldsymbol{q}$$

where $\odot$ denotes pointwise multiplication. A natural way to solve these equations is to recursively update $\boldsymbol{u}$ and $\boldsymbol{v}$ by

$$\boldsymbol{u}^{(k+1)} = \frac{\boldsymbol{p}}{K\boldsymbol{v}^{(k)}} \qquad \boldsymbol{v}^{(k+1)} = \frac{\boldsymbol{q}}{K^\top \boldsymbol{u}^{(k+1)}} \tag{9.7}$$

with the initialisation $v^{(0)} = \mathbb{1}_n$. The division operator above is implemented pointwise. This is known as Sinkhorn's algorithm.

The iteration's (9.7) can also be seen as projections onto the hyperplanes $\mathcal{C}_1$ and $\mathcal{C}_2$. The projection onto C with respect to the Kullback-Leibler divergence is defined by

$$P_C^{\mathrm{KL}}(\pi) = \operatorname*{argmin}_{\gamma \in C} \mathrm{KL}(\gamma|\pi).$$

Proposition 9.2. *Let $\boldsymbol{p} \in \mathbb{R}_+^m$ and $\boldsymbol{q} \in \mathbb{R}_+^n$ with $\sum_{i=1}^m p_i = 1 = \sum_{j=1}^n q_j$ and define $\mathcal{C}_1$, $\mathcal{C}_2$ by (9.3-9.4). Then the projection onto $\mathcal{C}_1$ and $\mathcal{C}_2$ with respect to the Kullback-Leibler divergence is given by*

$$P_{\mathcal{C}_1}^{\mathrm{KL}}(\pi) = \mathrm{diag}\left(\frac{\boldsymbol{p}}{\pi \mathbb{1}_n}\right)\pi, \qquad P_{\mathcal{C}_2}^{\mathrm{KL}}(\pi) = \mathrm{diag}\left(\frac{\boldsymbol{q}}{\pi^\top \mathbb{1}_m}\right)\pi$$

where division in the above operators is understood to be pointwise.

Proof. If $\gamma \in \mathcal{C}_1$ then $\sum_{j=1}^n \gamma_{ij} = p_i$ for all i. Hence we treat γ_{ij} for $j = 1, \dots, n-1$ as free variables and define $\gamma_{in} = p_i - \sum_{j=1}^{n-1} \gamma_{ij}$. Then,

$$\begin{aligned} \mathrm{KL}(\gamma|\pi) &= \sum_{i=1}^m \sum_{j=1}^n \gamma_{ij} \log \frac{\gamma_{ij}}{\pi_{ij}} \\ &= \sum_{i=1}^m \left(\sum_{j=1}^{n-1} \gamma_{ij} \log \frac{\gamma_{ij}}{\pi_{ij}} + \left(p_i - \sum_{j=1}^{n-1} \gamma_{ij}\right) \log \frac{p_i - \sum_{j=1}^{n-1} \gamma_{ij}}{\pi_{in}} \right). \end{aligned}$$

So, for all $j \neq n$,

$$\begin{aligned} \frac{\partial}{\partial \gamma_{ij}} \mathrm{KL}(\gamma|\pi) &= \log \frac{\gamma_{ij}}{\pi_{ij}} + 1 - \log \frac{p_i - \sum_{j=1}^{n-1} \gamma_{ij}}{\pi_{in}} - \frac{p_i - \sum_{j=1}^{n-1} \gamma_{ij}}{p_i - \sum_{j=1}^{n-1} \gamma_{ij}} \\ &= \log \frac{\gamma_{ij}}{\pi_{ij}} - \log \frac{p_i - \sum_{j=1}^{n-1} \gamma_{ij}}{\pi_{in}}. \end{aligned}$$

Hence $\frac{\partial}{\partial \gamma_{ij}} \mathrm{KL}(\gamma|K) = 0$ implies $\frac{\gamma_{ij}}{\pi_{ij}} = \frac{\gamma_{in}}{\pi_{in}}$. Since the right hand side is independent of j we have that there exists some α_i such that $\gamma_{ij} = \alpha_i \pi_{ij}$ for all $j = 1, \dots, n-1$.

Repeating the argument where we instead treat γ_{ij} for $j = 2, \dots, n$ as free variables and define $\gamma_{i1} = p_i - \sum_{j=2}^n \gamma_{ij}$ we have that there exists α' such that $\gamma_{ij} = \alpha_i' \pi_{ij}$ for all $j = 2, \dots, n$. It must follows that $\alpha_i = \alpha_i'$ and therefore $\gamma_{ij} = \alpha_i \pi_{ij}$ for all $j = 1, \dots, n$.

Since $\sum_{j=1}^n \gamma_{ij} = p_i$ then $\alpha_i = \frac{p_i}{\sum_{j=1}^n \pi_{ij}}$ which implies $\gamma_{ij} = \frac{p_i \pi_{ij}}{\sum_{j=1}^n \pi_{ij}}$. Equivalently this can be written $\gamma = \mathrm{diag}\left(\frac{\boldsymbol{p}}{\pi \mathbb{1}_n}\right)\pi$. The argument is analogous for $P_{\mathcal{C}_2}^{\mathrm{KL}}$. $\square$

If we let $\pi = \mathrm{diag}(\boldsymbol{u}) K \mathrm{diag}(\boldsymbol{v})$ we see that

$$\begin{aligned} P_{\mathcal{C}_1}^{\mathrm{KL}}(\pi) &= \mathrm{diag}\left(\frac{\boldsymbol{p}}{\boldsymbol{u} \odot (Kv)}\right) \mathrm{diag}(\boldsymbol{u}) K \mathrm{diag}(\boldsymbol{v}) \\ &= \mathrm{diag}\left(\frac{\boldsymbol{p}}{K\boldsymbol{v}}\right) K \mathrm{diag}(\boldsymbol{v}) \\ &= \mathrm{diag}(\boldsymbol{u}') K \mathrm{diag}(\boldsymbol{v}) \end{aligned}$$

where $\boldsymbol{u}' = \frac{\boldsymbol{p}}{K\boldsymbol{v}}$. Similarly $P_{\mathcal{C}_1}^{\mathrm{KL}}(\pi) = \mathrm{diag}(\boldsymbol{u}) K \mathrm{diag}(\boldsymbol{v}')$ where $v' = \frac{\boldsymbol{q}}{K^\top \boldsymbol{u}}$. Hence, (9.7) are the projections of $\pi^{(k)} = \mathrm{diag}(\boldsymbol{u}^{(k)}) K \mathrm{diag}(\boldsymbol{v}^{(k)})$ onto $\mathcal{C}_1$ and $\mathcal{C}_2$.

9.3 A Flow Minimisation Approach

Section references: we follow the derivation given in [11].

Let $\Omega \subset \mathbb{R}^d$ be a compact domain with smooth boundary and let μ, ν be two probabilitiy measures on Ω that admit densities with respect to the Lebesgue measure. Abusing notation we write $\mathrm{d}\mu(x) = \mu(x)\mathrm{d}x$ and $\mathrm{d}\nu(x) = \nu(x)\mathrm{d}x$. The Monge optimal transport problem with quadratic cost is to find the map $T : \Omega \to \Omega$, which minimises the following energy

$$\mathcal{E}(T) = \int_\Omega |T(x) - x|^2 \mu(x)\,\mathrm{d}x \tag{9.8}$$

subject to $T_\# \mu = \nu$. We assume the following polar factorization of T. Let $s : \Omega \times [0, \infty) \to \Omega$ and call the second coordinate time. We further assume for any fixed t, $[s(\cdot, t)]_\# \mu = \mu$. That is, $s(\cdot, t) : \Omega \to \Omega$ is a mass preserving rearrangement of μ. Let $T^0 : \Omega \to \Omega$ be an initial mass preserving map between μ and ν, i.e. $T^0_\# \mu = \nu$. We assume that $s(\cdot, t)$ is invertible in x for every t and with an abuse of notation we write s^{-1} for this inverse, i.e.

$$s^{-1}(s(x,t),t) = x = s(s^{-1}(x,t),t) \quad \text{for all } x \in \Omega \text{ and for all } t \in [0, \infty). \tag{9.9}$$

We require that $T = T^0 \circ s^{-1}$. The strategy is to evolve $s(\cdot, t)$ using a gradient descent step such that it converges to a minimiser of $\mathcal{E}$ satisfying the appropriate constraint as $t \to \infty$. We first consider sufficient conditions on s in order to guarantee that $[s(\cdot, t)]_\# \mu = \mu$ holds for all $t > 0$.

Proposition 9.3. *Let $\Omega \subset \mathbb{R}^d$ be a compact domain with a smooth boundary and $\mu, \nu \in \mathcal{P}(\Omega)$. Assume that μ and ν have C^1 densities with respect to the Lebesgue measure on Ω and with an abuse of notation write $\mathrm{d}\mu(x) = \mu(x)\,\mathrm{d}x$ and $\mathrm{d}\nu(x) = \nu(x)\,\mathrm{d}x$ Let χ be a C^1 vector field on Ω satisfying $\mathrm{div}(\chi) = 0$ on Ω and $\chi \cdot n = 0$ on $\partial\Omega$ where n is the normal to the boundary of Ω. Assume $s : \Omega \times [0, \infty) \to \Omega$ is differentiable and invertible in the sense of (9.9), and T^0 satisfies $T^0_\# \mu = \nu$. If, $s(\cdot, 0) = \mathrm{Id}$ and for all $t > 0$*

$$\frac{\partial s}{\partial t}(x,t) = \frac{1}{\mu(s(x,t))} \chi(s(x,t)) \tag{9.10}$$

then $[s(\cdot, t)]_\# \mu = \mu$. Furthermore $\frac{\partial T^t}{\partial t} = -\frac{1}{\mu} \nabla T^t \chi$ and $T^t_\# \mu = \nu$ where

$$T^t = T^0(s^{-1}(\cdot, t)). \tag{9.11}$$

Proof. We show the following:

1. $[s(\cdot,t)]_{\#}\mu = \mu$ for all $t > 0$;
2. $T^t_{\#}\mu = \nu$ for all $t > 0$; and
3. $\frac{\partial T^t}{\partial t} = -\frac{1}{\mu}\nabla T^t \chi$.

1. Let $\omega(\cdot,t) = [s^{-1}(\cdot,t)]_{\#}\mu$, and with an abuse of notation we again write $\mathrm{d}\omega(x,t) = \omega(x,t)\,\mathrm{d}x$. We show that ω is independent of time. By Exercise 1.3 and Exercise 4.3,

$$
\begin{aligned}
\frac{\partial}{\partial t}\omega(x,t) &= \frac{\partial}{\partial t}\left(|\det \nabla_x s(s,t)|\mu(s(x,t))\right) \\
&= \left(\frac{\partial}{\partial t}|\det \nabla_x s(x,t)|\right)\mu(s(x,t)) + |\det(\nabla_x s(x,t))|\left(\frac{\partial}{\partial t}\mu(s(x,t))\right) \\
&= |\det(\nabla_x s(x,t))|\left[\mu \mathrm{div}\left(\frac{\partial s}{\partial t}(s^{-1}(\cdot,t),t)\right)\right]\circ s(x,t) \\
&\qquad + |\det(\nabla s(x,t))|\nabla_x \mu(s(x,t))\cdot\frac{\partial s}{\partial t}(x,t) \\
&= |\det(\nabla_x s(x,t))|\left[\mu \mathrm{div}\left(\frac{\partial s}{\partial t}(s^{-1}(\cdot,t),t)\right) + \nabla_x\mu\cdot\frac{\partial s}{\partial t}(s^{-1}(\cdot,t),t)\right]\circ s(x,t) \\
&= |\det(\nabla_x s(x,t))|\left[\mathrm{div}\left(\mu\frac{\partial s}{\partial t}(s^{-1}(\cdot,t),t)\right)\right]\circ s(x,t).
\end{aligned}
$$

Now, $\mu(y)\frac{\partial s}{\partial t}(s^{-1}(y,t),t) = \chi(y)$, hence

$$
\mathrm{div}\left(\mu(y)\frac{\partial s}{\partial t}(s^{-1}(y,t),t)\right) = \mathrm{div}\chi(y) = 0.
$$

It follows that ω is independent of time. Since $\omega(\cdot,0) = \mu$ then $\omega(\cdot,t) = \mu$ for all $t > 0$. We have shown $[s^{-1}(\cdot,t)]_{\#}\mu = \mu$, hence $\mu = [s(\cdot,t)]_{\#}\mu$ by the composition of maps.

2. By the composition of maps (see Proposition 1.5) and part 1 we have

$$
\mu = [s_\chi^{-1}(\cdot,t)\circ s_\chi(\cdot,t)]_{\#}\mu = [s_\chi^{-1}(\cdot,t)]_{\#}\left([s_\chi(\cdot,t)]_{\#}\mu\right) = [s_\chi^{-1}(\cdot,t)]_{\#}\mu.
$$

And therefore, by the composition of maps again,

$$
T^t_{\#}\mu = [T^0\circ s_\chi^{-1}(\cdot,t)]_{\#}\mu = T^0_{\#}\left([s_\chi^{-1}(\cdot,t)]_{\#}\mu\right) = T^0_{\#}\mu = \nu.
$$

3. We have $T^0(x) = T^t(s(x,t))$, so differentiating by t implies,

$$\begin{aligned} 0 &= \frac{\partial T^t}{\partial t}(s(x,t)) + \nabla T^t(s(x,t))\frac{\partial s}{\partial t}(x,t) \\ &= \frac{\partial T^t}{\partial t}(s(x,t)) + \frac{1}{\mu(s(x,t))}\nabla T^t(s(x,t))\chi(s(x,t)) \end{aligned}$$

where we use step 1. Hence, $\frac{\partial T^t}{\partial t}(y) = -\frac{1}{\mu(y)}\nabla T^t(y)\chi(y)$ as required. □

If we restrict ourselves to look for transport maps of the form (9.10-9.11) then we must decide how to choose χ. Let us define $s_\chi : \Omega \times [0,\infty) \to \Omega$ by (9.10) with $s_\chi(\cdot,0) = \mathrm{Id}$ and $T^t_\chi : \Omega \to \Omega$ by (9.11) with $s = s_\chi$. An obvious criterion is to choose χ so that $\mathcal{E}(T^t_\chi)$ decreases quickest over all choices of χ. To this end we compute the derivative of $\mathcal{E}(T^t_\chi)$ with respect to t.

Lemma 9.4. *In addition to the assumptions and notation of Proposition 9.3 let $f \in C^1(\Omega;\mathbb{R}^m)$ and $g \in L^2(\nu)$ and define $\mathcal{E}$ by* (9.8)*. Define $s_\chi : \Omega \times [0,\infty) \to \Omega$ by* (9.10) *with $s_\chi(\cdot,0) = \mathrm{Id}$ and $T^t_\chi : \Omega \to \Omega$ by* (9.11) *with $s = s_\chi$. Then we have*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(T^t_\chi) = -2\int_\Omega T^t(x)\cdot\chi(x)\,\mathrm{d}x.$$

Proof. By a change of variables $y = s_\chi^{-1}(x,t)$, and since $[s_\chi(\cdot,t)]_{\#}\mu = \mu$ we have by the change of variables formula

$$\begin{aligned} \mathcal{E}(T^t_\chi) &= \int_\Omega \left|T^0(s_\chi^{-1}(x,t)) - x\right|^2 \mu(x)\,\mathrm{d}x \\ &= \int_\Omega \left(|T^0(s_\chi^{-1}(x,t))|^2 + |x|^2 - 2T^0(s_\chi^{-1}(x,t))\cdot x\right)\mu(x)\,\mathrm{d}x \\ &= \int_\Omega \left(|T^0(x)|^2 + |x|^2\right)\mu(x)\,\mathrm{d}x - 2\int_\Omega T^0(y)\cdot s_\chi(y,t)\mu(y)\,\mathrm{d}y. \end{aligned}$$

Differentiating the above we obtain,

$$\begin{aligned} \frac{\mathrm{d}\mathcal{E}}{\mathrm{d}t}(T^t_\chi) &= -2\int_\Omega T^0(y)\cdot\frac{\mathrm{d}s_\chi}{\mathrm{d}t}(y,t)\mu(y)\,\mathrm{d}y \\ &= -2\int_\Omega T^0(y)\cdot\frac{\chi(s_\chi(y,t))}{\mu(s_\chi(y,t))}\mu(y)\,\mathrm{d}y \\ &= -2\int_\Omega T^t_\chi(x)\cdot\chi(x)\,\mathrm{d}x \end{aligned}$$

as required. □

When $d = 2$ by the Helmholtz decomposition (in 2D) we can find, for each $t > 0$, two scalar fields $w : \Omega \to \mathbb{R}$ and $\alpha : \Omega \to \mathbb{R}$ such that $2T^t = \nabla w + \nabla^\perp\alpha$ and $\alpha = 0$ on $\partial\Omega$ where $\nabla^\perp f = \left(-\frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_1}\right)$ for a function $f(x) = f(x_1,x_2)$ (w and α will of course depend on t but we suppress this dependence as t is fixed). To find the direction of steepest descent we let $\psi = \nabla^\perp \alpha$ and $\chi = \nabla^\perp \beta$ and compute

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(T^t_\chi) &= -\int_\Omega (\nabla w(x) + \psi(x)) \cdot \chi(x)\,\mathrm{d}x \\
&= -\int_\Omega (\mathrm{div}(w\chi) - w\mathrm{div}(\chi))\,\mathrm{d}x - \int_\Omega \psi(x)\cdot\chi(x)\,\mathrm{d}x \\
&= -\int_{\partial\Omega} w(x)\chi(x)\cdot n(x)\,\mathrm{d}S(x) - \int_\Omega \psi(x)\cdot\chi(x)\,\mathrm{d}x \\
&= -\int_\Omega \psi(x)\cdot\chi(x)\,\mathrm{d}x \\
&= -\int_\Omega \nabla\alpha(x)\cdot\nabla\beta(x)\,\mathrm{d}x
\end{aligned}$$

where the third line follows from the divergence theorem and since $\mathrm{div}(\chi) = 0$ on Ω, and the fourth line follows from $\chi(x)\cdot n(x) = 0$ on $\partial\Omega$. It follows that the direction of steepest descent is $\alpha = \beta$.

To find α, we need to observe that $\nabla\alpha = -2(T^t)^\perp - \nabla^\perp w$ where $\perp$ is rotation clockwise by $\pi/2$, i.e. $Q^\perp = (-Q_2, Q_1)$. Taking the divergence we have

$$\Delta\alpha = \mathrm{div}(\nabla\alpha) = \mathrm{div}(-2(T^t)^\perp - \nabla^\perp w) = -2\mathrm{div}(T^t)^\perp.$$

Hence, α solves the Poisson equation with Dirichlet boundary conditions:

$$\Delta\alpha = -\mathrm{div}((T^t)^\perp) \tag{9.12}$$

$$\alpha = 0 \ \partial\Omega. \tag{9.13}$$

To summarise, the flow minimization scheme, given a step size τ and an initial map T^0 is as follows.

1. Set $t = 0$.
2. Find α by solving (9.12-9.13).
3. Update $T^{t+\tau} = T^t - \frac{\tau}{\mu}\nabla T^t \nabla^\perp \alpha$.
4. Set $t \mapsto t + \tau$.
5. Repeat 2-4 until convergence.

Bibliography

[1] L. Ambrosio. *Mathematical Aspects of Evolving Interfaces: Lectures given at the C.I.M.-C.I.M.E. joint Euro-Summer School held in Madeira, Funchal, Portugal, July 3-9, 2000*, chapter Lecture Notes on Optimal Transport Problems, pages 1–52. Springer Berlin Heidelberg, 2003.

[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[3] L. Ambrosio and A. Pratelli. *Optimal Transportation and Applications: Lectures given at the C.I.M.E. Summer School, held in Martina Franca, Italy, September 2-8, 2001*, chapter Existence and stability results in the L1 theory of optimal transportation, pages 123–160. Springer Berlin Heidelberg, 2003.

[4] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[5] Y. Brenier. Décomposition polaire et réarrangement monotone des champs de vecteurs. *Comptes rendus de l'Académie des Sciences, Paris, Série I*, 305:805–808, 1987.

[6] M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems (NIPS)*, pages 2292–2300, 2013.

[7] S. Daneri and G. Saveré. Lecture notes on gradient flows and optimal transport. *(preprint) arXiv:1009.3737*, 2010. Based on lectures given at the summer school "Optimal Transporation: Theory and Applications" in Grenoble, June 2009.

[8] I. Ekeland. On the variational principle. *Journal of Mathematical Analysis and Applications*, 47(2):324–353, 1974.

[9] L. C. Evans and W. Gangbo. *Differential equations methods for the Monge-Kantorovich mass transfer problem*, volume 653. American Mathematical Society, 1999.

[10] W. Gangbo and R. J. McCann. The geometry of optimal transportation. *Acta Mathematica*, 177(2):113–161, 1996.

[11] S. Haker, L. Zhu, A. Tannenbaum, and S. Angenent. Optimal mass transport for registration and warping. *International Journal of computer vision*, 60(3):225–240, 2004.

[12] L. V. Kantorovich. On translation of mass (in Russian), C. R. Doklady. *Proceedings of the USSR Academy of Sciences*, 37:199–201, 1942.

[13] S. Kolouri, S. R. Park, M. Thorpe, D. Slepčev, and G. K. Rohde. Optimal mass transport: Signal processing and machine-learning applications. *IEEE Signal Processing Magazine*, 34(4):43–59, 2017.

[14] S. Kolouri and G. K. Rohde. Optimal transport a crash course. IEEE ICIP 2016 Tutorial Slides: Part 1, 2016.

[15] B. Lévy and E. L. Schwindt. Notions of optimal transport theory and how to implement them on a computer. *Computers & Graphics*, 72:135–148, 2018.

[16] G. Monge. *Mémoire sur la théorie des déblais et des remblais*. De l'Imprimerie Royale, 1781.

[17] M. Muratori and G. Savaré. Gradient flows and evolution variational inequalities in metric spaces. I: structural properties. *(preprint) arXiv:1810.03939*, 2018.

[18] A. M. Oberman and Y. Ruan. An efficient linear programming method for optimal transportation. *(preprint) arXiv:1509.03668*, 2015.

[19] G. Peyré and M. Cuturi. Computational optimal transport. *(preprint) arXiv:1803.00567*, 2018.

[20] F. Santambrogio. *Optimal transport for applied mathematicians*. Birkäuser Springer, Basel, 2015.

[21] B. Simon. *Convexity: An Analytic Viewpoint*. Cambridge University Press, 2011.

[22] V. N. Sudakov. Geomtric problems in the theory of infinite-dimensional probability distributions. *Proceedings of the Steklov Institute of Mathematics*, 141:1–178, 1979.

[23] C. Villani. *Topics in Optimal Transportation*. American Mathematical Society, 2003.

[24] C Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.