Università degli Studi di Pisa
Scuola di Dottorato in Matematica
XVII ciclo

Ph.D. Thesis in Mathematics

Mass Transportation Problems and Minimal Measures

by
Dr. Luca Granieri

Advisors
Prof. Giuseppe Buttazzo
Dr. Luigi De Pascale

Ph.D. School Director
Prof. Fabrizio Broglia

2005

Alla mia cara mamma

Contents

Ringraziamenti

Ringrazio con piacere il prof. Giuseppe Buttazzo per la sua disponibilità e per essere stato un riferimento costante durante questi anni di dottorato.

Desidero inoltre ringraziare il prof. Luigi Ambrosio, il prof. Sergio Solimini e il dott. Franco Maddalena per l'interesse mostrato verso le ricerche contenute in questo lavoro, e tutti i profesori e ricercatori che ho avuto modo di incontrare ed apprezare durante il mio soggiorno a Pisa per il clima sereno e amichevole in cui ho potuto vivere e lavorare.

Desidero anche ringraziare la dott.ssa Maria Stelli Gelli con la quale ho avuto il piacere di collaborare, per la sua amicizia, i suoi consigli e le piacevoli chiacchierate.

Un ringraziamento particolare voglio rivolgere al dott. Luigi De Pascale che mi ha introdotto ai temi di ricerca sviluppati in questa tesi. Molta parte di questo lavoro non sarebbe stata possibile senza i suoi preziosi suggerimenti e commenti, i suoi consigli e costanti stimoli, la sua amicizia, premura e pazienza. Grazie mille Gigi!

Ringrazio volentieri anche Angel, Michele, Francesca, Stefania, Vittoria, Marina, Marco, Antonio, Patrizia, Barbara, Chiara e tutti gli amici con cui ho condiviso l'avventura della ricerca matematica in questi anni di dottorato.

Molte altre persone hanno contribuito, in maniera più o meno diretta, alla realizzazione di questo lavoro.

In primo luogo il buon Dio che ha avuto la felice idea, almeno per quanto mi riguarda, di mettermi al mondo.

La mia ragazza Maria Angela per aver sopportato con pazienza e amore questi anni di lontananza (sperando che siano finiti!).

I miei genitori, le mie tre sorelle e mio fratello con le rispettive famiglie, per il loro insostituibile affetto e aiuto.

Il caro don Giò, i gifrini di Noicattaro e Pisa e tutti gli amici che mi hanno sostenuto con il loro affetto e la loro stima.

Le persone che vorrei ringraziare non finiscono certo qui e una lista completa avrebbe sicuramente una lunghezza smisurata. Anche a tutte queste persone dedico questo lavoro, nella speranza che mi perdonino per tutto il tempo che ho dovuto sottrarre loro.

Introduction

In the last ten years mass transportation problems, also known as Monge-Kantorovich problems, have been intensively studied in particular with respect to numerous and important applications and connections with other fields such as PDE, material sciences, probability and economics. One of these connections, proposed by L. C. Evans in some recent works (see [36, 40]) is concerned with Mather's theory of minimal measures in Lagrangian dynamic. The principal aim of this thesis is to investigate some aspects of these connections.

The Monge-Kantorovich and the Mather problem come from quite different settings. A first connection between them will be a formulation of both problems in terms of normal 1-currents. This formulation generalizes some results of [7].

Once the first bridge between the two problems is established, it is also natural to think of a formulation of the Monge-Kantorovich problem in terms of probability measures on the phase-space minimizing a suitable action functional. In the Mather problem the minimal measures have to satisfy a closure constraint and an homological one (see chapter 2 for further details). In the Monge-Kantorovich problem a suitable boundary condition has to be satisfied but, as we will see, this is not always enough. Therefore we also investigate this formulation.

Now, let us briefly describe the Monge-Kantorovich and the Mather problem and the main results of this work in more details.

We will consider a compact Riemannian manifold (M, g) without boundary and a Lagrangian $L : TM \to \mathbb{R}$ which satisfies the following properties:

1. L is regular (say $\mathcal{C}^2$) and positive,
2. $L(x, \cdot)$ is superlinear in the fiber for all x,
3. $L(x, \cdot)$ is strictly convex in the fiber for all x.

Finally, $H : T^*M \to \mathbb{R}$ will denote the convex conjugate of L on the fibers, that is

$$H(x, \cdot) = L^*(x, \cdot).$$

The Monge-Kantorovich problem

Our description will be restricted to the setting of a compact Riemannian manifold M, however many of the concepts of this section could be formulated in general metric spaces or even in a topological setting. For a detailed presentation the reader may look at the book and lecture notes [1, 19, 37, 73, 82], the paper [48] and for some of the applications [11, 12, 18, 20, 56]. In chapter 1 we provide the basic tool of the optimal mass transportation theory that we will need. We pay particular attention to duality and to Kantorovich potentials. We shortly discuss the problem of finding potentials which are distance functions and we also provide an explicit counterexample (see section 1.3). Then we describe the Brenier's formulation of the optimal mass transportation problems.

Let $c : M \times M \to \mathbb{R}^+$ be a positive Borel function. The Monge problem is formulated has follows: given two probability measures μ, ν find a map $t : M \to M$ such that $t_\sharp \mu = \nu$ ($\sharp$ denotes the push-forward operator of measures) and such that t minimizes the transportation cost

$$\int_M c(x, t(x)) d\mu$$

among the maps with the same property. It may happens that the set of admissible maps is empty (e.g. this occurs if we take $\mu = \delta_x$ and $\nu = \frac{1}{2}(\delta_y + \delta_z)$). Moreover, because of the strong non-linearity, existence of minimizers is a difficult matter. In order to avoid this difficulties the problem can be reformulated in its Kantorovich's relaxed form which amounts to:

$$\mathbf{MK} := \min \left\{ \int_{M \times M} c(x,y) d\gamma(x,y) \mid \gamma \in \mathcal{P}(M \times M), \pi^1_\sharp \gamma = \mu, \pi^2_\sharp \gamma = \nu \right\}. \tag{1}$$

Here $\mathcal{P}$ stands for the space of probability measures and $\pi^{1,2}$ are the projections on the factors of $M \times M$. The admissible measures γ for problem $\mathbf{MK}$ are called transport plans. Observe that if t is admissible for the Monge problem then the measure $\gamma = (id \times t)_\sharp \mu$ is a transport plan for $\mathbf{MK}$. Furthermore, the class of transport plans is never empty as it contains for instance the measure $\gamma = \mu \otimes \nu$. The Kantorovich problem $\mathbf{MK}$ is now linear and the semicontinuity of c is enough to prove existence of minimizers. Moreover this is true also if c takes the value $+\infty$ (see chapter 1 for details).

If c is a distance then the cost of transportation

$$d_c(\mu, \nu) = \min \left\{ \int_{M \times M} c(x,y) d\gamma(x,y) \mid \pi^1_\sharp \gamma = \mu, \pi^2_\sharp \gamma = \nu \right\}$$

defines a distance on $\mathcal{P}(M)$. If d is the geodesic distance on the manifold M then for $p \geq 1$ also

$$W_p(\mu, \nu) = \left(\min \left\{ \int_{M \times M} d^p(x,y) d\gamma(x,y) \mid \pi^1_\sharp \gamma = \mu, \pi^2_\sharp \gamma = \nu \right\} \right)^{1/p}$$

defines a distance on $\mathcal{P}(M)$ called Wasserstein distance of order p, or p-Wasserstein distance. Moreover the space $(\mathcal{P}(M), W_p)$ is complete and separable and W_p metrizes the weak$*$ convergence of measures. The metric structure induced on $\mathcal{P}(M)$ is extremely rich and will play an important role in the sequel. Many properties of this space are discussed in chapter 1 and chapter 4.

We will always consider a "length" cost $c(x,y)$, namely

$$c(x,y) = \inf_{\substack{\gamma \in AC([0,1],M) \\ \gamma(0)=x, \gamma(1)=y}} \int_0^1 L(\gamma(t), \dot\gamma(t)) dt.$$

In this case Brenier introduced a third formulation for the mass transportation problem (see [8, 19]) which is the following: minimize the cost

$$K(\rho, E) := \int_0^1 \int_M L(x, v(x,t)) d\rho_t dt,$$

among the pairs (ρ, E), with $\rho : [0,1] \to \mathcal{P}(M)$ a curve, E_t a vectorial measure on M defined ρ_t almost everywhere, and $v(x,t)$ the Radon-Nikodym derivative of E with respect to ρ, which satisfy the continuity equation:

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \nabla \cdot E = 0 & on\ (0,1) \times M, \\ \rho_0 = \mu, \quad \rho_1 = \nu. \end{cases} \tag{2}$$

We discuss in details this formulation in chapter 1 and chapter 3.

Mather's minimal measures

A measure $\mu \in \mathcal{M}(TM)$ is said to be closed if for all exact forms ω it results

$$\int_{TM} \langle \omega(x), v \rangle d\mu = 0. \tag{3}$$

According to this definition we set

$$\mathcal{M}^c = \left\{ \mu \in \mathcal{M}(TM) \mid \mu \text{ is closed, } \mu \in \mathcal{P}(TM), \int_{TM} L(x,v)d\mu < \infty \right\}.$$

To each measure $\mu \in \mathcal{M}^c$ we can associate the homology class of μ which we will denote by $\rho(\mu) \in H_1(M, \mathbb{R})$ (by duality with H^1, see section 2.4). Indeed, thanks to the fact that L is superlinear and $\int_{TM} L(x,v)d\mu < \infty$, μ acts in a natural way on the set of the closed 1-forms on M by

$$\omega \mapsto \int_{TM} \langle \omega(x), v \rangle d\mu \tag{4}$$

and thanks to condition (3) this action passes to the quotient by the exact forms.

Once we fix an homology class $[h]$, Mather's variational problem amounts to:

$$\beta([h]) := \min_{\mathcal{M}^c} \left\{ \int_{TM} L(x,v)d\mu \mid \rho(\mu) = [h] \right\}. \tag{5}$$

The relevance of Mather's theory is due to the fact that it is a theory of invariant sets and measures under assumptions extremely weaker than those of KAM theory. Actually, it is strictly related to the weak KAM theory developed by Fathi (see [43]).

A remarkable property of problem (5) is the following: we minimize an action functional which depends on L on measures which are merely closed. However it turns out that the minimal measures are also invariant for the flow associated to the Lagrangian L, see for example [7, 26, 44]. We provide some basic tools of Mather's theory in chapter 2.

Description of the main results

In this work we address to the two main following questions:

1. Is it possible to describe Mather minimal measures in terms of the fundamental objects of the theory of optimal transportation?

2. Is it possible to describe the Monge-Kantorovich problem as

$$\min \left\{ \int_{TM} L(x,v)d\mu \mid \mu \in \mathcal{K} \right\},$$

   where $\mathcal{K}$ is a suitable convex subset of $\mathcal{P}(TM)$?

Problem 1 is investigated in chapter 3. In particular we extend some results of [7] to the case of Lagrangians not necessarily homogeneous in v and investigate the relations with the mass transportation. Then we give a description of minimal measures in terms of curves on the space of probability measures which are solutions of equation (2). This provides in some sense an Eulerian description of the minimal measures as opposed to the Lagrangian description of [7]. More in details in section 3.1 we show that problem (5) is equivalent to a variational problem on the space $\mathcal{N}_1(M)$ of normal 1-currents (thus extending some result of [7] where this equivalence was proved for $L(x,v) = g(v,v)$). Due to the non-homogeneity of L, the definition of the action functional on the space of normal 1-currents requires some justifications and some technical accuracy (see Lemma 3.1.1 and Lemma 3.1.3). The definition we give is mostly inspired by the work of Brenier on transport problems.

As by-product of this first equivalence we obtain that each minimal measure is supported on a graph. This last result was originally proved by Mather; here we provide an extremely simple proof.

In section 3.2 we prove that the problem on currents, introduced in section 3.1, is in turn equivalent to the problem of minimizing a convex functional on the set of periodic solutions of equation (2), which satisfy a suitable topological constraint. The idea of this last equivalence stems from the study of the transport density (see for example [12, 34]). In the same section we also extend this equivalence to a larger set of solutions of equation (2) which are in some sense quasi-periodic curves in the space of probability measures. The problem introduced in section 3.2 is very similar to Brenier's formulation of the Monge-Kantorovich problem and, roughly speaking, is obtained to a problem similar to the Brenier's formulation of the mass transportation replacing the boundary constraint in (2) with a suitable homological constraint.

Problem 2 is investigate in chapter 4. In particular we first investigate the following problem:

$$\mathbf{M} := \min\left\{\int_{TM} L(x,v)d\mu \mid \mu\in\mathcal{P}(TM),\ \partial p(\mu) = f^+ - f^-\right\}, \tag{6}$$

where $f^+, f^- \in \mathcal{P}(M)$, and $p(\mu)$ is the current associated to μ by (4) (for the basic facts about currents we refer to the appendix A). In chapter 4 we prove that problem $\mathbf{M}$ is equivalent to the Monge-Kantorovich one in the case $L(x,v) = g(v,v)^{1/2}$. Then we prove that for superlinear Lagrangians the convex set $\{\mu \in \mathcal{P}(TM) : \partial p(\mu) = f^+ - f^-\}$ is too big to describe the Monge-Kantorovich problem. Therefore we have to consider an additional constraint which in general leads to the strict subfamily

$$\Lambda \subset \left\{\mu\in\mathcal{P}(TM) \mid \partial p(\mu) = f^+ - f^-\right\}.$$

In Theorem 4.0.15 is proved the following equality:

$$\mathbf{MK} = \min\left\{\int_{TM} L(x,v)d\mu \mid \mu\in\Lambda\right\}.$$

This result is obtained adding an extra "time" dimension, passing to a dual problem on currents as for the Mather problem and then by a disintegration argument.

In order to this strategy works, we establish an Hamiltonian duality formula for the transportation (Theorem 4.0.9) which on the other hand allow us to state the equivalence of Monge-Kantorovich problem with Brenier's formulation for cost associated to general Lagrangians.

Finally, we show that in the case of $f^+ = f^-$, in which case the current $p(\mu)$ is closed, the additional constraint Λ does not play any role, see Lemma 4.0.18, and requiring that $\rho(p(\mu)) = h$ problem $\mathbf{M}$ reduces to the Mather problem $\beta(h)$. Finally, as a by-product of the problem studied in chapter 4, we obtain a characterization of the Mañé critical value, see Proposition 4.0.19, which in our opinion deserves some attention by itself.

The investigation of the minimality properties of the Eulerian representation of Mather's minimal measures brought us to investigate some possible definition of barycenters of probability measures on the manifold M. The attempt to use a Kirszbraun type theorem in this setting leads to the counterexample we present in section 5.3.3. More precisely, in Chapter 5 we discuss some geometric properties of the space of probability measures endowed with the 2-Wasserstein distance, especially with respect to curvature properties, showing that the assumption of target space of global bounded curvature in the Kirszbraun theorem stated in [60] cannot be reduced to a local one.

The main results of this work are contained in the two forthcoming papers [33, 53].

After this thesis was completed, the advisors informed the author of the existence of the preprint [9] of Bernard and Buffoni which also deals with the connections between mass transportation problems and Mather theory addressing some questions related to those studied in this thesis. In fact, in that paper the authors exploit the instruments of the weak KAM theory of Albert Fathi to introduce transport problem related to a generic Lagrangian (even time depending) and to relate this transport problems to the Mather's one. In particular, under some more regularity assumptions, in [9] is proved that every Eulerian representation ρ of a minimal measure μ is also locally minimal for the mass transportation cost. More precisely, if we set

$$c_s^t(x,y) = \inf_{\substack{\gamma \in AC([s,t],M) \\ \gamma(s)=x, \gamma(t)=y}} \int_s^t L(\gamma(\tau), \dot\gamma(\tau))d\tau$$

and take the mass transportation cost

$$C_s^t(\mu,\nu) = \min\left\{ \int_{M\times M} c_s^t(x,y)d\gamma(x,y) \mid \gamma \in \mathcal{P}(M\times M), \pi^1_\sharp\gamma = \mu, \pi^2_\sharp\gamma = \nu \right\},$$

then every Eulerian representation is locally minimal in the sense that for all $s \le t \le \tau$ it results

$$C_s^\tau(\rho_s,\rho_\tau) = C_s^t(\rho_s,\rho_t) + C_t^\tau(\rho_t,\rho_\tau).$$

List of symbols

$\mathcal{B}(X)$ Borel σ-algebra,

$\mathcal{M}(X)$ Radon signed measures,

$\mathcal{M}^+(X)$ Radon positive measures,

$\mathcal{P}(X)$ probability measures, i.e. $\mu \in \mathcal{M}^+(X)$ $s.t.$ $\mu(X) = 1$,

$\mathcal{P}_p(X) := \{\mu \in \mathcal{P}(X) \mid \forall x_0 \in X : \int_X d(x_0, x)^p d\mu(x) < +\infty\}$, probability measures with finite p-moment,

$\mu \llcorner B$ restriction of μ to the Borel set B, i.e. $\mu \llcorner B(A) = \mu(A \cap B)$,

$\mu \otimes \nu$ tensor product of two Radon measures, i.e. $\mu \otimes \nu(A \times B) = \mu(A)\nu(B)$,

δ_x Dirac delta at point x, i.e. $\delta_x(B) = \begin{cases} 1 & x \in B, \\ 0 & x \notin B, \end{cases}$

$\mu \ll \nu$ μ absolutely continuous with respect to ν,

$\mathcal{L}^n$ n-dimensional Lebesgue measure,

$\mathcal{H}^n$ n-dimensional Hausdorff measure,

$\mathcal{M}^n(M)$ vectorial measures on the n-dimensional manifold M, i.e. the dual space of the continuous sections of the tangent bundle TM,

$|\mu|$ total variation measure of μ,

$\|\mu\|$ total mass of μ, i.e. $\|\mu\| = |\mu|(X)$,

$d(A, B) := \inf\{d(x, y) \mid x \in A,\ y \in B\}$,

A_ε ε-neighborhood of A, i.e. the set of points whose distance from A is less than ε,

$diam(A)$ diameter of A, i.e. $diam(A) = \sup\{d(x, y) \mid x, y \in A\}$,

χ_B characteristic function of B, i.e. $\chi_B(x) = \begin{cases} 1 & x \in B, \\ 0 & x \notin B, \end{cases}$

$\mathcal{C}(X)$ real valued continuous functions on X,

$\mathcal{C}_c(X)$ real valued continuous functions on X with compact support,

$\mathcal{C}_0(X)$ real valued continuous functions on X which vanish at infinity, i.e. which satisfy the condition: $\forall \varepsilon > 0\ \exists K \subset X$ compact s.t. $|f(x)| < \varepsilon$ for all $x \notin K$,

$\mathcal{C}_b(X)$ real valued continuous and bounded functions on X,

$\mathrm{Lip}_l(X)$ Lipschitz functions with Lipschitz constant not greater than $l \in \mathbb{R}$,

$\mathrm{Lip}_b(X)$ Lipschitz and bounded functions,

$\langle\cdot,\cdot\rangle$ scalar product or in general duality,

MK Monge-Kantorovich problem, see (1.3),

M transport problem on currents, see (1.44),

$\beta([h])$ Mather's beta function, see (2.19),

B Brenier's problem, see (1.43) and (4.4),

$\mathcal{P}$ dual problem on currents, see (3.16),

Φ_s Lagrangian flow, see Definition 2.1.1,

$\mathbb{T}^n$ n-dimensional flat torus, i.e. $\mathbb{T}^n = \frac{\mathbb{R}^n}{\mathbb{Z}^n}$.

Chapter 1

Mass Transportation problem

1.1 The Monge-Kantorovich problem

Assume that we are given a pile of soil and an excavation that we want to fill up with the soil. In 1781 Monge posed the question to find an optimal way to do this. We can model the pile of soil and the excavation by two measures μ, ν over a space X, which we will always assume to be a proper metric space, i.e. such that closed balls are compact. We denote by d the distance function on X. The pile and the excavation must have the same volume, so that we can assume that μ and ν are both probability measures.

We consider a measurable map $t : X \to X$ as a "transport" between μ and ν if the amount of mass of ν on a region B of X is the same coming from μ through the map t. Hence, if we consider a Borel set $B \subset X$ we require that

$$\nu(B) = \mu(t^{-1}(B)). \tag{1.1}$$

In other words we have that ν is the image measure of μ through the map t. We use the notation $t_\sharp\mu = \nu$ according to the following definition.

Definition 1.1.1 (Push-forward). *Let $t : X \to Y$ be a measurable map, and $\mu \in \mathcal{M}(X)$. Then $t_\sharp\mu$ is the measure on Y defined as follows:*

$$\textit{for every Borel set } B \subset X : t_\sharp\mu(B) = \mu(t^{-1}(B)).$$

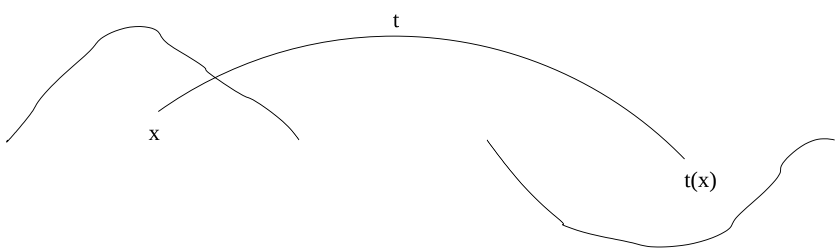


Figure 1.1: Transport problem.

The elementary work to move a particle from the location x to the destination $t(x)$ is given by $d(x, t(x))d\mu(x)$, so that the total cost amounts to $\int_X d(x, t(x))d\mu$. Hence, Monge problem can be written as follows:

$$\inf \left\{ \int_X d(x, t(x))d\mu \mid \nu = t_\sharp\mu \right\}, \tag{1.2}$$

where the unknown is the transport function t. Actually, Monge problem was originally formulated for $X = \mathbb{R}^n$, d the Euclidean distance and $\mu, \nu \ll \mathcal{L}^n$. The problem can be also formulated in the case in which the measure ν is over another space Y and replacing the distance with a cost function $c : X \times Y \to [0, +\infty]$, which we will always consider at least lower semicontinuous (l.s.c.).

Observe that the Monge problem is not always well posed. In fact if we consider for example the measures $\mu = \delta_x$ and $\nu = \frac{1}{2}(\delta_y + \delta_z)$, the Monge transport problem has no solutions simply because there is no map t such that $t_\sharp \mu = \nu$. Moreover, because of the non-linearity of the cost with respect to t, existence of minimizers in (1.2) is a difficult matter and the first rigorous existence theorems are relatively recent, see ([23, 39, 80, 81]), despite the long history of the problem. In order to avoid these difficulties the problem can be reformulated in its Kantorovich relaxed form. The basic idea is to consider a measure $\gamma(x, y)$ which, roughly speaking, establishes the amount of mass of the point x to be allocated at the destination y. More precisely we require that for a Borel set $B \subset Y$, the measure $\gamma \in \mathcal{P}(X \times Y)$ satisfies:

$$\nu(B) = \gamma(X \times B) = \gamma(\pi_2^{-1}(B)) = (\pi_2)_\sharp \gamma(B),$$

where π_2 is the projection of $X \times Y$ on the second factor. With respect to the other projection π_1 we also require that $\mu = (\pi_1)_\sharp \gamma$. If $c(x, y)$ is the cost to move the particle in x to the position y, the Monge-Kantorovich problem amounts to

$$\mathbf{MK} := \min \left\{ \int_{X \times Y} c(x, y) d\gamma(x, y) \mid \gamma \in \mathcal{P}(X \times Y) \ , (\pi_1)_\sharp \gamma = \mu, \ (\pi_2)_\sharp \gamma = \nu \right\}. \tag{1.3}$$

The admissible measures γ for problem **MK** are called transport plans. Observe that if t is admissible for the Monge problem then the measure $\gamma = (id \times t)_\sharp \mu$ is a transport plan for **MK**. Furthermore, the class of transport plans is never empty as it always contains $\mu \otimes \nu$. Note that the Kantorovich problem **MK** is now linear and existence is quite easy to obtain, see Theorem 1.1.7. Actually, Monge problem (1.2) can be considered as a deterministic problem and **MK** as a probabilistic version of it. These optimal transport problems, also known as Monge-Kantorovich problems, have been very intensively studied in the last 10 years. For a detailed discussion of these problems, their history and applications we refer the reader for example to [1, 19, 37, 48, 72, 73, 82].

Let us now state some preliminary lemmas.

Lemma 1.1.2. *Let $u : X \to [a, +\infty]$, $u \neq +\infty$.*
For $n \geq 1$, define $u_n(x) = \inf\{u(y) + nd(x, y) |\ y \in X\}$. Then

$$u_n \in \mathrm{Lip}_n(X, d), u_n \leq u_{n+1} \leq u.$$

Moreover we have

$$\lim_{n \to +\infty} u_n(x) = u(x),$$

whenever x is a point of lower semicontinuity for u.

Proof. The fact that u_n is n-Lipschitz and that $u_n \leq u_{n+1} \leq u$ follows directly from the definition of u_n. Let x be a point of lower semicontinuity for u. For every $n \geq 1$ we can find $x_n \in X$ such that

$$a + nd(x_n, x) \leq u(x_n) + nd(x_n, x) < u_n(x) + \frac{1}{2^n}. \tag{1.4}$$

If $\lim_{n \to +\infty} u_n(x) = +\infty$ the statement is true. Otherwise, by (1.4) we infer that $\lim_{n \to +\infty} x_n = x$. Hence, by lower semicontinuity we have

$$u(x) \leq \liminf_{n \to +\infty} u(x_n) \leq \lim_{n \to +\infty} u_n(x).$$

□

Remark 1.1.3. If we pose $\bar{u}_n(x) = \min\{u_n(x), n\}$ it results that the functions $\bar{u}_n$ satisfy all the claims of the above Lemma 1.1.2. Therefore, we can approximate u also by Lipschitz and bounded functions.

Corollary 1.1.4. *Lipschitz functions are dense in* $(\mathcal{C}_b(X), \|\cdot\|_\infty)$.

Proof. By Lemma 1.1.2 we find an increasing sequence of Lipschitz function u_n which converge pointwise to u. Applying the so called Dini's lemma, see for example Theorem 7.2.2 in [35], it results that the convergence is in fact uniform on compact subsets. □

Definition 1.1.5 (Narrow convergence). *A sequence μ_n in $\mathcal{P}(X)$ is narrowly convergent to $\mu \in \mathcal{P}(X)$, and we will use the notation $\mu_n \rightharpoonup \mu$, if*

$$\forall f \in \mathcal{C}_b(X) : \lim_{n\to+\infty} \int_X f(x)d\mu_n = \int_X f(x)d\mu. \tag{1.5}$$

The following lemma provides a sufficient condition to the narrow convergence.

Lemma 1.1.6 (Prokhorov). *Let (μ_n) be a sequence in $\mathcal{P}(X)$ such that*

$$\forall \varepsilon > 0 : \exists K_\varepsilon \subset X \text{ compact s.t. } \forall n \geq 1 : \mu_n(X \setminus K_\varepsilon) < \varepsilon. \tag{1.6}$$

Then there exists $\mu \in \mathcal{P}(X)$ and a subsequence s.t. $\mu_{h(n)} \rightharpoonup \mu$.

Proof. Let $\frac{1}{n} > 0$, by (1.6) we find compact sets K_n s.t. $\mu_h(X \setminus K_n) < \frac{1}{n}$. Since every proper metric space is σ-compact, we can suppose that the sequence (K_n) is increasing. Define $\mu_h^n = \chi_{K_n}\mu_h$. By weak-* compactness of the measures (μ_h^n) for each fixed n, and using a diagonal argument, we can find a subsequence and measures σ_n concentrated on K_n such that $\mu_{k(h)}^n \overset{*}{\rightharpoonup} \sigma_n$ for $h \to +\infty$. Since these measures are concentrated on the same compact set we obtain

$$\forall f \in \mathcal{C}_b(X) : \lim_{h\to+\infty} \int_X f(x)d\mu_{k(h)}^n = \int_X f(x)d\sigma_n.$$

If $m > n$ we have

$$\|\mu_{k(h)}^n - \mu_{k(h)}^m\| = |\mu_{k(h)}^n - \mu_{k(h)}^m|(X \setminus K_n) = |\mu_{k(h)}^n|(X \setminus K_n) \leq \mu_{k(h)}(X \setminus K_n) < \frac{1}{n}.$$

Passing to the limit for $h \to +\infty$ we obtain

$$\|\sigma_n - \sigma_m\| \leq \liminf_{h\to+\infty} \|\mu_{k(h)}^n - \mu_{k(h)}^m\| \leq \frac{1}{n}.$$

Hence (σ_n) is a Cauchy sequence in $(\mathcal{M}(X), \|\cdot\|)$. Then there exists $\sigma \in \mathcal{M}(X)$ such that $\lim_{n\to+\infty} \|\sigma_n - \sigma\| = 0$. Let us now prove that $\mu_{k(h)} \rightharpoonup \sigma$. Fixed $f \in \mathcal{C}_b(X)$ we can evaluate

$$|\int_X fd(\mu_{k(h)} - \sigma)| \leq |\int_X fd(\mu_{k(h)} - \mu_{k(h)}^n)| + |\int_X fd(\mu_{k(h)}^n - \sigma_n)| +$$

$$+|\int_X fd(\sigma_n - \sigma)| \leq \|f\|_\infty \mu_{k(h)}(X \setminus K_n) + |\int_X fd(\mu_{k(h)}^n - \sigma_n)| + \|f\|_\infty \|\sigma_n - \sigma\| \leq$$

$$\leq \|f\|_\infty \frac{1}{n} + |\int_X fd(\mu_{k(h)}^n - \sigma_n)| + \|f\|_\infty \|\sigma_n - \sigma\|.$$

Passing to the limit for $h \to +\infty$ we obtain

$$\limsup_{h\to+\infty} |\int_X fd(\mu_{k(h)} - \sigma)| \leq \frac{1}{n}\|f\|_\infty + \|f\|_\infty \|\sigma_n - \sigma\|.$$

Passing to the limit for $n \to +\infty$ the statement follows. □

Theorem 1.1.7 (Existence of optimal plans). *Let $c : X \times Y \to [0, +\infty]$ be a l.s.c. cost. Then **MK** admits at least a minimizer.*

Proof. Let $c_m \nearrow c$ with c_m continuous and bounded as in Remark 1.1.3 and γ_n a minimizing sequence for **MK**. Let $\varepsilon > 0$. We can find compact sets such that

$$\mu(X \setminus K_1) < \varepsilon \ , \ \nu(Y \setminus K_2) < \varepsilon.$$

Choosing $K = K_1 \times K_2$ we have

$$\gamma_n\left((X \times Y) \setminus K\right) = \int_{X \times Y} \chi_{(X \times Y) \setminus K}(x, y) d\gamma_n \leq$$

$$\leq \int_{X \times Y} [\chi_{X \setminus K_1}(x) + \chi_{Y \setminus K_2}(y)] d\gamma_n = \mu(X \setminus K_1) + \nu(Y \setminus K_2) < 2\varepsilon.$$

By Lemma 1.1.6 we can then assume that $\gamma_n \rightharpoonup \gamma$ with $\gamma \in \mathcal{P}(X)$. Observe that γ is a transport plan. In fact if $f \in \mathcal{C}_0(X)$ we have

$$\int_{X \times Y} f(x) d\gamma(x, y) = \lim_{n \to +\infty} \int_{X \times Y} f(x) d\gamma_n(x, y) = \lim_{n \to +\infty} \int_X f(x) d\mu.$$

Therefore $(\pi_1)_\sharp \gamma = \mu$. Analogously we have that $(\pi_2)_\sharp \gamma = \nu$. Since $\gamma_n \rightharpoonup \gamma$ we deduce

$$\int_{X \times Y} c_m(x, y) d\gamma = \lim_{n \to +\infty} \int_{X \times Y} c_m(x, y) d\gamma_n \leq \liminf_{n \to +\infty} \int_{X \times Y} c(x, y) d\gamma_n.$$

Letting $m \to +\infty$ the statement follows. □

If μ has no atoms and the cost c is continuous and real valued it is possible to show that **MK** coincide with the infimum of the Monge problem, while in general one has only the inequality **MK** $\leq$ (1.2), (see [1]). A detailed discussion of this problem is contained for example in [72]. We make use of the following

Lemma 1.1.8. *Let γ be an optimal transport plan for **MK**. Then for every Borel set $A \subset X \times Y$ the measure $\gamma_A := \gamma \llcorner A$ is an optimal transport plan between its marginals.*

Proof. Define $\mathcal{C}(\gamma) := \int_{X \times Y} c(x, y) d\gamma$. We observe that $\mathcal{C}(\gamma)$ is linear with respect to γ. Suppose that there exists a measure $\tilde{\gamma}$ with the same marginal as γ_A such that $\mathcal{C}(\tilde{\gamma}) < \mathcal{C}(\gamma)$. Then the measure $\eta = \tilde{\gamma} + \gamma \llcorner A^c$ has the same marginals of γ and $\mathcal{C}(\eta) < \mathcal{C}(\gamma)$ contradicting the fact that γ is optimal. □

1.2 Kantorovich duality

In this section we state some duality results for the Monge-Kantorovich problem. We will need a basic duality theorem in convex analysis, for the proof of which we refer to Theorem I.11 in [21].

Theorem 1.2.1. *Let E be a normed vector space, E^* its dual space and $A, B : E \to \mathbb{R} \cup \{+\infty\}$ two convex functions. Assume that there exists $z_0 \in E$ such that A is continuous at z_0 and $A(z_0), B(z_0)$ belongs to $\mathbb{R}$. Then*

$$\inf_{z \in E} \{A(z) + B(z)\} = \max_{z^* \in E^*} \{-A^*(-z^*) - B^*(z^*)\},$$

where A^*, B^* denote the convex conjugate of A and B respectively.

The Kantorovich dual formulation of the mass transportation problem is given by the following result.

Theorem 1.2.2. *Let $c(x,y)$ be a positive continuous cost on $X \times Y$. Then the minimum value of the Monge-Kantorovich problem **MK** is equal to*

$$\mathcal{D} := \sup_{(\varphi,\psi)\in\mathcal{C}_0(X)\times\mathcal{C}_0(Y)} \left\{ \int_X \varphi(x)d\mu + \int_Y \psi(y)d\nu \mid \varphi(x)+\psi(y) \le c(x,y) \right\}. \tag{1.7}$$

Proof. Let γ be a transport plan for **MK**. If $\varphi + \psi$ are admissible for $\mathcal{D}$, integrating we obtain

$$\int_{X\times Y} [\varphi(x)+\psi(y)]d\gamma(x,y) \le \mathbf{MK} \Rightarrow \int_X \varphi(x)d\mu + \int_Y \psi(y)d\nu \le \mathbf{MK}. \tag{1.8}$$

Taking the supremum with respect to (φ,ψ) we have that $\mathcal{D} \le \mathbf{MK}$.

Set $E = (\mathcal{C}_0(X\times Y), \|\cdot\|_\infty)$, so that $E^* = (\mathcal{M}(X\times Y), \|\cdot\|)$. First let us suppose that X, Y are compact. Define $A, B : E \to [0,+\infty]$ by

$$A(u) = \begin{cases} 0 & \text{if } u(x,y) \ge -c(x,y), \\ +\infty & \text{otherwise,} \end{cases}$$

$$B(u) = \begin{cases} \int_X \varphi(x)d\mu + \int_Y \psi(y)d\nu & \text{if } \exists\ (\varphi,\psi)\in\mathcal{C}_0(X)\times\mathcal{C}_0(Y) \ s.t.\ u(x,y)=\varphi(x)+\psi(y), \\ +\infty & \text{otherwise .} \end{cases}$$

A, B are convex functions with $A(0) = B(0) = 0$. Replacing the cost c with $c+\delta$, for any $\delta > 0$, if necessary, we can suppose that the cost $c(x,y)$ is strictly positive. Hence we have that

$$\|u\|_\infty \le \varepsilon \Rightarrow A(u) = 0.$$

By Theorem 1.2.1 we find that

$$\inf_{u\in E}\{A(u)+B(u)\} = \max_{u^*\in E^*}\{-A^*(-u^*) - B^*(u^*)\}. \tag{1.9}$$

The left side of (1.9) is

$$\inf \left\{ \int_X \varphi(x)d\mu + \int_Y \psi(y)d\nu \mid \varphi(x)+\psi(y) \ge -c(x,y) \right\} = \tag{1.10}$$

$$= -\sup \left\{ \int_X \varphi(x)d\mu + \int_Y \psi(y)d\nu \mid \varphi(x)+\psi(y) \le c(x,y) \right\}. \tag{1.11}$$

Put $u^* = \lambda \in \mathcal{M}(X\times Y)$. Let us compute

$$A^*(-\lambda) = \sup_{u\in E}\{\langle -\lambda, u\rangle - A(u)\} =$$

$$= \sup_{u\in\mathcal{C}_0(X\times Y)} \left\{ -\int_{X\times Y} u(x,y)d\lambda \mid u(x,y) \ge -c(x,y) \right\} =$$

$$= \sup_{u\in\mathcal{C}_0(X\times Y)} \left\{ \int_{X\times Y} u(x,y)d\lambda \mid u(x,y) \le c(x,y) \right\}.$$

If λ has a negative part we can find a negative function $v \in \mathcal{C}_0(X\times Y)$ such that $\int_{X\times Y} v(x,y)d\lambda > 0$. Then, choosing $u = rv$ and letting $r \to +\infty$, we find that $A^*(-\lambda) = +\infty$. In the case in which $\lambda \ge 0$, since the space $X \times Y$ is σ-compact (for the assumptions made at the beginning) we can approximate c from below by continuous and compactly supported functions and we obtain that $A^*(-\lambda) = \int_{X\times Y} c(x,y)d\lambda$. Hence

$$A^*(-\lambda) = \begin{cases} \int_{X\times Y} c(x,y)d\lambda & \text{if } \lambda \ge 0, \\ +\infty & \text{otherwise .} \end{cases} \tag{1.12}$$

On the other hand we have

$$B^*(\lambda) = \sup_{u \in \mathcal{C}_0(X\times Y)} \left\{ \int_{X\times Y} u(x,y)d\lambda - B(u) \right\} =$$

$$= \sup_{u \in \mathcal{C}_0(X\times Y)} \left\{ \int_{X\times Y} u(x,y)d\lambda - \int_X \varphi(x)d\mu - \int_Y \psi(y)d\nu \mid u(x,y) = \varphi(x) + \psi(y) \right\} =$$

$$= \sup_{(\varphi,\psi)\in \mathcal{C}_0(X)\times\mathcal{C}_0(Y)} \left\{ \int_{X\times Y} [\varphi(x)+\psi(y)]d\lambda - \int_X \varphi(x)d\mu - \int_Y \psi(y)d\nu \right\}.$$

Hence, arguing as for A^* we obtain

$$B^*(\lambda) = \begin{cases} 0 & \text{if } (\pi_1)_\sharp\lambda = \mu, (\pi_2)_\sharp\lambda = \nu, \\ +\infty & \text{otherwise} . \end{cases} \tag{1.13}$$

By (1.10), (1.12) and (1.13) the equality in (1.9) becomes

$$-\sup\left\{ \int_X \varphi(x)d\mu + \int_Y \psi(y)d\nu \mid \varphi(x)+\psi(y) \le c(x,y) \right\} =$$

$$= \sup\left\{ -\int_{X\times Y} c(x,y)d\lambda \mid \lambda \ge 0, (\pi_1)_\sharp\lambda = \mu, (\pi_2)_\sharp\lambda = \nu \right\} \Leftrightarrow$$

$$\Leftrightarrow \mathcal{D} = \mathbf{MK}.$$

In the general case, since X, Y are σ- compact, let $c_n \nearrow c$ with c_n continuous with compact support $X_n \times Y_n$. Let γ be an optimal transport plan and set $\gamma_n = \gamma \llcorner X_n \times Y_n$. By Lemma 1.1.8 γ_n is an optimal plan between its marginals. We have

$$\mathbf{MK} = \int_{X\times Y} c(x,y)d\gamma = \lim_{n\to +\infty} \int_{X\times Y} c_n(x,y)d\gamma = \int_{X_n\times Y_n} c_n(x,y)d\gamma_n \le \mathcal{D}.$$

□

As noticed for example in [3], the statement of Theorem 1.2.2 is true also if the cost $c(x,y)$ is only lower semicontinuous. In fact, one can pointwise approximate c from below by an increasing sequence of continuous and bounded functions c_n as in Remark 1.1.3. If γ_n are optimal plans for the transportation with the cost c_n, by Lemma 1.1.6 we can assume that $\gamma_n \rightharpoonup \gamma$. Observe that γ is a transport plan between μ and ν. Fixed $\delta > 0$ and $m \le n$, by (1.7) we find

$$\varphi_n(x) + \psi_n(y) \le c_n(x,y) \le c(x,y)$$

such that

$$\int_{X\times Y} c_m d\gamma_n \le \int_{X\times Y} c_n d\gamma_n \le \int_X \varphi_n d\mu + \int_Y \psi_n d\nu + \delta \le \mathcal{D} + \delta,$$

where $\mathcal{D}$ is defined as the value of the supremum in (1.7). Passing to the limit as $n \to +\infty$ we obtain that $\int_{X\times Y} c_m d\gamma \le \mathcal{D} + \delta$.

Passing to the limit as $m \to +\infty$ and then letting $\delta \to 0^+$ we conclude, since the opposite inequality is immediate as in (1.8).

Another remarkable fact is that the value of the supremum in (1.7) does not change if we replace the space $\mathcal{C}_0(X)$ and $\mathcal{C}_0(Y)$ by L^1_μ and L^1_ν respectively.

Proposition 1.2.3.

$$\sup_{(\varphi,\psi)\in\mathcal{C}_0(X)\times\mathcal{C}_0(Y)}\left\{\int_X \varphi d\mu+\int_Y \psi d\nu \mid \varphi(x)+\psi(y)\le c(x,y)\right\}=$$

$$=\sup_{(\varphi,\psi)\in L^1_\mu\times L^1_\nu}\left\{\int_X \varphi d\mu+\int_Y \psi d\nu \mid \varphi(x)+\psi(y)\le c(x,y)\right\},$$

where the inequality in the last supremum holds up to Borel sets $A\cup B\subset X\times Y$ s.t. $\mu(\pi_1(A))=0$ and $\nu(\pi_2(B))=0$.

Proof. Let $(\varphi,\psi)\in L^1_\mu\times L^1_\nu$ such that

$$\varphi(x)+\psi(y)\le c(x,y), \tag{1.14}$$

and γ be an optimal transport plan between μ and ν. Observe that the inequality (1.14) holds also $\gamma-a.e.$ Indeed, let N,M negligible sets with respect to μ and ν respectively and such that inequality (1.14) holds for $(x,y)\in(X\setminus N)\times(Y\setminus M)$. Since γ has μ and ν as marginals we have

$$\gamma(N\times Y)=\mu(N)=0,\ \gamma(X\times M)=\nu(M)=0.$$

and then $\gamma(X\setminus N\times Y\setminus M)=1$. We can evaluate

$$\int_X \varphi d\mu+\int_Y \psi d\nu=\int_{X\times Y}[\varphi(x)+\psi(y)]d\gamma\le\int_{X\times Y}c(x,y)d\gamma.$$

Taking the supremum with respect to $(\varphi,\psi)\in L^1_\mu\times L^1_\nu$ we deduce the inequality

$$\sup_{(\varphi,\psi)\in L^1_\mu\times L^1_\nu}\left\{\int_X \varphi d\mu+\int_Y \psi d\nu \mid \varphi(x)+\psi(y)\le c(x,y)\right\}\le\mathbf{MK}.$$

By Theorem 1.2.2 the statement follows since the opposite inequality is immediate. □

If $X=Y$ and the cost c is the distance d on X, then the duality formulation becomes simpler.

Theorem 1.2.4. *If $c(x,y)=d(x,y)$ then for every $\mu,\nu\in\mathcal{P}_1(X)$ the minimum value of the Monge-Kantorovich problem **MK** is equal to*

$$\sup\left\{\int_X ud(\mu-\nu)\mid u\in \mathrm{Lip}_1(X)\right\}.$$

Proof. Let $u\in\mathrm{Lip}_1(X)$ and take $\varphi=u,\ \psi=-u$. Since $\mu,\nu\in\mathcal{P}_1(X)$ it follows at once that $(\varphi,\psi)\in L^1_\mu\times L^1_\nu$. By Proposition 1.2.3 and Theorem 1.2.2 we deduce that

$$\sup\left\{\int_X ud(\mu-\nu)\mid u\in \mathrm{Lip}_1(X)\right\}\le\mathbf{MK}.$$

Vice versa, let $(\varphi,\psi)\in\mathcal{C}_0(X)\times\mathcal{C}_0(Y)$ such that

$$\varphi(x)+\psi(y)\le d(x,y). \tag{1.15}$$

Let M,N be the supports of μ and ν respectively. Define

$$\forall x\in M:\hat\varphi(x)=\inf_{y\in N}\{d(x,y)-\psi(y)\}. \tag{1.16}$$

By (1.15) we have

$$\forall x\in M:\varphi(x)\le\hat\varphi(x), \tag{1.17}$$

and

$$\forall x \in M, \forall y \in N : \hat{\varphi}(x) + \psi(y) \leq d(x,y). \tag{1.18}$$

If we define

$$\forall y \in N : \hat{\psi}(y) = \inf_{x \in M}\{d(x,y) - \hat{\varphi}(x)\}, \tag{1.19}$$

by (1.18) we also deduce that

$$\forall y \in N : \psi(y) \leq \hat{\psi}(y). \tag{1.20}$$

Observe that by (1.19) we have

$$\forall x \in M, \forall y \in N : \hat{\varphi}(x) + \hat{\psi}(y) \leq d(x,y). \tag{1.21}$$

By (1.16) and (1.20) it results $\hat{\varphi}(x) \geq \inf_{y \in N}\{d(x,y) - \hat{\psi}(y)\}$, for every $x \in M$. Hence, by (1.21) we obtain

$$\forall x \in M : \hat{\varphi}(x) = \inf_{y \in N}\{d(x,y) - \hat{\psi}(y)\}. \tag{1.22}$$

Now we use the right side of (1.22) to define $\hat{\varphi}(x)$ for every $x \in X$. Let us also prove that $\hat{\varphi} \in \mathrm{Lip}_1(X)$. Indeed, let $x_1, x_2 \in X$ and $\varepsilon > 0$. By (1.22) there exists $y_1 \in N$ such that

$$d(x_1, y_1) - \hat{\psi}(y_1) < \hat{\varphi}(x_1) + \varepsilon.$$

Hence

$$\hat{\varphi}(x_2) - \hat{\varphi}(x_1) < \hat{\varphi}(x_2) - d(x_1,y_1) + \hat{\psi}(y_1) + \varepsilon \leq$$
$$\leq d(x_2,y_1) - \hat{\psi}(y_1) - d(x_1,y_1) + \hat{\psi}(y_1) + \varepsilon \leq d(x_1,x_2) + \varepsilon.$$

Sending $\varepsilon \to 0^+$ we conclude. Similarly it results $\hat{\psi} \in \mathrm{Lip}_1(X)$. Observe that by (1.19) we have $\hat{\varphi}(x) + \hat{\psi}(x) \leq 0$, for every $x \in M$. Taking $u = -\hat{\psi}$ and by (1.17), (1.20) we obtain

$$\int_M \varphi d\mu + \int_N \psi d\nu \leq \int_M \hat{\varphi} d\mu + \int_N \hat{\psi} d\nu \leq \int_X u d(\mu - \nu).$$

Taking the supremum with respect to (φ, ψ) the statement follows. □

Actually, the supremum in Theorem 1.2.4 is attained and the optimal u are called Kantorovich potentials.

Theorem 1.2.5 (Kantorovich potentials). *Let $c(x,y) = d(x,y)$. Then for every $\mu, \nu \in \mathcal{P}_1(X)$, there exists $u \in \mathrm{Lip}_1(X)$ such that*

$$\boldsymbol{MK} = \int_X u d(\mu - \nu).$$

Moreover, if $\mu, \nu \in \mathcal{P}(X)$ are compactly supported then we can find a Kantorovich potential u which satisfies

$$\forall x \in M : u(x) = \min_{y \in N}\{u(y) + d(x,y)\} \ , \ \forall y \in N : u(y) = \max_{x \in M}\{u(x) - d(x,y)\}, \tag{1.23}$$

where M, N are the supports of μ, ν respectively.

Proof. Let $u_n \in \mathrm{Lip}_1(X)$ be a maximizing sequence. Adding a constant to u_n we can assume that $u_n(x_0) = 0$ at some point $x_0 \in X$. By the Ascoli-Arzelà theorem we can also assume that u_n converges pointwise to a function $u \in \mathrm{Lip}_1(X)$, with uniform convergence on compact subsets. Since $\mu, \nu \in \mathcal{P}_1(X)$, if $\varepsilon > 0$ we can find a large ball B'_h such that

$$\int_X u d(\mu - \nu) \leq \int_{B'_h} u d(\mu - \nu) + \varepsilon = \lim_{n \to +\infty} \int_{B'_h} u_n d(\mu - \nu) + \varepsilon \leq$$

$$\leq \liminf_{n\to+\infty} \int_X u_n d(\mu - \nu) + 2\varepsilon.$$

Sending $\varepsilon \to 0^+$ the first claim of the theorem follows. Let us now suppose that M, N are compact. Then the couple $(\varphi, \psi) = (u, -u)$ is maximizing for (1.7). As in the proof of Theorem 1.2.4 we can consider the functions $\hat{\varphi}, \hat{\psi} \in \mathrm{Lip}_1(X)$ defined by (1.16) and (1.19) respectively. Observe that by (1.19) it results $\hat{\varphi}(x) + \hat{\psi}(x) \leq 0$, for every $x \in M$. Actually, we prove that in fact $\hat{\varphi}(z) + \hat{\psi}(z) = 0$, for every $z \in M$. Indeed, if $\hat{\varphi}(z) + \hat{\psi}(z) < 0$, by compactness and (1.22) we find $x \in M, y \in N$ such that

$$\hat{\varphi}(z) = d(z, y) - \hat{\psi}(y)\ ,\ \hat{\psi}(z) = d(x, z) - \hat{\varphi}(x).$$

Hence

$$\hat{\varphi}(z) + \hat{\psi}(z) = d(z, y) - \hat{\psi}(y) + d(x, z) - \hat{\varphi}(x) \Rightarrow$$

$$\Rightarrow d(x, z) + d(z, y) < \hat{\varphi}(x) + \hat{\psi}(y) \leq d(x, y),$$

contradicting the triangular inequality. Thus, putting $\hat{u} = -\hat{\psi}$ and recalling (1.22), for $x \in M$ we have that

$$\hat{u}(x) = -\hat{\psi}(x) = \hat{\varphi}(x) = \min_{y\in N}\{d(x, y) + \hat{u}(y)\},$$

while for $y \in N$ it results

$$\hat{u}(y) = -\hat{\psi}(y) = -\min_{x\in M}\{d(x, y) - \hat{\varphi}(x)\} = -\min_{x\in M}\{d(x, y) + \hat{\psi}(x)\} =$$

$$= -\min_{x\in M}\{d(x, y) - \hat{u}(x)\} = \max_{x\in M}\{\hat{u}(x) - d(x, y)\}.$$

□

For a function $\varphi : X \to \mathbb{R}$ we define the c-transform as follows:

$$\varphi^c(y) = \inf\{c(x, y) - \varphi(x) \mid x \in X\}. \tag{1.24}$$

For general cost c the existence of maximizing pairs (φ, ψ) in Theorem 1.2.2 is not always assured. However, in many cases the supremum is just attained on couples of the form (φ, φ^c), (see for example [3, 48, 67]).

1.3 Kantorovich potentials as distance functions

Kantorovich potentials are an important tool in all the existence proofs for the Monge problem that are available at the moment. A problem analogous to the Monge-Kantorovich one is used the study the behaviour of sand piles, see for example [24]. In that setting it turns out that the Kantorovich potential is the distance from the boundary of the open set Ω. In this section we study whether or not there exists a closed set Γ such that the Kantorovich potential can be written as $u(x) = d(x, \Gamma)$.

In this section, although we could consider Riemannian manifolds, we take Ω a convex, open and bounded set of $\mathbb{R}^n$, and $f^+, f^- \in \mathcal{P}(\Omega)$ with M, N their respective supports. Therefore Monge problem amounts to

$$\min\left\{\int_\Omega |x - t(x)| df^+ \mid t_\sharp f^+ = f^-\right\}. \tag{1.25}$$

A key lemma to find optimal transport map for problem (1.25) is the following.

Lemma 1.3.1. *Let $u \in \mathrm{Lip}_1(\Omega)$, and $t : \Omega \to \Omega$ be a Borel map such that $t_\sharp f^+ = f^-$. If*

$$u(x) - u(t(x)) = |x - t(x)| \quad \text{for } f^+ - a.e.\ x \in \Omega, \tag{1.26}$$

then:
1) u is a Kantorovich potential,
2) t is an optimal transport map for problem (1.25),
3) every optimal map $\hat{t}$ and Kantorovich potential $\hat{u}$ also satisfy (1.26).

Proof. For any map $t : \Omega \to \Omega$ such that $t_\sharp f^+ = f^-$ and $u \in \mathrm{Lip}_1(\Omega)$ we have

$$\int_\Omega u d(f^+ - f^-) = \int_\Omega u df^+ - \int_\Omega u(t(x)) df^+ \leq \int_\Omega |x - t(x)| df^+. \tag{1.27}$$

On the other hand, if t and u satisfy (1.26), then we have equality in (1.27). This implies assertion 1) and 2). To check 3) take v a Kantorovich potential and t an optimal transport map. Then we have

$$\int_\Omega v d(f^+ - f^-) = \int_\Omega |x - t(x)| df^+ \Leftrightarrow$$
$$\Leftrightarrow \int_\Omega v df^+ - \int_\Omega v(t(x)) df^+ = \int_\Omega |x - t(x)| df^+ \Rightarrow$$
$$\Rightarrow v(x) - v(t(x)) = |x - t(x)| \quad \text{for } f^+ - a.e.\ x \in \Omega.$$

□

Condition (1.26) contains a useful geometric meaning. In fact, let $x, y \in \Omega$ such that

$$u(x) - u(y) = |x - y|. \tag{1.28}$$

If $z \in [x, y]$, since $u \in \mathrm{Lip}_1(\Omega)$ we have

$$|y - z| = |x - y| - |x - z| = u(x) - u(y) - |x - z| \leq u(x) - u(y) + u(z) - u(x) \Rightarrow$$
$$\Rightarrow u(z) = u(y) + |z - y|.$$

So the function u is decreasing with the maximum possible rate along the segment $[x, y]$. Furthermore, by the triangular inequality we find that for every $z_1, z_2 \in [x, y]$ it also results

$$u(z_1) - u(z_2) = |z_1 - z_2|.$$

Indeed

$$|z_1 - z_2| \leq |z_1 - x| + |x - y| + |y - z_2| = u(z_1) - u(x) + u(x) - u(y) + u(y) - u(z_2).$$

Observe that if u is a Kantorovich potential satisfying (1.23) and $x \in M$, there exists $y \in N$ such that $u(x) - u(y) = d(x, y)$, and so the data f^+, f^- are supported on the segments determined by the condition (1.28). We will call these segments transport rays. The precise definition is the following.

Definition 1.3.2 (Transport rays). *A transport ray $R_{x,y}$ is a segment joining x and y such that*
1) $x \in M,\ y \in N, x \neq y$,
2) $u(x) - u(y) = |x - y|$,
3) Maximality: set $a_t = x + t(y - x)$. Then for any $t < 0$ such that $a_t \in M$ we have

$$|u(a_t) - u(y)| < |a_t - y|,$$

and for any $t > 1$ such that $a_t \in N$ we have

$$|u(a_t) - u(x)| < |a_t - x|.$$

We call the points x, y the upper and lower ends of $R_{x,y}$ respectively. Hence, condition (1.26) tells us that any transport map moves the mass along the transport rays, namely that every point on M, N belongs to some transport ray.

Transport rays can be also defined on Riemannian manifolds replacing the segments with geodesics. For the transportation problems on manifolds we refer to [32, 46, 67, 71].

The notion of transport ray was introduced by Evans and Gangbo in [39]. In particular, in their PDE approach to the transportation, Evans and Gangbo derived the system of equations

$$\begin{cases} -div(\mu Du) = f^+ - f^- \\ |Du| = 1 \quad \mu - a.e. \end{cases} \tag{1.29}$$

where $u \in \mathrm{Lip}_1(\Omega)$ is a Kantorovich potential, while μ is a measure called transport density. Actually, roughly speaking, the measure μ establishes the amount of mass to move along a transport ray whose direction is given by Du. The equations in (1.29) are important in many different contexts. For theory and applications we refer to [11, 12, 13, 34, 47, 72]. These equations will be important also for our purposes. The rest of this section is just devoted to discuss this question. The following result gives a positive answer provided that the supports M, N are compact and sufficiently far. We denote by $d(M, N) := \inf\{d(x, y) \ : \ x \in X, \ y \in Y\}$ the distance between M and N.

Theorem 1.3.3 (Potentials as distance). *Let u be a Kantorovich potential which satisfy (1.23), and set $\Sigma_0 = \{x \in \Omega \mid u(x) = 0\}$. If*

$$d(M, N) \geq \min\left(diam(M), diam(N)\right), \tag{1.30}$$

then, up to addition of a constant, we have:

$$u = \begin{cases} d(\cdot, \Sigma_0) & on\ M, \\ -d(\cdot, \Sigma_0) & on\ N. \end{cases} \tag{1.31}$$

Proof. Suppose that $d(M, N) \geq diam(M)$ and let $m = \min_{x \in M} u(x)$. Adding a constant to u we can assume that $m = 0$, so that $u \geq 0$ on M. Fix $x_0 \in M$ such that $u(x_0) = 0$. Hence, for every $x \in M$ we have

$$u(x) = u(x) - u(x_0) \leq |x - x_0| \leq diam(M).$$

By (1.23) and (1.30), for every $y \in N$ we have

$$u(y) = \max_{x \in M}\{u(x) - |x - y|\} \leq diam(M) - d(M, N) \leq 0.$$

Then we have

$$\begin{cases} u \geq 0 & \text{on } M, \\ u \leq 0 & \text{on } N. \end{cases}$$

We obtain the same conclusion arguing in a similar way also if $d(M, N) \geq diam(N)$. Furthermore, the potential u satisfies the following inequalities:

$$\forall x \in M : d(x, \Sigma_0) \leq u(x) \ , \ \forall y \in N : d(y, \Sigma_0) \leq -u(y).$$

Indeed, if $u(x) = 0$ then $d(x, \Sigma_0) = 0$. If $u(x) > 0$, consider a transport ray $R_{a,b}$ with $x \in R_{a,b}$. Hence $u(x) = u(b) + |x - b|$. Now, if $u(b) = 0$, then $d(x, \Sigma_0) \leq |x - b| = u(x)$. Otherwise, by continuity of u along the transport ray, there exists $z \in R_{a,b}$ such that $u(z) = 0$. Then again $d(x, \Sigma_0) \leq |x - z| = u(x)$. The other inequality follows in a similar way. Define the function

$$v = \begin{cases} d(\cdot, \Sigma_0) & \text{on } M, \\ -d(\cdot, \Sigma_0) & \text{on } N. \end{cases}$$

Since $v(x) - v(y) = d(x, \Sigma_0) + d(y, \Sigma_0) \leq d(x, y)$, then $v \in \mathrm{Lip}_1(\Omega)$. Since u is a Kantorovich potential and since $|u(z)| \leq d(z, \Sigma_0)$, we have

$$\int_\Omega v d(f^+ - f^-) \leq \int_\Omega u d(f^+ - f^-) \Rightarrow$$

$$\Rightarrow 0 = \int_\Omega [u + d(\cdot, \Sigma_0)] df^- \leq \int_\Omega [u - d(\cdot, \Sigma_0)] df^+ \leq 0 \Rightarrow$$

$$\Rightarrow u = \begin{cases} d(\cdot, \Sigma_0) & \text{on } M, \\ -d(\cdot, \Sigma_0) & \text{on } N. \end{cases}$$

□

The condition (1.30) is not a necessary condition, as it follows by figure (1.2) in which

$$d(M, N) < \min(diam(M), diam(N)).$$

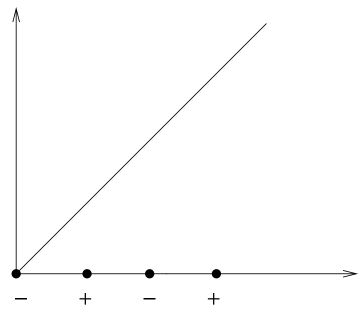


Figure 1.2: - and + denoted Dirac deltas of f^- and f^+ respectively.

Furthermore, there are potentials which cannot be expressed as a distance function (see figure (1.3)). The Kantorovich potentials are uniquely determined, up to addition of a constant, on

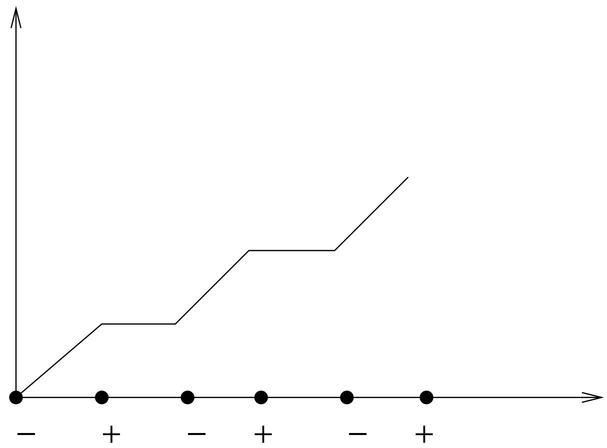


Figure 1.3: A potential which is not a distance.

transport rays, where they increase with maximum rate. In Figure 1.2, Figure 1.3, and Figure 1.4

the transport rays correspond to the segments connecting the Dirac deltas of f^- with those of f^+. Therefore, for distributions of masses as in Figure 1.3 we can also choose the potential $u(x) = x$. However, the situation is more involved since there are also distributions of masses for which none of the potentials is a distance function (see figure (1.4)).

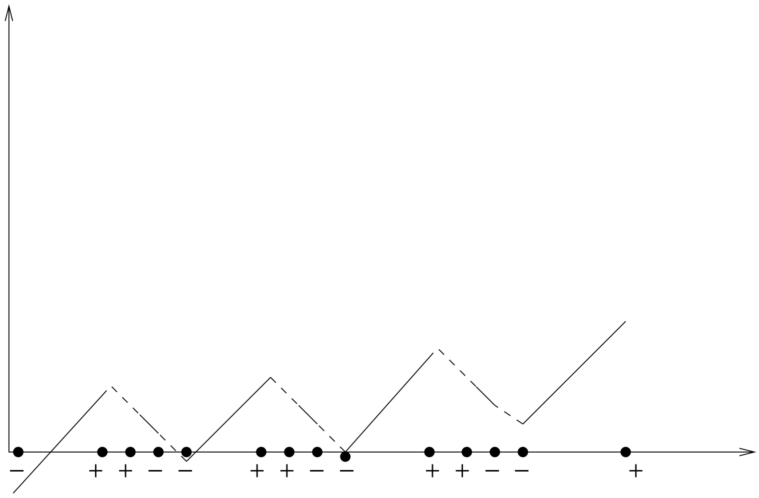


Figure 1.4: The drawn lines in the picture denote a possible Lipschitz continuation of u. It is clear that for any other potential and any continuation we can repeat a similar scheme of masses to lose the distance property (1.31).

At the moment we are no able to produce other geometric conditions on M, N which ensure a distance property as in (1.31). The main difficulty is the change of sign in $f = f^+ - f^-$. Actually, in the case $f \geq 0$, many interesting results are available. We quote [10] which has been our main inspiration for Theorem 1.3.3, [24, 50] and the reference therein for related results.

1.4 The Wasserstein distances

The transport of measures allows to define a distance on the space of probability measures. More precisely, we will define a family of distances by considering the transport with $d(x,y)^p$, $p \geq 1$, as cost.

Definition 1.4.1 (Wasserstein distances). *Let $p \geq 1$ and $\mu, \nu \in \mathcal{P}_p(X)$. We define the p-Wasserstein distance between μ and ν by*

$$W_p(\mu,\nu)^p := \min_{\gamma \in \mathcal{P}(X\times X)} \left\{ \int_{X\times X} d(x,y)^p d\gamma(x,y) \mid (\pi_1)_\sharp \gamma = \mu, (\pi_2)_\sharp \gamma = \nu \right\}.$$

These distances will be very important for our study especially for their geometrical properties. If for example we consider two Dirac measures δ_x, δ_y then $\|\delta_x - \delta_y\| = 2$ independently of the position of x and y, while one has $W_p(\delta_x, \delta_y) = d(x,y)$.

In order to check that definition 1.4.1 actually defines a distance we need some results of measure theory. First we give the notion of measurability of maps with values in the space of measures.

Definition 1.4.2. *A map $\lambda : X \to \mathcal{M}(Y)$ is said to be Borel if for any Borel set $B \subset Y$ the function $x \in X \mapsto \lambda(x)(B)$ is Borel. Equivalently, if for any Borel and bounded map $\varphi : X \times Y \to \mathbb{R}$ it results that the map*

$$x \in X \mapsto \int_Y \varphi(x,y) d\lambda(x)$$

is Borel.

The following theorem will be used several times in the sequel and a proof of it is in Theorem 2.28 of [2].

Theorem 1.4.3 (Disintegration theorem). *Let $f: X \to Y$ be a Borel map and $\mu \in \mathcal{M}^+(X)$. Then for $f_\sharp\mu -$ a.e. $y \in Y$ there exists $\mu_y \in \mathcal{P}(X)$ concentrated on $\{x \mid f(x) = y\}$ such that*
(i) the map $y \mapsto \mu_y$ is Borel,
(ii) $\forall \varphi \in \mathcal{C}_b(X): \int_X \varphi(x)d\mu = \int_Y \left(\int_X \varphi(x)d\mu_y(x)\right) df_\sharp\mu(y)$.
Moreover the measures μ_y are uniquely determined up to a negligible set with respect to $f_\sharp\mu$.

Remark 1.4.4. Denoting by $\sigma = f_\sharp\mu$, we will write $\mu = \sigma \otimes \mu_y$. Usually, we will use the above theorem in the particular case in which f is a projection of a product space on one of his factors. Moreover, writing $\sigma \otimes \mu_y$ we will always assume that μ_y satisfy the condition (i) and (ii) of Theorem 1.4.3.

The following lemma is useful to obtain the triangular inequality for W_p .

Lemma 1.4.5 (Dudley). *Let $\gamma_1 \in \mathcal{P}(X \times Y), \gamma_2 \in \mathcal{P}(Y \times Z)$ such that $(\pi_2)_\sharp\gamma_1 = (\pi_1)_\sharp\gamma_2 = \sigma$. Then there exists $\gamma \in \mathcal{P}(X \times Y \times Z)$ such that*

$$(\pi_{1,2})_\sharp\gamma = \gamma_1 \ , \ (\pi_{2,3})_\sharp\gamma = \gamma_2.$$

Proof. By Theorem 1.4.3 we find $\mu_y \in \mathcal{P}(X), \nu_y \in \mathcal{P}(Z)$ such that $\gamma_1 = \sigma \otimes \mu_y$ and $\gamma_2 = \sigma \otimes \nu_y$. Then it is sufficient to take $\gamma = \sigma \otimes (\mu_y \otimes \nu_y)$ to verify the assertion. □

It will be useful a countable version of Dudley's lemma.

Lemma 1.4.6. *Let $X_n,\ n \in \mathbb{N}$ a sequence of spaces and $\mu_n \in \mathcal{P}(X_n)$. Let $\gamma_{n,n+1}$ be transport plans between μ_n and μ_{n+1}. If we set $X = \Pi_{n\in\mathbb{N}}X_n$, then there exists $\tilde{\mu} \in \mathcal{P}(X)$ such that*

$$\forall n \in \mathbb{N}: (\pi_n, \pi_{n+1})_\sharp\tilde{\mu} = \gamma_{n,n+1},\ (\pi_n)_\sharp\tilde{\mu} = \mu_n.$$

Proof. By Dudley's lemma 1.4.5 we find $\tilde{\mu}_3 \in \mathcal{P}(X_1 \times X_2 \times X_3)$ such that

$$(\pi_{1,2})_\sharp\tilde{\mu}_3 = \gamma_{1,2} \ , \ (\pi_{2,3})_\sharp\tilde{\mu}_3 = \gamma_{2,3}.$$

For $n \geq 3$ let $\widetilde{X}_n := \Pi_{i=1}^n X_i$ and $\tilde{\pi}_n : \widetilde{X}_m \to \widetilde{X}_n, m \geq n$ be the projection onto the first n coordinates. Observe that by construction

$$(\pi_n)_\sharp\tilde{\mu}_n = (\pi_n)_\sharp(\pi_n, \pi_{n+1})_\sharp\tilde{\mu}_n = (\pi_n)_\sharp\gamma_{n-1,n} = \mu_n.$$

Applying recursively Dudley's lemma 1.4.5 to the triple $\widetilde{X}_{n-1}, X_n, X_{n+1}$, we can find a sequence of measures $\tilde{\mu}_{n+1} \in \mathcal{P}(\widetilde{X}_{n+1})$ satisfying:

$$(\tilde{\pi}_n)_\sharp\tilde{\mu}_{n+1} = \tilde{\mu}_n,\ (\pi_n, \pi_{n+1})_\sharp\tilde{\mu}_{n+1} = \gamma_{n,n+1}.$$

Kolmogorov's theorem, see cap. III no. 51 of [31], provides a measure $\tilde{\mu} \in \mathcal{P}(\widetilde{X})$ such that

$$\forall n \in \mathbb{N}: (\tilde{\pi}_n)_\sharp\tilde{\mu} = \tilde{\mu}_n.$$

Therefore

$$(\pi_n, \pi_{n+1})_\sharp\tilde{\mu} = (\pi_n, \pi_{n+1})_\sharp(\tilde{\pi}_{n+1})_\sharp\tilde{\mu} = (\pi_n, \pi_{n+1})_\sharp\tilde{\mu}_{n+1} = \gamma_{n,n+1}.$$

Furthermore

$$(\pi_n)_\sharp\tilde{\mu} = (\pi_n)_\sharp(\pi_n, \pi_{n+1})_\sharp\tilde{\mu} = (\pi_n)_\sharp\gamma_{n,n+1} = \mu_n.$$

We sketch that Kolmogorov's theorem can be proved by taking the functional $I(g) = \int_{\widetilde{X}} g_n\tilde{\mu}_n$ defined on functions of the form $g = g_n \circ \tilde{\pi}_n$ with $g_n \in \mathcal{C}(\widetilde{X}_n)$ and then applying the Daniell's representation theorem, which yields a measure $\tilde{\mu} \in \mathcal{P}(\widetilde{X})$ such that $I(g) = \int_{\widetilde{X}} g_n\tilde{\mu}$. See [31] for details. □

Theorem 1.4.7. *For every $p \geq 1$, W_p is a distance on $\mathcal{P}_p(X)$.*

Proof. $W_p(\mu,\nu) = W_p(\nu,\mu)$ by the symmetry of $d(x,y)$ and of course $W_p(\mu,\mu) = 0$. If $W_p(\mu,\nu) = 0$, then any optimal plan γ is concentrated on the diagonal of $X \times X$. Hence

$$\forall f \in \mathcal{C}_0(X) : \int_X f d\mu = \int_{X\times X} f(x) d\gamma(x,x) = \int_X f d\nu \Rightarrow \mu = \nu.$$

In order to check the triangular inequality fix $\mu_1, \mu_2, \mu_3 \in \mathcal{P}_p(X)$. Let $\gamma_{1,2}$ and $\gamma_{2,3}$ be optimal plans respectively between μ_1, μ_2, and μ_2, μ_3. Hence $(\pi_2)_\sharp \gamma_{1,2} = \mu_3 = (\pi_1)_\sharp \gamma_{2,3}$. By Lemma 1.4.5 applied to $X = Y = Z$ and $\gamma_{1,2}, \gamma_{2,3}$, we find $\gamma \in \mathcal{P}(X \times X \times X)$ such that $\gamma_{1,2} = (\pi_{1,2})_\sharp \gamma$, $\gamma_{2,3} = (\pi_{2,3})_\sharp \gamma$. Let us pone $\gamma_{1,3} = (\pi_{1,3})_\sharp \gamma$. Observe that

$$(\pi_1)_\sharp \gamma_{1,3} = (\pi_1)_\sharp (\pi_{1,3})_\sharp \gamma = (\pi_1)_\sharp (\pi_{1,2})_\sharp \gamma = (\pi_1)_\sharp \gamma_{1,2} = \mu_1,$$

$$(\pi_2)_\sharp \gamma_{1,3} = (\pi_2)_\sharp (\pi_{1,3})_\sharp \gamma = (\pi_2)_\sharp (\pi_{2,3})_\sharp \gamma = (\pi_2)_\sharp \gamma_{2,3} = \mu_3.$$

Then we can evaluate

$$\begin{aligned} W_p(\mu_1,\mu_3) &\leq \left(\int_{X\times X} d(x_1,x_3)^p d\gamma_{1,3} \right)^{\frac{1}{p}} = \left(\int_{X\times X\times X} d(x_1,x_3)^p d\gamma \right)^{\frac{1}{p}} \leq \\ &\leq \left(\int_{X\times X\times X} [d(x_1,x_2) + d(x_2,x_3)]^p d\gamma \right)^{\frac{1}{p}} \leq \\ &\leq \left(\int_{X\times X\times X} d(x_1,x_2)^p d\gamma \right)^{\frac{1}{p}} + \left(\int_{X\times X\times X} d(x_2,x_3)^p d\gamma \right)^{\frac{1}{p}} = \\ &= \left(\int_{X\times X} d(x_1,x_2)^p d(\pi_{1,2})_\sharp \gamma \right)^{\frac{1}{p}} + \left(\int_{X\times X} d(x_2,x_3)^p d(\pi_{2,3})_\sharp \gamma \right)^{\frac{1}{p}} = \\ &= \left(\int_{X\times X} d(x_1,x_2)^p d\gamma_{1,2} \right)^{\frac{1}{p}} + \left(\int_{X\times X} d(x_2,x_3)^p d\gamma_{2,3} \right)^{\frac{1}{p}} = \\ &= W_p(\mu_1,\mu_2) + W_p(\mu_2,\mu_3). \end{aligned}$$

□

The following proposition collects some important inequalities for the Wasserstein distances.

Proposition 1.4.8. *Let $1 \leq p \leq q$. Then*
1) $W_p \leq W_q$.
2) $W_q \leq W_p^{\frac{p}{q}} diam(X)^{1-\frac{p}{q}}$.

Proof. Point 1) is a direct application of Hölder's inequality. To check 2) let γ be an optimal plan between its marginal μ, ν for W_p. Hence

$$\begin{aligned} W_q &\leq \left(\int_{X\times X} d(x,y)^q d\gamma \right)^{\frac{1}{q}} = \left(\int_{X\times X} d(x,y)^p d(x,y)^{q-p} d\gamma \right)^{\frac{1}{q}} \leq \\ &\leq W_p^{\frac{p}{q}} diam(X)^{\frac{q-p}{q}} = W_p^{\frac{p}{q}} diam(X)^{1-\frac{p}{q}}. \end{aligned}$$

□

Therefore, on bounded metric spaces all the Wasserstein distances produce the same topology.

The following theorem characterizes the Wasserstein distances in terms of weak convergence of measures.

Theorem 1.4.9. *Let $1 \leq p < +\infty$ and $\mu_n,\ \mu \in \mathcal{P}_p(X)$. Then the following statements are equivalent:*

1. $\lim_{n\to+\infty} W_p(\mu_n, \mu) = 0$,

2. $\mu_n \overset{*}{\rightharpoonup} \mu$ *and for some (and thus any)* $x_0 \in X$*:*

$$\forall R > 0,\ \lim_{n\to+\infty} \int_{X\setminus B(x_0,R)} d(x_0,x)^p d\mu_n = \int_{X\setminus B(x_0,R)} d(x_0,x)^p d\mu,$$

3. $\forall \varphi \in \mathcal{C}(X)$ *s.t.* $|\varphi(x)| \leq C[1 + d(x_0,x)^p]$, $\lim_{n\to+\infty} \int_X \varphi d\mu_n = \int_X \varphi d\mu$.

Proof. 1) $\Rightarrow$ 2). We make use of the following elementary inequality:

$$\forall \varepsilon > 0, \exists C_\varepsilon \text{ s.t. } \forall a, b \geq 0, (a+b)^p \leq (1+\varepsilon)a^p + C_\varepsilon b^p. \tag{1.32}$$

(It suffices to establish the inequality for $b = 1$ and then take $C_\varepsilon = \max\{(a+1)^p - (1+\varepsilon)a^p \mid a \geq 0\}$). By the triangular inequality we have

$$d(x_0,x)^p \leq (1+\varepsilon)d(x_0,y)^p + C_\varepsilon d(x,y)^p.$$

If γ_n is an optimal plan between μ_n and μ, integrating the above inequality we obtain

$$\int_X d(x_0,x)^p d\mu_n(x) \leq (1+\varepsilon)\int_X d(x_0,y)^p d\mu(y) + C_\varepsilon \int_{X\times X} d(x,y)^p d\gamma_n(x,y).$$

Passing to the limit as $n \to +\infty$ we have

$$\limsup_{n\to+\infty} \int_X d(x_0,x)^p d\mu_n \leq (1+\varepsilon)\int_X d(x_0,y)^p d\mu(y).$$

Letting $\varepsilon \to 0^+$ we obtain

$$\limsup_{n\to+\infty} \int_X d(x_0,x)^p d\mu_n \leq \int_X d(x_0,x)^p d\mu. \tag{1.33}$$

To complete the proof let us first suppose that X is compact. Hence, by Proposition 1.4.8 we can restrict our attention to the case $p = 1$ without loss of generality. Therefore we can use the duality with the 1-Lipschitz functions. Indeed, if $f \in \mathrm{Lip}(X)$, by Theorem 1.2.4 we have

$$|\int_X f d(\mu_n - \mu)| \leq \mathrm{Lip}(f) W_1(\mu_n, \mu).$$

Since Lipschitz functions are dense in $\mathcal{C}(X)$ with respect to the uniform convergence, we deduce that $\mu_n \overset{*}{\rightharpoonup} \mu$. If X is not compact, fix $f \in \mathcal{C}_c(X)$ and large balls B_h containing the support of f. Let γ_n be an optimal plan between μ_n and μ for W_1. Observe that by Lemma 1.1.8 the transport plans $\gamma_{n\lfloor B_h\times B_h}$ are optimal between their marginals, say $\gamma_n^{h^+}$ and $\gamma_n^{h^-}$. Since, by Proposition 1.4.8, $W_1 \leq W_p$, we have

$$W_1(\gamma_n^{h^+}, \gamma_n^{h^-}) = \int_{B_h\times B_h} d(x,y) d\gamma(x,,y) \leq W_1(\mu_n,\mu) \leq W_p(\mu_n,\mu) \overset{n\to+\infty}{\longrightarrow} 0,$$

hence $(\gamma_n^{h^+} - \gamma_n^{h^-}) \overset{*}{\rightharpoonup} 0$. As in the proof of Theorem 1.1.7, taking large balls and using (1.33), by Lemma 1.1.6 we can assume that $\gamma_n \rightharpoonup \gamma$. We can evaluate

$$|\int_X f d(\mu_n - \mu)| = |\int_{B_h} f d(\mu_n - \mu)| \leq$$

$$\leq |\int_{B_h} fd(\mu_n - \gamma_n^{h^+})| + |\int_{B_h} fd(\gamma_n^{h^+} - \gamma_n^{h^-})| + |\int_{B_h} fd(\gamma_n^{h^-} - \mu)| \leq$$

$$\leq \|f\|_\infty [|\mu_n - \gamma_n^{h^+}|(B_h) + |\mu - \gamma_n^{h^-}|(B_h)] + |\int_{B_h} fd(\gamma_n^{h^+} - \gamma_n^{h^-})| \leq$$

$$\leq \|f\|_\infty [\gamma_n(B_h, X) - \gamma_n(B_h, B_h) + \gamma_n(X, B_h) - \gamma_n(B_h, B_h)] + |\int_{B_h} fd(\gamma_n^{h^+} - \gamma_n^{h^-})|.$$

Passing to the limit as $n \to +\infty$ we obtain

$$\limsup_{n\to+\infty} |\int_X fd(\mu_n - \mu)| \leq \|f\|_\infty [\gamma(B_h', X) - \gamma(B_h, B_h) + \gamma(X, B_h') - \gamma(B_h, B_h)].$$

Letting $h \to +\infty$ we obtain that $\mu_n \overset{*}{\rightharpoonup} \mu$. By lower semicontinuity of the weak$*$ convergence we also have

$$\int_X d(x_0, x)^p d\mu \leq \liminf_{n\to+\infty} \int_X d(x_0, x)^p d\mu_n,$$

which together with (1.33) complete the proof (since the same arguments also works for $X \setminus B(x_0, R)$ in place of X).

2) $\Rightarrow$ 1). Observe that by (1.32) we have

$$W_p(\mu_n, \mu)^p \leq \int_{X\times X} d(x, y)^p d\gamma_n \leq (1 + \varepsilon) \int_X d(x_0, x)^p d\mu_n + C_\varepsilon \int_X d(x_0, y)^p d\mu.$$

Since $\mu \in \mathcal{P}_p(X)$ and by 2) we can find a large ball such that

$$\int_{X\setminus B(x_0,R)} d(x_0, x)^p d\mu < \varepsilon \ , \ |\int_{X\setminus B(x_0,R)} d(x_0, x)^p d(\mu_n - \mu)| < \varepsilon.$$

Hence the contribution outside the ball $B(x_0, R)$ is negligible in $W_p(\mu_n, \mu)$. Then we can assume again that X is compact and then we have to prove that $W_1(\mu_n, \mu) \to 0$ as $n \to +\infty$. Let $u_n \in \mathrm{Lip}_1(X)$ be Kantorovich potentials for $W_1(\mu_n, \mu)$. Adding a constant we can assume that $u_n(x_0) = 0$. By the Ascoli-Arzelà theorem we can assume that u_n converges uniformly to a function $u \in \mathrm{Lip}_1(X)$. Hence

$$W_1(\mu_n, \mu) = \int_X u_n d(\mu_n - \mu) \leq |\int_X (u_n - u) d(\mu_n - \mu)| + |\int_X u d(\mu_n - \mu)| \leq$$

$$\leq 2\|u_n - u\|_\infty + |\int_X u d(\mu_n - \mu)| \overset{n\to+\infty}{\longrightarrow} 0.$$

2) $\Rightarrow$ 3). Take $\varepsilon > 0$ and φ as in point 3). By 2) we can find $R > 0$ such that for large n it results

$$\int_{X\setminus B(x_0,R)} d(x_0, x)^p d\mu < \varepsilon \ , \ |\int_{X\setminus B(x_0,R)} d(x_0, x)^p d(\mu_n - \mu)| < \varepsilon,$$

$$|\int_{B(x_0,R)} \varphi d(\mu_n - \mu)| < \varepsilon \ , \ \mu_n(X \setminus B(x_0, R)) < \varepsilon \ , \ \mu(X \setminus B(x_0, R)) < \varepsilon.$$

Then we can evaluate

$$|\int_X \varphi d(\mu_n - \mu)| \leq |\int_{B(x_0,R)} \varphi d(\mu_n - \mu)| + |\int_{X\setminus B(x_0,R)} \varphi d(\mu_n - \mu)| <$$

$$< \varepsilon + C\mu_n(X \setminus B(x_0, R)) + C\mu(X \setminus B(x_0, R)) +$$

$$+C\int_{X\setminus B(x_0,R)} d(x_0,x)^p d\mu_n + C\int_{X\setminus B(x_0,R)} d(x_0,x)^p d\mu <$$

$$< \varepsilon + 2C\varepsilon + 2C\varepsilon + C\varepsilon.$$

3) $\Rightarrow$ 2). It suffices to estimate

$$|\int_{X\setminus B(x_0,R)} d(x_0,x)^p d(\mu_n-\mu)| \leq$$

$$\leq |\int_X d(x_0,x)^p d(\mu_n-\mu)| + |\int_{B(x_0,R)} d(x_0,x)^p d(\mu_n-\mu)|.$$

□

Remark 1.4.10. If X is compact then Theorem 1.4.9 implies that $\mathcal{P}_p(X)$ is also compact. However in the general case $\mathcal{P}_p$ is not even locally compact.

Proposition 1.4.11. *For every $p \geq 1$, $(\mathcal{P}_p(\mathbb{R}^n), W_p)$ is not locally compact.*

Proof. Fix $r > 0$ and a sequence of points x_n such that $\frac{1}{n}|x_n|^p = r^p$. Set $\mu_n = (1-\frac{1}{n})\delta_0 + \frac{1}{n}\delta_{x_n}$. Observe that $\mu_n \overset{*}{\rightharpoonup} \delta_0$, while $W_p(\mu_n,\delta_0) = (\frac{1}{n}|x_n|^p)^{\frac{1}{p}} = r$. On the other hand we have that $\int_{\mathbb{R}^n} |x|^p d\mu_n = \frac{1}{n}|x_n|^p = r^p$. Hence, by Theorem 1.4.9 there are no convergent subsequences. □

The following convexity property is an important feature of the Wasserstein distances.

Proposition 1.4.12. *Let $p \geq 1$ and $\mu_1,\mu_2,\nu_1,\nu_2 \in \mathcal{P}_p(X)$. Then $\forall\lambda \in [0,1]$:*

$$W_p\left(\lambda\mu_1 + (1-\lambda)\mu_2, \lambda\nu_1 + (1-\lambda)\nu_2\right)^p \leq \lambda W_p(\mu_1,\nu_1)^p + (1-\lambda)W_p(\mu_2,\nu_2)^p.$$

Proof. Let γ_i be an optimal plan between μ_i and ν_i for $i = 1, 2$. Observe that $\lambda\gamma_1 + (1-\lambda)\gamma_2$ is a transport plan between $\lambda\mu_1 + (1-\lambda)\mu_2$ and $\lambda\nu_1 + (1-\lambda)\nu_2$. Then

$$W_p\left(\lambda\mu_1 + (1-\lambda)\mu_2, \lambda\nu_1 + (1-\lambda)\nu_2\right)^p \leq \lambda W_p(\mu_1,\nu_1)^p + (1-\lambda)W_p(\mu_2,\nu_2)^p.$$

□

Convexity properties of the distance are related to important geometric notions, in particular to curvature properties. In the last chapter we will derive some other geometric properties of $\mathcal{P}_p(X)$ in the case of a Riemannian manifold X.

1.5 Brenier's formulation of the Monge-Kantorovich problem

In this section we want to describe a PDE formulation of the Monge-Kantorovich problem introduced by Brenier (see [8, 17, 19]) which will be important for our purposes. In this setting we consider a compact Riemannian manifold (M, g) without boundary as ambient space. We denote by d the geodesic distance on M.

The basic idea is to look at paths connecting the two measures $f^+, f^- \in \mathcal{P}(M)$ that we want to transport. If, in the simplest case of $f^+ = \delta_x$ and $f^- = \delta_y$, we take γ a regular curve connecting x to y, then the measure $\rho_t = \delta_{\gamma(t)}$ and the velocity field $E_t = \dot\gamma(t)\delta_{\gamma(t)}$ are related by the equation

$$\frac{\partial\rho}{\partial t} + \nabla\cdot E = 0 \quad \text{on } (0,1)\times M, \tag{1.34}$$

in the sense of distribution. Indeed, if we take a test function $\varphi(t,x) \in \mathcal{C}_c^\infty((0,1)\times M)$, we have

$$(\frac{\partial \rho}{\partial t} + \nabla \cdot E)(\varphi) =$$

$$= -\int_0^1 \int_M \partial_t \varphi(t,x) d\rho_t dt - \int_0^1 \int_M d_x\varphi(t,x) dE_t dt =$$

$$= -\int_0^1 \left[\partial_t \varphi(t,\gamma(t)) + \langle d_x\varphi(\gamma(t)), \dot\gamma(t)\rangle\right] dt = -\int_0^1 \frac{d}{dt}[\varphi(t,\gamma(t))] dt = 0.$$

More generally, we can consider couples $(\rho_t, E_t) \in \mathcal{P}(M) \times \mathcal{M}^n(\mathcal{M})$ which satisfy the equation (1.34) as admissible configurations. Observe that if ψ is a regular function on M, then (1.34) implies

$$\frac{d}{dt}\left(\int_M \psi(x) d\rho_t\right) = \int_M d_x\psi(x) dE_t \tag{1.35}$$

in the sense of distributions. Indeed, because of (1.34) the right side of (1.35) is in fact a distribution. Then, if $\varphi \in \mathcal{C}_c^\infty(0,1)$, we have

$$\frac{d}{dt}\left(\int_M \psi(x) d\rho_t\right)\varphi = -\int_0^1 \dot\varphi(t) \int_M \psi d\rho_t dt =$$

$$-\int_0^1 \int_M \dot\varphi(t)\psi d\rho_t dt = \frac{\partial \rho}{\partial t}(\varphi\psi) = -\nabla \cdot E(\varphi\psi) =$$

$$= \int_0^1 \varphi(t) \int_M d_x\psi(x) dE_t dt = \left(\int_M d_x\psi(x) dE_t\right)\varphi.$$

We remark that if $\int_0^1 \|E_t\| dt < +\infty$, then for every $\varphi \in \mathcal{C}^\infty(M)$ the map $t \mapsto \int_0^1 \varphi d\rho_t \in W^{1,1}_{loc}(0,1)$. Actually, Proposition 1.37 below allows to consider ρ_t as an absolutely continuous curve in the space $\mathcal{P}(M)$ endowed with the 1-Wasserstein distance. For an analysis in metric spaces we refer to [3, 6]. We will need the following standard approximation lemma:

Lemma 1.5.1. $\mathrm{Lip}_1(M) = \overline{\{u \in \mathcal{C}^1(M) : |du(x)| \le 1\}}^{\|\cdot\|_\infty}$.

For Lipschitz function on a metric space (M,d) we introduce the metric derivative according to the following definition.

Definition 1.5.2. *Given a curve $\rho : [a,b] \to (M,d)$ we define the metric derivative at the point $t \in]a,b[$ as the limit*

$$\lim_{h\to 0} \frac{d(\rho(t+h),\rho(t))}{h} \tag{1.36}$$

whenever it exists and in this case we denote it by $|\dot\rho|(t)$.

If $\rho : [a,b] \to (M,d)$ is a Lipschitz curve, by Rademacher Theorem the metric derivative of ρ exists at $\mathcal{L}^1$-a.e. point in $[a,b]$. Furthermore, the length of the Lipschitz curve ρ is given by

$$l(\rho) = \int_a^b |\dot\rho|(t) dt.$$

Theorem 1.5.3. *Let $\rho : [0,1] \to \mathcal{P}(M)$ be a solution of the equation (1.34) with $E_t \in \mathcal{M}^n(M)$ such that $\int_0^1 \|E_t\| dt < +\infty$. Then the curve ρ is absolutely continuous (a.c.) with respect to the 1-Wasserstein distance W_1. Furthermore, the metric derivative of ρ satisfies*

$$|\dot\rho|(t) \le \|E_t\| \quad \text{for } \mathcal{L}^1 - a.e.\ t \in [0,1]. \tag{1.37}$$

Conversely, for every a.c. curve ρ there exists E_t such that (ρ, E) satisfies (1.34) and equality holds in (1.37).

Proof. Let $t_1, t_2 \in [0,1]$ and $u \in \mathrm{Lip}_1(M)$ be a Kantorovich potential for ρ_{t_1} and ρ_{t_2}. By approximation, we can suppose that $u \in C^1(M)$. Hence

$$W_1(\rho_{t_1}, \rho_{t_2}) = \int_M ud(\rho_{t_1} - \rho_{t_2}) = \int_{t_1}^{t_2} \frac{d}{dt}[\int_M ud\rho_t]dt =$$
$$= \int_{t_1}^{t_2} \int_M d_x u dE_t dt \leq \int_{t_1}^{t_2} \|E_t\| dt.$$

Since $\int_0^1 \|E_t\| dt < +\infty$, we deduce the absolute continuity of ρ and also (1.37). Indeed we have

$$\lim_{h \to 0} \frac{W_1(\rho_{t+h}, \rho_t)}{h} \leq$$
$$\leq \lim_{h \to 0} \frac{1}{h} \int_t^{t+h} \|E_s\| ds = \|E_t\| \quad \text{for } \mathcal{L}^1 - a.e.\ t \in [0,1].$$

Vice versa, without loss of generality, let us suppose that the curve ρ is Lipschitz. Consider the space F of $C^1(M)$ functions such that $u(x_0) = 0$ at a fixed point $x_0 \in M$ endowed with the norm $\|\phi\| = \mathrm{Lip}(\phi)$. Hence we have that $\mathcal{P}(M) \subset F^*$. Observe that if $\mu, \nu \in \mathcal{P}(M)$, and $\|\phi\| \leq 1$ we have

$$|\langle \mu - \nu, \phi \rangle| = |\int_M \phi d(\mu - \nu)| \leq W_1(\mu, \nu).$$

On the other hand, again by approximation, if $u \in \mathrm{Lip}_1(M)$ it results:

$$W_1(\mu, \nu) = \int_M ud(\mu - \nu) \leq \|\mu - \nu\|_{F^*}.$$

Hence $W_1(\mu, \nu) = \|\mu - \nu\|_{F^*}$, so that $(\mathcal{P}(M), W_1)$ can be isometrically embedded in F^*. By a general result proved in [5], any Lipschitz function from $[0,1]$ into the dual of a separable Banach space is weakly$*$-differentiable for $\mathcal{L}^1 - a.e.\ t$. In addition, if we set $\mu_t \in F^*$ such weak$*$-differential, it results $|\dot{\rho}|(t) = \|\mu_t\|_{F^*}$. Using the Hahn-Banach theorem one can check that any $\mu \in F^*$ can be represented as the divergence of a vectorial measure $E \in \mathcal{M}^n(M)$ with $\|\mu\|_{F^*} = \|E\|$. If we denote by $\mu_t = divE_t$, it remains to check that the couple $(\rho_t, -E_t)$ is an admissible configuration. Let $\phi \in F$. Since ρ is Lipschitz, then the function $t \mapsto \int_M \phi d\rho_t$ is Lipschitz too. Then, by the weak$*$-differentiability, for $\mathcal{L}^1 - a.e.\ t$ we have

$$\frac{d}{dt} \int_M \phi d\rho_t = -\langle \nabla \cdot E_t, \phi \rangle = \int_M d_x \phi dE_t. \tag{1.38}$$

If $\varphi \in C_c^\infty((0,1) \times M)$, we can evaluate

$$0 = \int_0^1 \frac{d}{dt} \int_M \varphi d\rho_t dt = \int_0^1 \int_M \partial_t \varphi(t,x) d\rho_t dt + \int_0^1 \int_M d_x \varphi(t,x) dE_t dt. \tag{1.39}$$

The last equation in (1.39) is obtained writing

$$\frac{1}{h}(\int_M \varphi(t+h, x) d\rho_{t+h} - \int_M \varphi(t,x) d\rho_t) =$$
$$= \frac{1}{h}\left[\int_M (\varphi(t+h,x) - \varphi(t,x)) d\rho_t + \int_M \varphi(t+h,x) d(\rho_{t+h} - \rho_t)\right],$$

and observing that the first term of the above formula goes to $\int_M \partial_t \varphi(t,x) d\rho_t$ as $h \to +\infty$, while the second can be divided in two different terms:

$$\frac{1}{h} \int_M \varphi(t,x) d(\rho_{t+h} - \rho_t),$$

from which we obtain $\int_M d_x\varphi(t,x)dE_t$ because of (1.38), and

$$\frac{1}{h}\int_M [\varphi(t+h,x)-\varphi(t,x)]d(\rho_{t+h}-\rho_t),$$

which is negligible since

$$\frac{1}{h}\left|\int_M [\varphi(t+h,x)-\varphi(t,x)]d(\rho_{t+h}-\rho_t)\right| \le Lip(\varphi)W_1(\rho_{t+h},\rho_t) \overset{h\to 0^+}{\to} 0.$$

□

If $\int_0^1 \|E_t\|dt < +\infty$, by Theorem 1.5.3 ρ is an a.c. curve between its endpoints. Furthermore, as follows by the proof, the equation (1.35) holds pointwise for $\mathcal{L}^1 - a.e.\ t$ and it turns out to be equivalent to (1.34).

The equation (1.34) has been used in the Monge-Kantorovich theory since its beginning for many applications. Recently, in [3] it has been pointed out that this equation also characterizes the a.c. curves on the space $(\mathcal{P}_p(M),W_p)$, as shown in the following theorem proved in [3].

Theorem 1.5.4. *Let $\rho : [0,1] \to (\mathcal{P}(M),W_p)$ be a curve. If ρ is a.c. and $|\dot\rho| \in L^1(0,1)$ is its metric derivative, then there exists a Borel vector field $v : (t,x) \mapsto v_t(x)$ such that*

$$v_t \in L^p(\rho_t) \quad \textit{and} \quad \|v_t\|_{L^p(\rho_t)} \le |\dot\rho|(t) \quad \textit{for} \quad \mathcal{L}^1 - a.e.\ t\in[0,1] \tag{1.40}$$

and the continuity equation

$$\frac{\partial\rho}{\partial t} + \nabla\cdot v\rho = 0 \quad \textit{in} \quad (0,1)\times M, \tag{1.41}$$

is satisfied in the sense of distributions.

Conversely, if ρ satisfies equation (1.41) for some vector fields v_t such that $\|v_t\|_{L^p(\rho_t)} \in L^1(0,1)$, then the map $t\mapsto\rho_t$ is a.c. and

$$|\dot\rho|(t) \le \|v_t\|_{L^p(\rho_t)} \quad \textit{for} \quad \mathcal{L}^1 - a.e.\ t\in[0,1].$$

Therefore, given an a.c. curve $\rho : [0,1] \to (\mathcal{P}_p(M),W_p)$ we can find a Borel vector field v such that $|\dot\rho|(t) = \|v_t\|_{L^p(\rho_t)}$ for $\mathcal{L}^1 - a.e.\ t\in[0,1]$. We refer to this vector field as the *tangent* vector associated to ρ. For the precise meaning of the word *tangent* we refer to section 7.6 of [3].

The equation (1.34) is often called transport equation or continuity equation. We will return on this continuity equation several times in the sequel.

Now we want to associate to the admissible configurations a cost of transportation. If we look again to the example of the two Dirac measures δ_x,δ_y of the beginning of this section, it appears clear that for any curve γ connecting the points x to y the cost associated to the transportation has to be equal to the length of γ. So if $\rho_t = \delta_{\gamma(t)}$ and $E_t = \dot\gamma(t)\delta_{\gamma(t)}$ is the admissible configuration we have

$$\int_0^1 |\dot\gamma(t)|dt = \int_0^1 \|E_t\|dt.$$

By analogy we consider the quantity $\int_0^1 \|E_t\|dt$ as the cost associated to the configuration (ρ,E). Then the Brenier formulation of the Monge-Kantorovich problem with the cost $c(x,y)=d(x,y)$ amounts to minimize the functional

$$K(\rho,E) := \int_0^1 \|E_t\|dt$$

among maps $\rho : [0,1]\to\mathcal{P}(M)$ and $E : [0,1] \to \mathcal{M}^n(M)$ which solve

$$\begin{cases} \dfrac{\partial\rho}{\partial t} + \nabla\cdot E = 0 & \textit{on } (0,1)\times M, \\ \rho_0 = f^+, \quad \rho_1 = f^-. \end{cases} \tag{1.42}$$

We set

$$\mathbf{B}_{1/2} := \min\{K(\rho, E) \mid (\rho, E) \text{ solves } (1.42)\} \tag{1.43}$$

The equivalence between the Brenier problem and **MK** is a well known fact. Moreover, Brenier formulation make sense whenever $c(x,y)$ is a "length" cost, namely

$$c(x,y) = \inf_{\substack{\gamma \in AC([0,1],M) \\ \gamma(0)=x, \gamma(1)=y}} \int_0^1 L(\gamma(t), \dot\gamma(t))dt,$$

for a given Lagrangian $L : TM \to \mathbb{R}$. Actually, the Brenier formulation was originally introduced for the quadratic Lagrangian $|v|^2$ (see for example [8]). Furthermore, this equivalence can be easily generalized to Lagrangian of the type $\varphi(|v|)$ (see [1, 19]). In chapter 3, see Corollary 4.0.10, we will prove the equivalence between Brenier problem and **MK** for a general Lagrangian $L(x,v)$. The Brenier problem formulated in this section corresponds to the Lagrangian $L(x,v) = g(v,v)^{\frac{1}{2}}$. To point out the dependence on the Lagrangian we denote the Monge-Kantorovich problem by $\mathbf{MK}_{1/2}$. To recover the equivalence between these two problems let us introduce a third problem, formulated on the space $\mathcal{N}_1(M)$ of normal 1-currents (for the basic facts about currents see the appendix A) as follows:

$$\mathbf{M}_{1/2} := \min\left\{M(T) \mid T \in \mathcal{N}_1(M), \partial T = f^+ - f^-\right\}. \tag{1.44}$$

We have the following

Theorem 1.5.5. $\boldsymbol{M}_{1/2} = \boldsymbol{B}_{1/2} = \boldsymbol{MK}_{1/2}$.

Proof. Take γ an optimal plan for $\mathbf{MK}_{1/2}$ and η a geodesic connecting x to y. Define (ρ, E) as follows:

$$\rho_t = \int_{M\times M} \delta_{x-\eta(t)} d\gamma(x,y) \ , \ E_t = \int_{M\times M} \dot\eta(t)\delta_{x-\eta(t)} d\gamma(x,y). \tag{1.45}$$

We claim that (ρ, E) is admissible for $\mathbf{B}_{1/2}$. Indeed, if $\varphi(t,x)$ is a test function, we have

$$\left[\frac{\partial\rho}{\partial t} + \nabla\cdot E\right](\varphi) = -\int_{M\times M}\int_0^1 \left[\partial_t\varphi(t,\eta(t)) + \langle d_x\varphi(t,\eta(t)), \dot\eta(t)\rangle\right] dt d\gamma(x,y) =$$

$$= -\int_{M\times M}\int_0^1 \frac{d}{dt}[\varphi(t,\eta(t))]dt d\gamma(x,y) = 0.$$

Furthermore: $\langle\rho_0, \varphi\rangle = \int_{M\times M}\varphi(x)d\gamma(x,y) = \langle f^+, \varphi\rangle$. Hence $\rho_0 = f^+$ and analogously $\rho_1 = f^-$.

Hence

$$\mathbf{B}_{1/2} \le \int_0^1 \|E_t\| dt \le \int_{M\times M}\int_0^1 |\dot\eta(t)| dt d\gamma(x,y) =$$

$$= \int_{M\times M} d(x,y)d\gamma(x,y) = \mathbf{MK}_{1/2}.$$

Now consider (ρ, E) admissible for $\mathbf{B}_{1/2}$ and put

$$T(\omega) = \int_0^1\int_M \omega dE_t dt. \tag{1.46}$$

We have

$$\partial T(\varphi) = T(d\varphi) = \int_0^1\int_M d\varphi(x)dE_t dt = \int_0^1 \frac{d}{dt}\int_M \varphi(x)d\rho_t dt =$$

$$= \langle\rho_1 - \rho_0, \varphi\rangle = \langle f^+ - f^-, \varphi\rangle.$$

Hence $\partial T = f^+ - f^-$. We evaluate

$$|T(\omega)| \leq \int_0^1 \|E_t\| dt \|\omega\|_\infty \Rightarrow M(T) \leq \mathbf{B}_{1/2},$$

from which we obtain $\mathbf{M}_{1/2} \leq \mathbf{B}_{1/2}$.

To check the opposite inequalities, let us consider a Kantorovich potential $u \in Lip_1(M)$.

Taking a current T such that $\partial T = f^+ - f^-$ we have

$$\mathbf{MK}_{1/2} = \int_M ud(f^+ - f^-) = \partial T(u) = T(du) \leq M(T).$$

Taking the minimum with respect to T the statement follows. □

Remark 1.5.6. Observe that the couple (ρ, E) defined in (4.6) is optimal for $\mathbf{B}_{1/2}$ while the correspondent current T defined by (1.46) is optimal for $\mathbf{M}_{1/2}$.

Remark 1.5.7. The equivalence between the three different problems stated in Theorem 1.5.5 is already known since in this case $\mathbf{M}_{1/2}$ can be seen as a shape optimization problem (see [11], [12]). Indeed, in terms of vectorial measures, the value $\mathbf{M}_{1/2}$ amounts to

$$\min\{M(T) \mid T \in \mathcal{N}_1(M), \partial T = f^+ - f^-\} = \min\left\{\int_M d|\nu| : \ \nu \in \mathcal{M}^n(M), -div(\nu) = f^+ - f^-\right\}.$$

For this different setting, and of course with a different proof, one can look to [71], where this question is formulated in terms of finslerian metric and a careful comparison of these problems on Riemannian manifolds is studied. Another proof is available in [54].

The introduction of an auxiliary problem on the space of normal currents will be crucial to study some relationship between the Monge-Kantorovich problem and the Mather's minimal measures of Lagrangian dynamic as we will see in the next chapters.

Chapter 2

Minimal measures and weak KAM theory

In this chapter we present some aspects of the theory of minimal measures in Lagrangian dynamic, first introduced by Mather, and relations to the weak KAM theory recently developed by Fathi. Some references for this subjects are the recent book and lecture notes [26, 43] and the papers [7, 40, 44, 62, 64, 65]. We will use some tools of these theories to study Brenier's formulation of the Monge-Kantorovich problem in a general Lagrangian setting (see Corollary 4.0.10). A useful "Hamiltonian" duality formula for the mass transportation will be also derived (Theorem 4.0.9).

Throughout this and the following chapters, unless otherwise stated, we will consider a compact Riemannian manifold (M, g) without boundary and a Lagrangian $L : TM \to \mathbb{R}$ which satisfies the following properties:

1. L is regular (say $\mathcal{C}^2$) and positive,
2. $L(x, \cdot)$ is superlinear in the fiber for all x, i.e. $\lim_{\|v\|\to+\infty} \frac{L(x,v)}{\|v\|} = +\infty$.
3. $L(x, \cdot)$ is strictly convex in the fiber for all x, i.e. $L(x, tv+(1-t)w) \leq tL(x, v)+(1-t)L(x, w)$ for all v, w and $t \in (0, 1)$, equality holding if and only if $v = w$.

The Hamiltonian, denoted by $H : T^*M \to \mathbb{R}$, is the convex conjugate of L, i.e.

$$H(x, w) = \sup_{v\in T_xM} \{\langle w, v\rangle - L(x, v)\}. \tag{2.1}$$

It turns out that H is regular, superlinear and strictly convex as well. Under this assumption it also results that the Legendre transform $\mathcal{L} : TM \to T^*M$, defined by

$$\mathcal{L}(x, v) = (x, \frac{\partial}{\partial v}L(x, v)),$$

is a $\mathcal{C}^1$-diffeomorfism. By definition we have the Young inequality

$$L(x, v) + H(x, \omega) \geq \langle \omega, v\rangle \qquad \forall x \in M, \forall v \in T_xM, \forall \omega \in T_x^*M. \tag{2.2}$$

Moreover, we have equality in (2.2) just on the Legendre transform

$$H(x, \frac{\partial}{\partial v}L(x, v)) = \langle \frac{\partial}{\partial v}L(x, v), v\rangle - L(x, v). \tag{2.3}$$

For an introduction to the Legendre transform see for example [43].

2.1 Preliminaries from the Calculus of Variations

A classical problem in the Calculus of Variation is to minimize the action of curves with fixed endpoints. So, fixed $x, y \in M$, we have the problem

$$\min \left\{ \mathcal{A}(\gamma) := \int_a^b L(\gamma, \dot{\gamma}) dt \mid \gamma \in \mathcal{C}^1([a,b], M), \gamma(a) = x, \gamma(b) = y \right\}. \tag{2.4}$$

For the Calculus of Variation in one dimension we refer to [22, 26, 43, 57]. A necessary condition to problem (2.4) is the vanishing of the so called first variation. Critical points of the minimization problem (2.4) which are $\mathcal{C}^1$ curves are also called extremal. If γ is $\mathcal{C}^2$ then γ is extremal if and only if γ solves the Euler-Lagrange equation:

$$\frac{d}{dt}\frac{\partial}{\partial v} L(\gamma, \dot{\gamma}) - \frac{\partial}{\partial x} L(\gamma, \dot{\gamma}) = 0. \tag{2.5}$$

More precisely, (2.5) is a second order system of ODE. The convexity assumption on the Lagrangian L ensures that (2.5) can be written as a system of normal equations of first order (Hamilton's ODE). In fact, the canonical Hamilton's ODE amounts to

$$\begin{cases} \dot{\gamma}(t) = \dfrac{\partial}{\partial \omega} H(\gamma(t), \omega(t)) \\ \\ \dot{\omega}(t) = -\dfrac{\partial}{\partial x} H(\gamma(t), \omega(t)), \end{cases} \tag{2.6}$$

where $\omega(t) = \frac{\partial}{\partial v} L(\gamma(t), \dot{\gamma}(t))$. Hence, by ODE theory, fixed $(x, v) \in TM$ as initial condition, (2.5) admits a unique solution. This fact leads to the following

Definition 2.1.1 (Lagrangian flow). *Let $t \in \mathbb{R}$. The Lagrangian flow is the function*

$$\phi_t : TM \to TM$$

defined as follows:

$$\phi_t(x, v) = (\gamma(t), \dot{\gamma}(t)),$$

where γ solves (2.5) with the initial conditions $\gamma(0) = x, \dot{\gamma}(0) = v$.

Remark 2.1.2. If $\gamma \in \mathcal{C}^2([a,b], M)$, by definition it turns out that γ is extremal if and only if

$$\forall s \in [a,b] : \phi_{s-a}(\gamma(a), \dot{\gamma}(a)) = (\gamma(s), \dot{\gamma}(s)).$$

Since M is compact, our hypothesis on the Lagrangian are sufficient to ensure that the flow is defined for all times as stated in Lemma 2.1.3 below. In this case the flow is said to be complete. In order to check this fact it will be important to consider the energy function $E(x, v)$ defined by

$$E(x, v) = \langle \frac{\partial}{\partial v} L(x, v), v \rangle - L(x, v).$$

This energy is constant along the extremal curves as comes out from the proof of the following

Lemma 2.1.3. *The flow ϕ_t is complete.*

Proof. If $\gamma \in \mathcal{C}^2$ is an extremal then

$$\frac{d}{dt} E(\gamma, \dot{\gamma}) = \frac{d}{dt} [\langle \frac{\partial}{\partial v} L(\gamma, \dot{\gamma}), \dot{\gamma} \rangle - L(\gamma, \dot{\gamma})] =$$

$$= \langle \frac{d}{dt}\frac{\partial}{\partial v}L(\gamma,\dot\gamma),\dot\gamma\rangle + \langle \frac{\partial}{\partial v}L(\gamma,\dot\gamma),\ddot\gamma\rangle - \langle \frac{\partial}{\partial v}L(\gamma,\dot\gamma),\ddot\gamma\rangle - \langle \frac{\partial}{\partial x}L(\gamma,\dot\gamma),\dot\gamma\rangle = 0.$$

Hence the energy is constant along the extremals. Observe that by convexity we have

$$\langle \frac{\partial}{\partial v}L(x,v),v\rangle \geq L(x,v) - L(x,0).$$

Then

$$\frac{1}{\|v\|}\langle \frac{\partial}{\partial v}L(x,v),v\rangle \geq \frac{L(x,v)}{\|v\|} - \frac{L(x,0)}{\|v\|}.$$

By superlinearity we infer that

$$\lim_{\|v\|\to+\infty} \|\frac{\partial}{\partial v}L(x,v)\| = +\infty. \tag{2.7}$$

Fix $(x,v) \in TM$, since the energy is constant along extremals we have

$$\forall t \in (a,b) : E(\phi_t(x,v)) = c.$$

By (2.7) and (2.3) we have that there exists a compact set $K \subset TM$ such that

$$\forall t \in (a,b) : (\gamma(t),\dot\gamma(t)) \in K.$$

Hence, by ODE theory, we obtain global existence in time for the flow $\phi_t(x,v)$. □

Sometimes we consider the flow as a function

$$\phi : \mathbb{R} \times TM \to TM,$$

defined by $\phi(t,x,v) = \phi_t(x,v)$. Observe also that by ODE theory the flow is continuous, in fact locally $\mathcal{C}^1$.

Minimizers in (2.4) are usually obtained considering a.c. functions as set of admissible curves and then applying the direct methods, essentially following the approach introduced since Tonelli's works on the Calculus of Variations (see [22]). Sometimes we refer to minimizers of (2.4) as Tonelli minimizers. By regularity theory one has that extremal curves are in fact $\mathcal{C}^2$ and then they solve the Euler-Lagrange equation (2.5). Moreover, Tonelli minimizers are extremal.

We end this section with some useful lemmas.

Lemma 2.1.4. *There exists an increasing function $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that*

$$\forall (x,v) \in TM : \|\frac{\partial}{\partial v}L(x,v)\| \leq f(\|v\|).$$

Proof. By Young inequality (2.2) and (2.3) we have

$$\langle \frac{\partial}{\partial v}L(x,v),v'\rangle \leq L(x,v') + H(x,\frac{\partial}{\partial v}L(x,v)) = L(x,v') + E(x,v).$$

Take $l_0 = \max_{x\in M} E(x,0), c = \max_{(x,v)\in TM}\{L(x,v) \mid \|v\| \leq 1\}$, and define

$$g(r) = \sup_{(x,v)\in TM}\left\{w \cdot \frac{\partial^2}{\partial v^2}L(x,v)\cdot w \mid \|w\| = 1, \|v\| \leq r\right\}.$$

We can evaluate

$$E(x,v) = E(x,0) + \int_0^{\|v\|} \frac{d}{ds}E(x,s\frac{v}{\|v\|})ds =$$

$$= E(x,0) + \int_0^{\|v\|} s\frac{v}{\|v\|}\cdot\frac{\partial^2}{\partial v^2}L(x,s\frac{v}{\|v\|})\cdot\frac{v}{\|v\|}ds \Rightarrow$$

$$\Rightarrow E(x,v) \le l_0 + \int_0^{\|v\|} sg(s)ds \le l_0 + \frac{1}{2}g(\|v\|)\|v\|^2.$$

Then for $\|v'\| \le 1$ we have

$$\langle \frac{\partial}{\partial v}L(x,v), v'\rangle \le c + l_0 + \frac{1}{2}g(\|v\|)\|v\|^2 = f(\|v\|).$$

Taking the supremum with respect to $\|v'\| \le 1$ the statement follows. □

If $x, y \in M$ and $T > 0$ let us define

$$\mathcal{C}_T(x,y) := \{\gamma : [0,T] \to M \ a.c. \mid \gamma(0) = x, \gamma(T) = y\}.$$

We have the following

Lemma 2.1.5. *Let $\varepsilon > 0$ and $T \ge \varepsilon$. Then there exists a constant c_ε such that for every Tonelli minimizer $\gamma \in \mathcal{C}_T(x,y)$ it results*

$$\forall t \in [0,T] : \|\dot\gamma(t)\| \le c_\varepsilon.$$

Proof. Let σ be a geodesic between x and y defined on $[0,T]$ and parametrized by constant speed, so that

$$\|\dot\sigma(t)\| = \frac{d(x,y)}{T} \le \frac{d(x,y)}{\varepsilon} \le \frac{diam(M)}{\varepsilon}.$$

Take $b_\varepsilon = \max\left\{L(x,v) \mid \|v\| \le \frac{diam(M)}{\varepsilon}\right\}$. Hence

$$\int_0^T L(\gamma,\dot\gamma)dt \le \int_0^T L(\sigma,\dot\sigma)dt \le b_\varepsilon T.$$

By superlinearity we deduce that $\|\dot\gamma(t)\| \le c_\varepsilon$ for some constant c_ε. □

Lemma 2.1.6. *For every $t > 0$ there exists a compact set $K_t \subset TM$, such that for every Tonelli minimizer $\gamma : [a,b] \to M$, with $b - a \ge t$ it results*

$$\forall s \in [a,b] : (\gamma(s), \dot\gamma(s)) \in K_t.$$

Proof. It is sufficient to prove the statement in the case $[a,b] = [0,t]$. Indeed, fixed $t_0 \in [a,b]$, we can find a constant c such that $t_0 \in [c, c+t] \subset [a,b]$. If we define $\gamma_c(s) = \gamma(s+c)$ we can apply Lemma 2.1.6 to the curve $\gamma_c : [0,t] \to M$ to obtain the result. Hence let us prove the lemma for $\gamma : [0,t] \to M$. Take $x = \gamma(0), y = \gamma(t)$, and σ a geodesic between x, y such that $\|\dot\sigma(s)\| = \frac{d(x,y)}{t}$. Consider the compact set

$$A_t = \left\{(x,v) \in TM \mid \|v\| \le \frac{diam(M)}{t}\right\}.$$

Since $(\sigma(s), \dot\sigma(s)) \in A_t$ there exists a constant c_t such that

$$\int_0^t L(\gamma,\dot\gamma)dt \le \int_0^t L(\sigma,\dot\sigma)dt \le tc_t.$$

By the mean value theorem we can find $s_0 \in [0,t]$ such that

$$L(\gamma(s_0), \dot\gamma(s_0)) \le c_t.$$

By superlinearity the set $B_t = \{(x,v) \in TM \mid L(x,v) \le c_t\}$ is compact. By continuity of the flow we can then take the compact set

$$K_t = \phi([-t,t] \times B_t) = \bigcup_{|s|\le t} \phi_s(B_t).$$

Since γ is an extremal curve, by remark 2.1.2 we have

$$\forall s \in [0,t] : (\gamma(s), \dot\gamma(s)) = \phi_{s-s_0}(\gamma(s_0), \dot\gamma(s_0)) \in \phi_{s-s_0}(B_t) \subset K_t.$$

□

2.2 The Lax-Oleinik semigroup

We introduce a well known semigroup of non-linear operators $T_t : \mathcal{C}(M) \to \mathcal{C}(M)$. This semigroup, known as Lax-Oleinik semigroup, has many applications in PDE and in Calculus of Variations. Recently it was used to obtain weak KAM solution as shown by Fathi (see [42, 43]). The properties of this semigroup will be used to extend Brenier's formulation of the Monge-Kantorovich problem to the Lagrangian settings.

Definition 2.2.1. *Let $u \in \mathcal{C}(M)$ and $t > 0$. The Lax-Oleinik semigroup is defined as follows*

$$T_t u(x) = \inf_{\gamma} \left\{ u(\gamma(0)) + \int_0^t L(\gamma(s), \dot\gamma(s)) ds \right\},$$

where the infimum is taken over all a.c. curves $\gamma : [0,t] \to M$ such that $\gamma(t) = x$.

Observe that the definition is well posed since, by testing as usual against a geodesic, for some constant c_t we have

$$-\|u\|_\infty \le T_t u(x) \le \|u\|_\infty + c_t t,$$

and so $T_t u(x) \in \mathbb{R}$. Directly from the definition we have the following

Proposition 2.2.2. *The Lax-Oleinik semigroup has the following properties:*

1. *(Semigroup property)* $\forall t, t' \ge 0 : T_{t+t'} = T_t \circ T_{t'}$,
2. *(monotonicity)* $\forall u, v \in \mathcal{C}(M), \forall t > 0 : u \le v \Rightarrow T_t u \le T_t v$,
3. $\forall u \in \mathcal{C}(M), \forall c \in \mathbb{R} : T_t(u + c) = T_t u + c$,
4. *(non-expansive property)* $\forall u, v \in \mathcal{C}(M), \forall t > 0 : \|T_t u - T_t v\|_\infty \le \|u - v\|_\infty$.

Lemma 2.2.3. *If $t > 0$, $u \in \mathcal{C}(M)$, and $x \in M$ are given then there exists an extremal curve $\gamma : [0,t] \to M$ such that $\gamma(t) = x$ and*

$$T_t u(x) = u(\gamma(0)) + \int_0^t L(\gamma(s), \dot\gamma(s)) ds.$$

Furthermore, $\gamma : [0,t] \to M$ is a Tonelli minimizer among all a.c. curves $\sigma : [0,t] \to M$ with the same endpoints.

Proof. Fix $y \in M$ and $\gamma_y \in \mathcal{C}_t(y,x)$ a Tonelli minimizer. We have

$$T_t u(x) = \inf_{y \in M} \left\{ u(\gamma_y(0)) + \int_0^t L(\phi_s(\gamma_y(0), \dot\gamma_y(0)) ds \right\}.$$

Take a sequence of points such that

$$T_t u(x) = \lim_{n\to+\infty} \left\{ u(\gamma_{y_n}(0)) + \int_0^t L(\phi_s(\gamma_{y_n}(0), \dot\gamma_{y_n}(0))ds \right\}.$$

By Lemma 2.1.6 we find a compact set K_t such that

$$(\gamma_{y_n}(0), \dot\gamma_{y_n}(0)) \in K_t.$$

Then we can assume that

$$(\gamma_{y_n}(0), \dot\gamma_{y_n}(0)) \to (y_\infty, v_\infty) \in K_t, \text{ for } n \to +\infty.$$

By continuity of the flow we have

$$T_t u(x) = u(y_\infty) + \int_0^t L(\phi_s(y_\infty, v_\infty))ds.$$

It turns out that the curve $\gamma(s) = \pi_x(\phi_s(y_\infty, v_\infty))$ satisfy all the claims. $\square$

If $k \in \mathbb{R}$, $T > 0$ and $x, y \in M$ are given, it is useful to introduce the time-potential as follows

$$\phi_k(x, y, T) := \min_{\mathcal{C}_T(x,y)} \left\{ \int_0^T L(\gamma(t), \dot\gamma(t))dt + kT \right\}. \tag{2.8}$$

We have the following

Theorem 2.2.4. *For every $\varepsilon > 0$ and $k \in \mathbb{R}$ the time-potential ϕ_k is Lipschitz on $M \times M \times [\varepsilon, +\infty[$.*

Proof. As a first step, we prove that the function $t \in [\varepsilon, +\infty[\mapsto \phi_k(x, y, t)$ is uniformly Lipschitz. Let $T > \varepsilon$, consider a Tonelli minimizer γ on $\mathcal{C}_T(x, y)$ and set $h(s) := \phi_k(x, y, s)$ for any $s > 0$. If we take $\gamma_s(t) = \gamma(\frac{T}{s}t)$ for $t \in [0, s]$ then

$$h(s) \le \int_0^s L(\gamma_s(t), \dot\gamma_s(t))dt + ks := B(s). \tag{2.9}$$

Observe that $B(s) = \int_0^s L(\gamma(\frac{T}{s}t), \frac{T}{s}\dot\gamma(\frac{T}{s}t))dt + ks$, and that $h(T) = B(T)$. Differentiating we obtain

$$B'(T) = k + L(\gamma(T), \dot\gamma(T)) + \int_0^T \frac{d}{ds} L\left(\gamma_s, \frac{T}{s}\dot\gamma(\frac{T}{s}t)\right)_{s=T} dt =$$

$$= k + L(\gamma(T), \dot\gamma(T)) + \int_0^T L_x(\gamma, \dot\gamma)(-\frac{t}{T})\dot\gamma dt + \int_0^T L_v(\gamma, \dot\gamma)(-\frac{1}{T}\dot\gamma(t) - \frac{t}{T}\ddot\gamma(t))dt =$$

$$= k + L(\gamma(T), \dot\gamma(T)) - \frac{1}{T}\int_0^T \frac{d}{dt}[L(\gamma, \dot\gamma)t]dt + \frac{1}{T}\int_0^T [L(\gamma, \dot\gamma) - \langle L_v(\gamma, \dot\gamma), \dot\gamma\rangle]dt =$$

$$= \frac{1}{T}\int_0^T [k - E(\gamma, \dot\gamma)]dt.$$

By Lemma 2.1.5 we obtain

$$B'(T) \le C_\varepsilon + |k|.$$

The same estimate also holds for every $t > \varepsilon$. If $S > T > \varepsilon$, by (2.9) we have

$$\phi_k(x, y, S) - \phi_k(x, y, T) = h(S) - h(T) \le B(S) - B(T) =$$

$$= \int_T^S B'(t)dt \le (C_\varepsilon + |k|)|S - T|.$$

Reversing the role of S, T we have the uniform Lipschitz condition. Now, fix $y_0, x_0 \in M$ and $T > 0$. We prove the Lipschitz condition for the map $y \mapsto \phi_k(x_0, y, T)$ in a neighborhood of y_0. The proof for a neighborhood of x_0 will be similar. Since M is compact this will be enough to conclude. Let $\gamma \in \mathcal{C}_T(x_0, y_0)$ be a Tonelli minimizer and $\sigma : [0,1] \to M$ be a geodesic from y_0 to y. Since the flow is locally $\mathcal{C}^1$ we can find a smooth variation of extremals $\eta(s,t) : [0,1] \times [T-\delta, T] \to M$ such that

$$\eta(0,t) = \gamma(t), \quad \eta(s,T) = \sigma(s), \quad \eta(s, T-\delta) = \gamma(T-\delta).$$

Define

$$\mathcal{E}(s) = \int_{T-\delta}^{T} L\left(\eta(s,t), \partial_t \eta(s,t)\right) dt + k\delta.$$

Hence

$$\mathcal{E}'(s) = \int_{T-\delta}^{T} \left[L_x\left(\eta(s,t), \partial_t\eta(s,t)\right) \frac{\partial \eta}{\partial s}(s,t) + L_v\left(\eta(s,t), \partial_t\eta(s,t)\right) \frac{\partial^2 \eta}{\partial s \partial t}(s,t) \right] dt =$$

$$= \left[L_v\left(\eta(s,t), \partial_t\eta(s,t)\right) \frac{\partial \eta}{\partial s}(s,t) \right]_{T-\delta}^{T} +$$

$$+ \int_{T-\delta}^{T} \left[L_x\left(\eta(s,t), \partial_t\eta(s,t)\right) - \frac{d}{dt} L_v\left(\eta(s,t), \partial_t\eta(s,t)\right) \right] \frac{\partial \eta}{\partial s}(s,t) dt =$$

$$= \langle L_v\left(\sigma(s), \partial_t\eta(s,T)\right), \dot\sigma(s) \rangle.$$

Where we used the fact that $\eta(s,t)$ is an extremal in the last equality. Furthermore, if $\delta > 0$ and $d(y, y_0)$ are small enough the extremal $\eta(s,t)$ is in fact a Tonelli minimizer (see for example section 3.5 in [43]). Then, by Lemma 2.1.5, we can assume that $\|\partial_t \eta(s,T)\| \le c_\delta$. By Lemma 2.1.4 we have that

$$\|L_v(\sigma(s), \partial_t\eta(s,t))\| \le f(c_\delta).$$

Thus

$$\phi_k(x_0, y, T) \le \int_0^{T-\delta} L(\gamma, \dot\gamma) dt + \int_{T-\delta}^{T} L(\eta(1,t), \partial_t \eta(1,t)) dt + kT =$$

$$= \phi_k(x_0, y_0, T) + \mathcal{E}(1) - \mathcal{E}(0) = \phi_k(x_0, y_0, T) + \int_0^1 \mathcal{E}'(s) ds \le$$

$$\le \phi_k(x_0, y_0, T) + \int_0^1 f(c_\delta) \|\dot\sigma(s)\| ds = \phi_k(x_0, y_0, T) + f(c_\delta) d(y_0, y).$$

Changing the role of y_0 and y, $f(c_\delta)$ is a local Lipschitz constant for the map $y \mapsto \phi_k(x_0, y, T)$. $\square$

Observe that by Lemma 2.2.3 and (2.8)

$$T_t u(x) = \min_{z \in M} \{u(z) + \phi_0(z, x, t)\}.$$

We have the following regularizing property of the Lax-Oleinik semigroup.

Theorem 2.2.5. *For every $\varepsilon > 0$ and $u, v \in \mathcal{C}(M)$ there exists $K_\varepsilon > 0$ s.t.*
1) $T_t u$ is K_ε-Lipschitz for all $t \ge \varepsilon$,
2) $\forall s, t \ge \varepsilon : |T_t u(x) - T_s v(x)| \le \|u - v\|_\infty + K_\varepsilon |s - t|$.

Proof. Given $u \in \mathcal{C}(M)$, $x, y \in M$ and $t \geq \varepsilon$, fix K_ε a Lipschitz constant for $\phi_0(x,y,t)$ as in Theorem 2.2.4 on $M \times M \times [\varepsilon, +\infty[$. Let $z \in M$ such that

$$T_t u(y) = u(z) + \phi_0(z,y,t).$$

Hence

$$T_t u(x) \leq u(z) + \phi_0(z,x,t) \leq u(z) + \phi_0(z,y,t) + |\phi_0(z,x,t) - \phi_0(z,y,t)| \leq$$
$$\leq T_t u(y) + K_\varepsilon d(x,y).$$

Reversing the role of x and y we obtain that

$$|T_t u(x) - T_t u(y)| \leq K_\varepsilon d(x,y).$$

To check the second point, take $u, v \in \mathcal{C}(M)$ and $s, t \geq \varepsilon$. Let $z \in M$ such that

$$T_t u(x) = u(z) + \phi_0(z,x,t).$$

Then

$$T_s v(x) \leq v(z) + \phi_0(z,x,s) \leq |v(z) - u(z)| + u(z) + \phi_0(z,x,t) +$$
$$+ |\phi_0(z,x,s) - \phi_0(z,x,t)| \leq T_t u(x) + \|u - v\|_\infty + K_\varepsilon |s-t|.$$

Reversing the role of u and v the theorem is proved. ☐

Corollary 2.2.6. *For $u \in \mathcal{C}(M)$, set $\phi(t,x) = T_t u(x)$.*
Then for every $\varepsilon > 0$: $\phi(t,x)$ is K_ε-Lipschitz on $[\varepsilon, +\infty[\times M$.

Proof. By Theorem 2.2.5 we can estimate

$$|\phi(t,x) - \phi(s,y)| \leq |T_t u(x) - T_s u(x)| + |T_s u(x) - T_s u(y)| \leq K_\varepsilon |s-t| + K_\varepsilon d(x,y).$$

☐

By Rademacher's theorem the function $T_t u(x)$ is differentiable for $\mathcal{L}^n$-a.e. $x \in M$. The following lemma state a useful differentiability property of the Lax-Oleinik semigroup.

Lemma 2.2.7. *Let $u \in \mathcal{C}(M)$, $t > 0$ and $x \in M$ a point of differentiability of $T_t u$.*
If $\gamma : [0,t] \to M$ is a Tonelli minimizer as in Lemma 2.2.3 then

$$d_x T_t u(x) = \frac{\partial L}{\partial v}(x, \dot{\gamma}(t)).$$

Proof. As in Lemma 2.2.3 γ is such that

$$T_t u(x) = u(\gamma(0)) + \int_0^t L(\gamma(s), \dot{\gamma}(s)) ds.$$

Fix $v \in T_x M$ and take a smooth variation of extremals $\eta(s,\tau)$ such that

$$\eta(s,0) = \gamma(0), \quad \eta(t,\tau) = \gamma(\tau), \quad \frac{\partial \eta}{\partial s}(t,t) = v. \tag{2.10}$$

Define $A(s) = \int_0^t L(\eta(s,\tau), \partial_\tau \eta(s,\tau)) d\tau$. Hence

$$A'(t) = \int_0^t \left[L_x(\eta(t,\tau), \partial_\tau \eta(t,\tau)) \frac{\partial \eta}{\partial s}(t,\tau) + L_v(\eta(t,\tau), \partial_\tau \eta(t,\tau)) \frac{\partial^2 \eta}{\partial s \partial \tau}(t,\tau) \right] d\tau =$$

$$= \left[L_v(\eta(t,\tau), \partial_\tau \eta(t,\tau)) \frac{\partial \eta}{\partial s}(t,\tau) \right]_0^{\tau = t} + \int_0^t (L_x - \frac{d}{d\tau} L_v) \frac{\partial \eta}{\partial s}(t,\tau) d\tau =$$

$$= \langle L_v(\gamma(t), \dot\gamma(t)), v\rangle. \tag{2.11}$$

Since γ is a Tonelli minimizer and by (2.10), for $s \le t$ we can evaluate

$$\frac{1}{t-s}[T_t u(\eta(s,t)) - T_t u(x)] =$$

$$= \frac{1}{t-s}[T_t u(\eta(s,t)) - u(\gamma(0)) - \int_0^t L(\gamma(s), \dot\gamma(s))ds] \le$$

$$\le \frac{1}{t-s}[u(\eta(s,0)) + \int_0^t L(\eta(s,\tau), \partial_\tau \eta(s,\tau))d\tau - u(\gamma(0)) - \int_0^t L(\gamma(s), \dot\gamma(s))ds] =$$

$$= \frac{1}{t-s}[A(s) - A(t)].$$

Letting $s \to t^-$ we obtain

$$\langle d_x(T_t u(\eta(t,t)), \partial_s \eta(t,t)\rangle \le A'(t).$$

Hence by (2.11) and by (2.10) we have

$$\langle d_x T_t u(x), v\rangle \le \langle L_v(x, \dot\gamma(t)), v\rangle.$$

Applying the same inequality to $-v$ and by arbitrariness of v the conclusion follows. □

2.3 The weak KAM theorem

Let $M = \mathbb{T}^n$ be the n-dimensional flat torus. In this case we can take the Hamiltonian as a function $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that $H(\cdot, p)$ is periodic. The canonical Hamilton's ODE are

$$\begin{cases} \dot x(t) = D_p H(x,p) \\ \dot p(t) = -D_x H(x,p). \end{cases} \tag{2.12}$$

Let us suppose that for all $P \in \mathbb{R}^n$ there exists a real number $\bar H(P)$ such that the Hamilton-Jacobi equation

$$H(x, D_x S) = \bar H(P) \tag{2.13}$$

admits a smooth solution $S(x,P)$. Let us also suppose that the function $\bar H(P)$ is smooth and that the relations

$$p = D_x S(x,P), \quad X = D_P S(x,P) \tag{2.14}$$

define a smooth change of coordinates. Differentiating the first relation of (2.14) and by (2.12) we have

$$-D_x H = \dot p = D^2_{xx} S \dot x + D^2_{Px} S \dot P \Rightarrow \dot P D^2_{Px} S = -(D_x H + D^2_{xx} S D_p H).$$

On the other hand, differentiating (2.13) we have

$$0 = D_x \bar H(P) = D_x H + D_p H D^2_{xx} S.$$

Hence we infer that $\dot P = 0$. Now, differentiating the second relation in (2.14) we obtain

$$\dot X = D^2_{Px} S \dot x + D^2_{PP} S \dot P = D_p H D^2_{Px} S = D_P \bar H.$$

Hence in the new coordinates (2.14) the Hamilton's ODE become

$$\begin{cases} \dot X(t) = D_P \bar H(P) \\ \dot P(t) = 0. \end{cases} \tag{2.15}$$

An Hamiltonian system which can be reduced to the form of (2.15) is said to be completely integrable, while the function $S(x,P)$ is called generating function and the function $\bar{H}$ is referred as the effective Hamiltonian. Usually one looks for generating functions of the form $S = \langle P, x\rangle + u(x)$ with u a periodic function, so that the Hamilton-Jacobi equation to solve is

$$H(x, P + D_x u) = \bar{H}(P). \tag{2.16}$$

Completely integrable systems can be explicitly solved. Actually it is possible to show that the phase-space $\{(x,p) \in \mathbb{R}^{2n}\}$ is foliated by tori that are invariant under the motion and, on each such torus, the motion is given by straight lines. This classical strategy in general may fail since the equation (2.16) does not usually admit a smooth solution and, even if it does, the transformations (2.14) are not usually globally defined. Only special Hamiltonians are integrable in this sense. The celebrated KAM (Kolmogorov-Arnold-Moser) theorem applies for Hamiltonian of the form $H(x,p) = H_0(p) + \delta H_1(x,p)$ finding a smooth solution $u(x,P)$ of the equation (2.16) for small perturbation $\delta > 0$ and suitable condition on H. In general situations one have to look for solutions of the Hamilton-Jacobi equations like (2.16) in a weak sense. Typically this weak sense is understood in the sense of viscosity solutions. For a brief introduction on this type of solutions we refer to [38] while for the connections of this solutions with the Hamiltonian systems we refer to [41, 52]. An extensive study on weak KAM solutions and the connections with the Lagrangian dynamic is due to Fathi, see [42, 43, 44]. For these problems see also [25, 26, 27, 28]. We will derive a weak KAM solution essentially using the properties of the Lax-Oleinik semigroup as in [43].

Lemma 2.3.1. *Let E be a Banach space and $\varphi : E \to E$ a continuous function such that $\varphi(E)$ is relatively compact. Then φ admits a fixed point.*

Proof. Take $K = \overline{co}(\varphi(E))$ which is a closed and totally bounded set (see Theorem 3.20 in [74]) and hence compact. Then we can apply the Schauder-Tychonoff theorem (see Theorem 5.28 in [74]) to the function $\varphi : K \to K$. ◻

Theorem 2.3.2. *There exists $u \in \mathrm{Lip}(M)$ and a unique $c_0 \in \mathbb{R}$ such that*

$$\forall t \geq 0 \quad u = T_t u + c_0 t.$$

Proof. Consider the equivalence relation: $u \sim v \Leftrightarrow \exists a \in \mathbb{R}$ *s.t.* $u = v + a$, and take the quotient space $E = \frac{\mathcal{C}(M)}{\sim}$. Let $t > 0$. Since by Proposition 2.2.2 $T_t(u+a) = T_t u + a$, the semigroup T_t induces on the quotient space the semigroup $\widetilde{T}_t[u] = [T_t u]$. Comparing with a geodesic σ we have

$$\|\widetilde{T}_t[u]\| = \|[T_t u]\| = \inf_{a \in R} \|T_t u + a\|_\infty \leq \int_0^t L(\sigma, \dot{\sigma}) ds \leq Ct.$$

Where $C = \sup\{L(x,v) | x \in M, \|v\| \leq \frac{diam(M)}{t}\}$. Since $T_t u$ is K_t-Lipschitz by Theorem 2.2.5, by Ascoli-Arzelà theorem we infer that $\widetilde{T}_t(E)$ is relatively compact. By Lemma 2.3.1 we find a fixed point $[u_t]$ for $\widetilde{T}_t$. Observe that by the semigroup property of Proposition 2.2.2 $[u_t]$ is a fixed point also for $\widetilde{T}_{kt}, \forall k \in \mathbb{N}$. If we take $t_n = 1/2n$ and $[u_{t_n}]$ fixed point for $\widetilde{T}_{t_n}$, again by Ascoli-Arzelà theorem we can suppose that $u_{t_n} \to u$ uniformly as $n \to +\infty$. Let $q \in \mathbb{Q}$, we can find a subsequence $t_{h(n)}$ such that $q = k_n t_{h(n)}$ with $k_n \in \mathbb{N}$. Hence, by Proposition 2.2.5 we have

$$\|\widetilde{T}_q[u] - [u]\| \leq \|\widetilde{T}_{k_n t_{h(n)}}[u] - \widetilde{T}_{k_n t_{h(n)}}[u_{h(n)}]\| + \|\widetilde{T}_{k_n t_{h(n)}}[u_{h(n)}] - [u]\| \leq$$

$$\leq \|T_{k_n t_{h(n)}} u - T_{k_n t_{h(n)}} u_{h(n)}\|_\infty + \|[u_{h(n)}] - [u]\| \leq 2\|u_{h(n)} - u\|_\infty \to 0, \quad \text{as } n \to +\infty.$$

If $q_n \searrow t$ by Proposition 2.2.5 we have

$$\|\widetilde{T}_t[u] - [u]\| \leq \|\widetilde{T}_t[u] - \widetilde{T}_t q_n[u]\| + \|\widetilde{T}_{q_n}[u] - [u]\| \leq$$

$$\leq \|T_t u - T_{q_n} u\|_\infty \leq K_t |t - q_n| \to 0, \quad \text{as } n \to +\infty.$$

Then there exists a common fixed point for $\widetilde{T}_t$ for every $t > 0$. Coming back to T_t we have that there exists a constant $a_t \in \mathbb{R}$ and $u \in \mathrm{Lip}(M)$ such that $T_t u = u + a_t$. Observe that by Proposition 2.2.2 it results $a_{t+t'} = a_t + a_{t'}$. Furthermore by Proposition 2.2.5 the function $t \mapsto a_t$ is continuous. Hence we deduce that a_t is linear and since $a_0 = 0$ we have that $a_t = ct$ with $c = a_1$. Taking $c_0 = -c$ the existence part of the statement follows. Let us now suppose to have two constants c_0 and c_0' and two Lipschitz functions u and v such that

$$u = T_t u + c_0 t, \quad v = T_t v + c_0' t.$$

We deduce that

$$u - v = T_t u - T_t v + (c_0 - c_0')t.$$

Since T_t is non-expansive by Proposition 2.2.2, Letting $t \to +\infty$ we obtain a contradiction unless $c_0 = c_0'$. □

We will show that the function $u \in \mathrm{Lip}(M)$ of Theorem 2.3.2 solves the Hamilton-Jacobi equation $H(x, d_x u) = c_0$. First we give the following

Definition 2.3.3. *Let $u \in \mathcal{C}(M)$ and $c \in \mathbb{R}$. We say that u is dominated by $L + c$, and we use the notation $u \prec L + c$, if for every $a \leq b \in R$ and for every $\mathcal{C}^1$ curve $\gamma : [a, b] \to M$ we have*

$$u(\gamma(b)) - u(\gamma(a)) \leq \int_a^b L(\gamma(s), \dot\gamma(s))ds + c(b - a).$$

The following lemma is a simple consequence of the definition.

Lemma 2.3.4. *Let $u \in \mathcal{C}(M)$ and $c \in \mathbb{R}$. Then*

$$u \prec L + c \Leftrightarrow \forall t \geq 0 \quad u \leq T_t u + ct.$$

Definition 2.3.5. *Let $u \in \mathcal{C}(M)$ and $c \in \mathbb{R}$. A $\mathcal{C}^1$ curve $\gamma : I \to M$ is said to be $(u - c)$-calibrated if*

$$\forall t \leq t' \in I \quad u(\gamma(t')) - u(\gamma(t)) = \int_t^{t'} L(\gamma(s), \dot\gamma(s))ds + c(t' - t).$$

The following result relates these two notions.

Lemma 2.3.6. *Let $\gamma : [a, b] \to M$ be a $\mathcal{C}^1$ curve, $u \in \mathcal{C}(M)$ and $c \in \mathbb{R}$ such that $u \prec L + c$. If*

$$u(\gamma(b)) - u(\gamma(a)) = \int_a^b L(\gamma(s), \dot\gamma(s))ds + c(b - a),$$

then γ is $(u - c)$-calibrated.

Proof. Let $t \leq t' \in [a, b]$. Since $u \prec L + c$

$$u(\gamma(b)) - u(\gamma(t')) \leq \int_{t'}^b L(\gamma(s), \dot\gamma(s))ds + c(b - t'), \tag{2.17}$$

$$u(\gamma(t')) - u(\gamma(t)) \leq \int_t^{t'} L(\gamma(s), \dot\gamma(s))ds + c(t' - t),$$

$$u(\gamma(t)) - u(\gamma(a)) \leq \int_a^t L(\gamma(s), \dot\gamma(s))ds + c(t - a).$$

Adding these inequalities we have

$$u(\gamma(b)) - u(\gamma(a)) \leq \int_a^b L(\gamma(s), \dot{\gamma}(s))ds + c(b-a)$$

which is an equality by hypothesis. Hence we deduce that all the inequalities in (2.17) are in fact equalities and then the Lemma is proved. $\square$

Theorem 2.3.7. *Let $u \in \mathrm{Lip}(M)$ as in Theorem 2.3.2. Then for every point of differentiability x of u it results*

$$H(x, d_x u) = c_0.$$

Proof. Fix $v \in T_x(M)$ and a $\mathcal{C}^1$ curve γ such that $\gamma(0) = x, \dot{\gamma}(0) = v$. By Lemma 2.3.4 we have that $u \prec L + c_0$ and then

$$u(\gamma(t)) \leq u(x) + \int_0^t L(\gamma(s), \dot{\gamma}(s))ds + c_0 t.$$

Dividing by t we obtain

$$\frac{u(\gamma(t)) - u(x)}{t} \leq \frac{1}{t}\int_0^t L(\gamma(s), \dot{\gamma}(s))ds + c_0.$$

Passing to the limit as $t \to 0^+$ we have

$$\langle d_x u(x), v\rangle \leq L(x, v) + c_0.$$

Taking the supremum with respect to v we obtain

$$H(x, d_x u) \leq c_0.$$

To check the reverse inequality take γ as in Lemma 2.2.3. Hence, by Theorem 2.3.2 we have

$$u(\gamma(t)) = u(\gamma(0)) + \int_0^t L(\gamma(s), \dot{\gamma}(s))ds + c_0 t.$$

By Lemma 2.3.6 γ is $(u - c_0)$-calibrated. Then for every $t' \leq t$ we have

$$u(\gamma(t)) - u(\gamma(t')) = \int_{t'}^t L(\gamma(s), \dot{\gamma}(s))ds + c_0(t - t'),$$

and then

$$\frac{u(\gamma(t)) - u(\gamma(t'))}{t - t'} = \frac{1}{t - t'}\int_{t'}^t L(\gamma(s), \dot{\gamma}(s))ds + c_0.$$

Passing to the limit as $t' \to t^-$ we obtain

$$\langle d_x u(x), \dot{\gamma}(t)\rangle = L(x, \dot{\gamma}(t)) + c_0 \Rightarrow c_0 \leq H(x, d_x u).$$

$\square$

2.4 Minimal measures and the Mañé's critical value

An important problem in dynamic is to study "minimal" orbits of Hamiltonian systems, where for "minimal" we mean action minimizing. In order to consider general situations in more degrees of freedom, Mather proposed the notion of minimal measures, see [64, 65]. The setting introduced by

Mather has the important feature that the results are formulated in terms of invariant measures. Here we briefly describe some aspects of these objects. For a more complete presentation we also refer to [7, 26, 43, 62].

A measure $\mu \in \mathcal{M}(TM)$ is said to be closed if for all exact forms ω it results

$$\int_{TM} \langle \omega(x), v \rangle d\mu = 0. \tag{2.18}$$

According to this definition we set

$$\mathcal{M}^c = \left\{ \mu \in \mathcal{P}(TM) \mid \mu \text{ is closed, } \int_{TM} L(x,v) d\mu < +\infty \right\}.$$

If $\mu \in \mathcal{M}^c$, thanks to the fact that L is superlinear and $\int_{TM} L(x,v)d\mu < +\infty$, μ acts in a natural way on the set of the closed 1-forms on M by

$$\omega \mapsto \int_{TM} \langle \omega(x), v \rangle d\mu$$

and thanks to condition (2.18) this action passes to the quotient by the exact forms. In other words, if we denote by $H^1(M,\mathbb{R})$ the first real cohomology group on M, the action of μ over the elements of $H^1(M,\mathbb{R})$ is well defined. Hence, by duality we can consider the homology class of μ, which we will denote by $\rho(\mu)$, in $H_1(M,\mathbb{R})$. In dynamic the homology class $\rho(\mu)$ is also called rotation vector. The action of $\rho(\mu)$ on a cohomology class $[\omega]$ is given by

$$\langle \rho(\mu), \omega \rangle = \int_{TM} \langle \omega(x), v \rangle d\mu.$$

The notion of action functional on curves is generalized by the notion of action functional on measures $\mu \in \mathcal{M}^c$ defined by

$$\mathcal{A}(\mu) := \int_{TM} L(x,v) d\mu.$$

Finally we define the β Mather's function $\beta : H_1(M,\mathbb{R}) \to \mathbb{R}$ by

$$\beta([h]) = \inf_{\mathcal{M}^c} \left\{ \mathcal{A}(\mu) \mid \rho(\mu) = [h] \right\}. \tag{2.19}$$

Theorem 2.4.1. *For every homology class $[h] \in H_1(M,\mathbb{R})$ the infimum in (2.19) is attained.*

Proof. Let μ_n a minimizing sequence for $\beta([h])$. By superlinearity we can assume that

$$\int_{TM} \|v\| d\mu_n \leq C < +\infty.$$

By compactness of measures up to subsequences $\mu_n \overset{*}{\rightharpoonup} \mu$. By lower semicontinuity we have that $\int_{TM} \|v\| d\mu \leq \liminf_{n\to+\infty} \int_{TM} \|v\| d\mu_n \leq C$. Hence, for each $\varepsilon > 0$, we can find $R > 0$ such that

$$\int_{\|v\|>R} \|v\| d\mu < \varepsilon, \quad L(x,v) \geq \frac{C}{\varepsilon}\|v\| \quad \text{ for } \|v\| > R.$$

We can evaluate

$$|\int_{TM} \|v\| d\mu_n - \int_{TM} \|v\| d\mu| \leq$$

$$\leq |\int_{\|v\|\leq R} \|v\| d\mu_n - \int_{\|v\|\leq R} \|v\| d\mu| + \int_{\|v\|>R} \|v\| d\mu_n + \int_{\|v\|>R} \|v\| d\mu <$$

$$< \varepsilon + \frac{\varepsilon}{C} \int_{\|v\|>R} L(x,v)d\mu_n + \varepsilon \leq 3\varepsilon.$$

By 3 of Theorem 1.4.9 we infer that

$$\lim_{n\to+\infty} \int_{TM} \varphi(x,v)d\mu_n = \int_{TM} \varphi(x,v)d\mu,$$

for every continuous function $\varphi(x,v)$ which grows at most linearly on the v variable. In particular we deduce that the limit measure μ is a closed probability measure and that $\rho(\mu) = [h]$. Finally the lower semicontinuity of the action $\mathcal{A}$ with respect to the weak$*$ convergence is enough to conclude. $\square$

The above result allows to give the following

Definition 2.4.2. *Let $[h] \in H_1(M, \mathbb{R})$. A measure $\mu \in \mathcal{M}^c$ with $\rho(\mu) = [h]$ is said to be minimal if $\beta([h]) = \mathcal{A}(\mu)$.*

A remarkable property of minimal measures is that they are invariant under the Lagrangian flow, see [7, 26, 44, 62]. In fact, the relevance of Mather theory is due to the fact that it is a theory of invariant sets and measures under assumptions extremely weaker than those of KAM theory. Actually, it is strictly related to the weak KAM theory as developed in [26, 42, 43, 44]. We point out that these measures are important to construct weak KAM solutions of the Hamilton-Jacobi equation $H(x, d_x u) = c_0$ as in Theorem 2.3.2, for which c_0 results a critical value in the sense that for $c < c_0$ the equation $H(x, d_x u) \leq c$ admits no subsolution, while the equation $H(x, d_x u) \leq c_0$ always admits $\mathcal{C}^1$ subsolutions (see [44]). For these problems we refer also to [26, 43]. This critical value c_0 is known as the Mañé's critical value and it can be characterized through the minimal measures as follows by the next theorem.

Theorem 2.4.3. $-c_0 = \min \{\mathcal{A}(\mu) \mid \mu \in \mathcal{M}^c\}$.

Proof. Let u be a $\mathcal{C}^1$ subsolution of $H(x, d_x u) \leq c_0$ which exists by the results of [44]. Then we have

$$\forall (x,v) \in TM : \langle d_x u(x), v\rangle - L(x,v) \leq H(x, d_x u) \leq c_0 \Rightarrow -c_0 \leq L(x,v) - \langle d_x u(x), v\rangle.$$

Taking $\mu \in \mathcal{M}^c$ and integrating we obtain

$$-c_0 \leq \int_{TM} L(x,v)d\mu.$$

Vice versa, fix $x \in M$ and a sequence $t_n \to +\infty$. By Lemma 2.2.3 and Theorem 2.3.2 there exist Tonelli minimizers $\gamma_n : [0, t_n] \to M$ and $u \in \mathrm{Lip}(M)$ such that

$$u(x) - u(\gamma_n(0)) = \int_0^{t_n} L(\gamma_n(s), \dot\gamma_n(s))ds + c_0 t_n. \tag{2.20}$$

By Lemma 2.1.6 there exists a common compact set $K_1 \subset TM$ such that

$$(\gamma_n(s), \dot\gamma_n(s)) \in K_1.$$

Define the measures $\mu_n \in \mathcal{P}(TM)$ through the formula

$$\int_{TM} f(x,v)d\mu_n = \frac{1}{t_n} \int_0^{t_n} f(\gamma_n(s), \dot\gamma_n(s))ds.$$

Up to subsequence, there exists $\mu \in \mathcal{P}(TM)$ such that $\mu_n \overset{*}{\rightharpoonup} \mu$. Observe that in fact $supp(\mu) \subset K_1$. In particular this implies that $\lim_{n\to+\infty} \int_{TM} f(x,v)d\mu_n = \int_{TM} f(x,v)d\mu$ for every continuous function $f : TM \to \mathbb{R}$. Since M is compact we have

$$\int_{TM} \langle d\varphi, v\rangle d\mu = \lim_{n\to+\infty} \frac{1}{t_n}\int_0^{t_n} \langle d\varphi(\gamma_n(s)), \dot\gamma_n(s)\rangle ds = \lim_{n\to+\infty} \frac{1}{t_n}[\varphi(\gamma_n(t_n)) - \varphi(\gamma_n(0))] = 0.$$

Then $\mu \in \mathcal{M}^c$. Furthermore by (2.20) we have

$$\int_{TM} L(x,v)d\mu = \lim_{n\to+\infty} \frac{1}{t_n}\int_0^{t_n} L(\gamma_n(s), \dot\gamma_n(s))ds =$$

$$= \lim_{n\to+\infty} \frac{1}{t_n}[u(x) - u(\gamma_n(0)) - c_0 t_n] = -c_0.$$

□

Remark 2.4.4. The minimal measure found in the proof of Theorem 2.4.3 is compactly supported. Actually this is a general property of minimal measures (see [64]). Moreover we have found such measure as the limit of measures defined on trajectories of the phase-space. This is also a general fact (see [26, 62]).

It is possible to show (see [26, 43, 64]) that the function β is convex and superlinear on $H_1(M,\mathbb{R})$. If we consider its dual transformation $\alpha : H^1(M,\mathbb{R}) \to \mathbb{R}$, defined by

$$\alpha([\omega]) = \max_{h\in H_1(M,\mathbb{R})} \{\langle \omega, h\rangle - \beta(h)\}, \tag{2.21}$$

where ω is any closed one-form whose cohomology class is $[\omega]$, we also have that

$$-\alpha([\omega]) = \min_{h\in H_1(M,\mathbb{R})} \{\beta(h) - \langle \omega, h\rangle\}. \tag{2.22}$$

Let ω be a closed one-form. Define the new Lagrangian $L - \omega$ putting

$$[L - \omega](x,v) = L(x,v) - \langle \omega, v\rangle,$$

and denote by $\mathcal{A}_\omega$ the associated action functional. If we take a minimizing $h \in H_1(M,\mathbb{R})$ in (2.22) and a minimal measure $\mu \in \mathcal{M}^c$ with $\rho(\mu) = h$, we have

$$-\alpha([\omega]) = \beta(h) - \langle \omega, h\rangle = \int_{TM} L(x,v)d\mu - \langle \omega, \rho(\mu)\rangle = \int_{TM} L - \omega d\mu.$$

Hence we infer that

$$\min\{\mathcal{A}_\omega \mid \mu \in \mathcal{M}^c\} \le -\alpha([\omega]).$$

On the other hand if $\mu \in \mathcal{M}^c$ we have

$$-\alpha([\omega]) \le \beta(\rho(\mu)) - \langle \omega, \rho(\mu)\rangle \le \int_{TM} L - \omega d\mu.$$

From the arbitrariness of $\mu \in \mathcal{M}^c$ we deduce that

$$-\alpha([\omega]) = \min\{\mathcal{A}_\omega \mid \mu \in \mathcal{M}^c\}. \tag{2.23}$$

Denoting with $c_0(L-\omega)$ the Mañé's critical value corresponding to the Lagrangian $L-\omega$, combining (2.23) with Theorem 2.4.3 we obtain the equality

$$\alpha([\omega]) = c_0(L - \omega).$$

Observe that considering the Hamiltonian H' with respect to the Lagrangian $L-\omega$, by Theorem 2.3.7 we can find a weak KAM solution of

$$H'(x, d_x u) = c_0(L-\omega). \tag{2.24}$$

On the other hand we have

$$H'(x,\omega') = \sup_{v\in T_xM}\{\langle\omega', v\rangle - L(x,v) - \langle\omega, v\rangle\} = H(x, \omega'-\omega).$$

Hence we obtain the equation

$$H(x, d_x u - \omega) = c_0(L-\omega). \tag{2.25}$$

Chapter 3

Minimal measures and mass transportation problems

Consider the n-dimensional flat torus $\mathbb{T}^n$. In this case it is known that the homology $H_1(M,\mathbb{R})$ and the cohomology $H^1(M,\mathbb{R})$ are n-dimensional vector spaces (see for example page 95 of [58]). Observe that $\mathcal{C}^1$ functions on $\mathbb{T}^n$ can be regarded as functions in $\mathcal{C}^1(\mathbb{R}^n)$ which are $[0,1]^n$-periodic and vice versa. Since the linear functions in $\mathbb{R}^n$ are not periodic, the 1-forms $dx_1, dx_2, \ldots, dx_n$ are closed but not exact on $\mathbb{T}^n$. Using the universal covering $\varphi : \mathbb{R}^n \to \mathbb{T}^n$ defined by

$$\varphi(x_1,\ldots,x_n) = (e^{2\pi i x_1},\ldots,e^{2\pi i x_n}),$$

we can express the forms $dx_1, dx_2, \ldots, dx_n$ as forms on $\mathbb{R}^n$. Hence they results linearly independent. Then the cohomology on $\mathbb{T}^n$ is generated by the classes $[dx_1],[dx_2],\ldots,[dx_n]$. In other words for every closed 1-forms ω on $\mathbb{T}^n$ there exists $P \in \mathbb{R}^n$ and $f \in \mathcal{C}^1(\mathbb{T}^n)$ such that $\omega = P + df$. Then the Hamilton-Jacobi equation (2.25) on $\mathbb{T}^n$ amounts to

$$H(x, P + d_x u) = c_0(L + P). \tag{3.1}$$

Hence the effective Hamiltonian $\bar{H}(P)$ of (2.16) is given, in the sense of (3.1), by Mañé's critical values.

Fixed $Q \in \mathbb{R}^n$, in this setting Mather's problem can be formulated as follows:

$$\min_{\mu} \int_{\mathbb{T}^n\times\mathbb{R}^n} L(x,v)d\mu$$

under the following constraints:

1. $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$,
2. $\forall f \in \mathcal{C}^1(\mathbb{T}^n) : \int_{\mathbb{T}^n\times\mathbb{R}^n} \langle df, v\rangle d\mu = 0$,
3. $\int_{\mathbb{T}^n\times\mathbb{R}^n} v d\mu = Q$.

Observe that the second constraint in the above list correspond to the closedness of μ while the third one is equivalent to fix the homology class $\rho(\mu)$. On this framework, in [36] L. C. Evans has developed some PDE methods for the weak KAM theory deriving the set of equations:

$$\begin{cases} -div(\sigma \partial_p H(x, d_x u)) = 0 \\ H(x, d_x u) = \bar{H}(P) \quad \sigma - a.e. \end{cases} \tag{3.2}$$

where $\sigma \in \mathcal{P}(\mathbb{T}^n)$ is the projection on $\mathbb{T}^n$ of a Mather's minimal measure and $u \in \mathrm{Lip}(\mathbb{T}^n)$ is a weak KAM solution. The equations (3.2) are similar to the equations (1.29) of transport problems and can be considered as the generalized transport equation and eikonal equation satisfied by u and σ. In [40] it is proposed a linear programming interpretation of Mather's problem, while a linear programming interpretation is a well known fact for the Monge-Kantorovich problem (see for example [37, 73]). In particular, for $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ define:

$$\phi \in \mathcal{C}^1(\mathbb{T}^n) \mapsto L_\mu(\phi) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \langle d\phi(x), v \rangle d\mu,$$

and consider the linear operator

$$A : \mathcal{M}(\mathbb{T}^n \times \mathbb{R}^n) \to \mathcal{C}^1(\mathbb{T}^n)^* \times \mathbb{R} \times \mathbb{R}^n$$

defined by

$$A\mu = \left(L_\mu,\ \mu(\mathbb{T}^n \times \mathbb{R}^n),\ \int_{\mathbb{T}^n \times \mathbb{R}^n} v d\mu \right).$$

If we denote by c^*, the linear functional associated to L by $\mu \mapsto \int_{TM} L(x,v) d\mu$, fixed $Q \in \mathbb{R}^n$ and posed $(0,1,Q) = b$, Mather's problem can be reformulated as

$$\begin{cases} \min \langle c^*, \mu \rangle \\ \mu \geq 0 \\ A\mu = b, \end{cases} \tag{3.3}$$

obtaining an infinite dimensional linear programming problem. Problem (3.3) admits a dual formulation:

$$\begin{cases} \max \langle y^*, b \rangle \\ A^* y^* \leq c^*. \end{cases} \tag{3.4}$$

This formulation is only formal since the space $C^1(\mathbb{T}^n)$ is not reflexive and no general theorem apply to this particular situation. However, in [40] it is shown that there is no gap between the two problem and that if we have a solution y^* of (3.4) such that $y^* = (u, \omega_0, P) \in \mathcal{C}^1(\mathbb{T}^n) \times \mathbb{R} \times \mathbb{R}^n$ then u solve the equation

$$H(x, P + d_x u) = \bar{H}(P) \quad \text{on } supp(\pi_x(\mu)),$$

with $\bar{H}(P) = \omega_0$ and μ a solution of (3.3). Hence the dual of Mather's problem predict the weak KAM equation (see also [51]). In this chapter we just investigate the relationship between the Monge-Kantorovich problem and Mather's problem. A rigorous connection between these two problem will be given by a careful formulation of both of them in terms of 1-currents, generalizing some results of [7].

3.1 Energy on currents and equivalence with Mather's problem

In this section we introduce an energy functional on the space of 1-dimensional currents $\mathcal{N}_1(M)$ associated to the Lagrangian L. For the basic facts about currents we refer to the appendix A. As L is convex, after some considerations one is convinced that the functional should depend not only on the current but also on the parametrization we choose for it. And indeed the definition we give turns out to be the natural one from many points of view that will be clear later on. However some reader may prefer to take as the definition the formula given in Lemma 3.1.3.

For each $\sigma \in \mathcal{P}(M)$ and each $T \in \mathcal{N}_1^c(M)$ we define the following functional:

$$\mathcal{L}(\sigma, T) := \sup_{\alpha, \omega} \int_M \alpha(x) d\sigma + T(\omega) \tag{3.5}$$

where the sup is taken for α continuous function and $\omega \in \Gamma^\infty(T^*M)$ which satisfy the constraint $\alpha(x) + H(x, \omega(x)) \leq 0$ pointwise.

Lemma 3.1.1. *$\mathcal{L}(\sigma,T)$ is positive and if it is finite then there exists a vector field $X \in L^1_\sigma(M)$ such that $T = \sigma \wedge X$.*

Proof. Since $L \geq 0$ one easily deduce that $H(x,0) = \sup\{-L(x,v) \mid v \in T_xM\} \leq 0$, i.e. the couple ($\alpha \equiv 0$, $\omega = 0$) is admissible for the supremum in (3.5) for any (σ,T). It follows that $\mathcal{L}(\sigma,T) \geq C$. Let us now fix (σ,T) and assume that $\mathcal{L}(\sigma,T) < +\infty$. By the representation theorem A.0.7 we can find $\tilde{\sigma} \in \mathcal{P}(M)$ and $X \in L^1_{\tilde{\sigma}}(M)$ such that $T = \tilde{\sigma} \wedge X$ and $g(X,X) = M(T)^2$ $\tilde{\sigma}$-a.e. . Assume by contradiction that $\tilde{\sigma}$ is not absolutely continuous w.r.t. σ, then we can decompose $\tilde{\sigma} = \tilde{\sigma}^a + \tilde{\sigma}^s$ where $\tilde{\sigma}^a << \sigma$ and $\tilde{\sigma}^s$ is singular w.r.t. σ, with $|\tilde{\sigma}^s| > 0$. Hence, we can rewrite the functional $\mathcal{L}(\sigma,T)$ as

$$\mathcal{L}(\sigma,T) = \sup_{\alpha,\omega} \int_M \alpha(x)\, d\sigma + \int_M \langle \omega(x), X(x) \rangle\, d\tilde{\sigma}. \tag{3.6}$$

Let us now consider the 1-form $\tilde{\omega}$ such that for $\tilde{\sigma}$-a.e. x

$$\langle \tilde{\omega}(x), X(x) \rangle = g(X(x),X(x))^{\frac{1}{2}} = M(T)$$

and let A be a Borel set on which $\tilde{\sigma}^s$ is concentrated. We define $\omega(x) := \tilde{\omega}(x)\chi_A(x)$. In general this form will be not continuous but, by using Lusin's theorem with respect to the measure $\sigma + \tilde{\sigma}^s$ on M, we can find regular forms $\omega_\varepsilon(x)$ such that $||\omega_\varepsilon(x)||_{L^\infty(\sigma+\tilde{\sigma}^s)} \leq 1$, $\omega_\varepsilon(x) \to \omega(x)$ pointwise $(\sigma + \tilde{\sigma}^s)$-a.e. and $(\sigma + \tilde{\sigma}^s)\big(\{x \in M \mid \omega_\varepsilon(x) \neq \omega(x)\}\big) < \varepsilon$. Let us define $\alpha_\varepsilon(x) = -H(x, w_\varepsilon(x))$. By the continuity of H we have that α_ε are continuous functions on M and they converge pointwise $(\sigma + \tilde{\sigma}^s)$-a.e. to $\alpha(x) := -H(x,\omega(x))$. Moreover, $(\alpha_\varepsilon, \omega_\varepsilon)$ is admissible for (3.5). Thus we get

$$\begin{aligned}
\mathcal{L}(\sigma,T) \geq \int_M \alpha_\varepsilon(x)\, d\sigma + \int_M \langle \omega_\varepsilon(x), X(x) \rangle\, d\tilde{\sigma} + \int_{\{\omega_\varepsilon \neq \omega\}} \alpha_\varepsilon(x)\, d\sigma - \\
- \int_{\{\omega_\varepsilon = \omega\}} H(x, w(x))\, d\sigma + \int_M \langle \omega_\varepsilon(x), X(x) \rangle\, d\tilde{\sigma} \geq \\
\geq -\|\alpha_\varepsilon\|_\infty \sigma(\{\omega_\varepsilon \neq \omega\}) - \int_{\{\omega_\varepsilon = \omega\}} H(x,0)\, d\sigma + \int_M \langle \omega_\varepsilon(x), X(x) \rangle\, d\tilde{\sigma} \geq \\
\geq -C\varepsilon + \int_M \langle \omega_\varepsilon(x), X(x) \rangle\, d\tilde{\sigma}
\end{aligned} \tag{3.7}$$

where $C := \sup_{M \times \{|\omega| \leq 1\}} |H(x,\omega)|$ and we have used the natural bounds $\|\alpha_\varepsilon\|_\infty \leq C$. Passing to the limit as $\varepsilon \to 0^+$ in (3.7), taking into account that

$$\lim_{\varepsilon \to 0^+} \int_M \langle \omega_\varepsilon(x), X(x) \rangle\, d\tilde{\sigma} = \int_M \langle \omega(x), X(x) \rangle\, d\tilde{\sigma}\,,$$

we get

$$\begin{aligned}
\mathcal{L}(\sigma,T) \quad &\geq \quad \int_M \langle \tilde{\omega}(x)\chi_A(x), X(x) \rangle\, d\tilde{\sigma} \\
&= \quad \int_A \langle \tilde{\omega}(x), X(x) \rangle\, d\tilde{\sigma}^s = M(T)|\tilde{\sigma}^s|.
\end{aligned} \tag{3.8}$$

Now it suffices to notice that for any positive number $K \in \mathbb{R}$, one can repeat the argument above replacing $\omega(x)$ with $K\omega(x)$, that is, considering $(\alpha^K_\varepsilon, K\omega\varepsilon)$ as a test couple in (3.5). The previous estimates still hold and (3.7) reads now

$$\begin{aligned}
\mathcal{L}(\sigma,T) \quad &\geq \quad \int_M \langle \tilde{\omega}(x)\chi_A(x), X(x) \rangle\, d\tilde{\sigma} \\
&= \quad -C_K\varepsilon + \int_M \langle K\omega_\varepsilon(x), X(x) \rangle\, d\tilde{\sigma}
\end{aligned}$$

where $C_K := \sup_{M\times\{|\omega|\le K\}} |H(x,\omega)|$. Passing to the limit as $\varepsilon \to 0^+$ we get

$$\mathcal{L}(\sigma,T) \ge \int_A \langle K\tilde\omega(x), X(x)\rangle \, d\tilde\sigma^s = KM(T)|\tilde\sigma^s|.$$

Since $|\tilde\sigma^s| > 0$, by letting $K \to +\infty$ we get a contradiction. □

Remark 3.1.2. In the definition of $\mathcal{L}(\sigma,T)$ we can require that the constraint is satisfied with equality rather than inequality i.e., taking the supremum on pairs $(\alpha,\omega) \in C(M)\times\Gamma^\infty(T^*M)$ such that $\alpha(x) + H(x,\omega(x)) = 0$ pointwise. Thus $\mathcal{L}(\sigma,T)$ can be rewritten also as

$$\mathcal{L}(\sigma,T) = \sup_\omega \int_M (-H(x,\omega(x)))\, d\sigma + T(\omega)$$

where the sup is taken for $\omega \in \Gamma^\infty(T^*M)$. Moreover, if $\mathcal{L}(\sigma,T)$ is finite, then we can enlarge the class of admissible ω as follows. Let

$$\Gamma_\sigma := \{\omega \text{ forms with coefficients in } L^\infty_\sigma\}$$

then

$$\mathcal{L}(\sigma,T) = \sup_{\omega\in\Gamma_\sigma} \int_M (-H(x,\omega(x)))\, d\sigma + T(\omega). \tag{3.9}$$

This fact can be easily proved following the same procedure of Lemma 3.1.1. Indeed, given $\omega \in \Gamma_\sigma$ by Lusin's theorem we can find $\omega_\varepsilon \in \Gamma^\infty(T^*M)$ such that $\omega_\varepsilon \to \omega$ in L^∞_σ. Then

$$\mathcal{L}(\sigma,T) \ge \int_M (-H(x,\omega_\varepsilon(x)) + \langle\omega_\varepsilon(x), X(x)\rangle)\, d\sigma \tag{3.10}$$

where we have taken $X(x)$ as in the thesis of Lemma 3.1.1. It suffices then to pass to the limit as $\varepsilon \to 0$ and notice that $H(x,\omega_\varepsilon(x)) \to H(x,\omega(x))$ in $L^1(\sigma)$ and $\lim_{\varepsilon\to 0}\int_M\langle\omega_\varepsilon, X\rangle\, d\sigma = \int_M \langle\omega, X\rangle\, d\sigma$ to get (3.9), the inverse inequality being obvious.

Lemma 3.1.3. *If $\mathcal{L}(\sigma,T)$ is finite then*

$$\mathcal{L}(\sigma,T) = \int_M L(x,X(x))d\sigma,$$

where X is the vector field such that $T = \sigma\wedge X$ as in Lemma 3.1.1.

Proof. Let us fix (σ,T) such that $\mathcal{L}(\sigma,T) < +\infty$, and let $X(x) \in L^1_\sigma$ be such that $T = \sigma\wedge X$ as in Lemma 3.1.1. By Young inequality (2.2) we have

$$L(x,v) + H(x,\omega) \ge \langle\omega, v\rangle \qquad \forall x\in M,\ \forall v\in T_xM,\ \forall\omega\in T^*_xM \tag{3.11}$$

with equality assumed in the case of $(x,\omega) = \mathcal{L}(x,v)$, where $\mathcal{L}$ is the Legendre transform. Thus for any couple (α,ω) such that $\alpha(x) + H(x,\omega(x)) = 0$ we have

$$\int_M (\alpha(x) + \langle\omega, X(x)\rangle)\, d\sigma = \int_M (-H(x,\omega(x)) + \langle\omega, X(x)\rangle)\, d\sigma \le \int_M L(x,X(x))d\sigma. \tag{3.12}$$

Taking supremum w.r.t. (α,ω) we deduce

$$\mathcal{L}(\sigma,T) \le \int_M L(x,X(x))d\sigma.$$

Let now $\omega(x)$ be such that $(x,\omega(x)) = \mathcal{L}(x,X(x))$ and let $\omega_k(x) = \omega(x)$ if $|\omega(x)| \le k$, $\omega_k(x) = 0$ otherwise. Up to fix a representant for X in L^1_σ we can think ω as defined for every $x \in M$ so that $\{x \in M : |\omega(x)| < \infty\}$ coincides with M (in general it will be M minus a set σ-negligible). By the regularity of $\mathcal{L}$ it is easy to prove that ω is Borel so, in particular, $\omega_K \in L^\infty_\sigma$. Moreover, defined $A_K := \{x \in M : |\omega(x)| \leq K\}$ we have that A_K converges to 0 in measure (or equivalently $\chi_{A_K} \to 1$ in L^1_σ).

Taking into account Remark 3.1.2 we can take ω_K as admissible form and compute

$$\begin{aligned}
\mathcal{L}(\sigma, T) &\geq \int_M (-H(x, \omega_K(x)) + \langle \omega_K, X(x) \rangle)\, d\sigma = \\
&= \int_{A_k} (-H(x, \omega(x)) + \langle \omega, X(x) \rangle)\, d\sigma + \int_{A_k^c} (-H(x, \omega_K(x)) + \langle \omega_K, X \rangle)\, d\sigma = \\
&= \int_M L(x, X(x)) \chi_{A_K}(x)\, d\sigma + \int_M (-H(x,0))(1 - \chi_{A_K}(x))\, d\sigma.
\end{aligned}$$

Hence, since $L \geq 0$ and from the boundedness of $H(x,0)$, we can take the limit as $K \to +\infty$ and get

$$\mathcal{L}(\sigma, T) \geq \int_M L(x, X(x))\, d\sigma. \tag{3.13}$$

□

To each measure $\mu \in \mathcal{M}^c(TM)$ we can apply the disintegration Theorem 1.4.3 and write $\mu = \sigma \otimes \nu_x$ for a suitable family of probability measures ν_x concentrated on the fiber T_xM and $\sigma = \pi_\sharp \mu$ where $\pi : TM \to M$ is the projection $(x, v) \mapsto x$. Moreover, as seen in the introduction, it is possible to associate to μ a closed normal 1-current $p(\mu)$ by the formula

$$p(\mu)(\omega) := \int_{TM} \langle \omega(x), v \rangle d\mu(x, v) = \int_M \int_{T_xM} \langle \omega(x), v \rangle d\nu_x(v) d\sigma(x). \tag{3.14}$$

Then, thanks to Lemmas 3.1.1 and 3.1.3, using the notation: $M(x) = \int_{T_xM} v d\nu_x$, we have:

$$\mathcal{L}(\sigma, p(\mu)) = \int_M L(x, M(x)) d\sigma. \tag{3.15}$$

To each current T in $\mathcal{N}^c_1(M)$ we can associate its homology class $[T]$. Once an homology class $[h]$ is fixed it makes sense to study the following problem:

$$\mathcal{P} := \min \left\{ \mathcal{L}(\sigma, T) \ : \ \sigma \in \mathcal{P}(M), \ T \in N^c_1(M), \ [T] = [h] \right\}. \tag{3.16}$$

Notice that by Lemma 3.1.3 for each homology class the minimum in problem $\mathcal{P}$ is finite.

Proposition 3.1.4. *For each homology class $[h]$ there exists at least a minimizer for $\mathcal{P}$.*

Proof. Let (σ_n, T_n) a minimizing sequence. We can assume that $\mathcal{L}(\sigma_n, T_n) \leq C$. By Lemma 3.1.1 $T_n = \sigma_n \wedge X_n$, with $X_n \in L^1_{\sigma_n}$. We may also assume that $\sigma_n \stackrel{*}{\rightharpoonup} \sigma \in \mathcal{P}(M)$. Thanks to Remark 3.1.2 we can take a form $\omega_n(x)$ such that $\langle \omega_n(x), X_n(x) \rangle = |X_n(x)|$, with $\|\omega_n\|_\infty \leq 1$. Observe that the functions $\alpha_n(x) = -H(x, \omega_n(x))$ are equibounded. We have

$$\int_M \alpha_n(x) d\sigma_n + \int_M |X_n(x)| d\sigma_n = \int_M \alpha_n(x) d\sigma_n + T_n(\omega_n) \leq \mathcal{L}(\sigma_n, T_n) \leq C.$$

Hence $\int_M |X_n(x)| d\sigma_n \leq C$. For any form ω such that $\|\omega\|_\infty \leq 1$, we can evaluate:

$$T_n(\omega) = \int_M \langle \omega(x), X_n(x) \rangle d\sigma_n \leq \int_M |X_n(x)| d\sigma_n \leq C.$$

Then $M(T_n) \leq C$. Compactness of 1-currents ensures that $\exists T \in \mathcal{N}_1(M)$ such that $T_n \rightharpoonup T$. It is now immediate to check that T is closed and $[T] = [h]$. If a pair (α, ω) is admissible for $\mathcal{P}$ we have:

$$\int_M \alpha(x) d\sigma + T(\omega) = \lim_{n \to +\infty} \int_M \alpha(x) d\sigma_n + T_n(\omega) \leq \lim_{n \to +\infty} \mathcal{L}(\sigma_n, T_n).$$

Taking the supremum with respect to (α, ω) we obtain $\mathcal{L}(\sigma, T) \leq \lim_{n \to +\infty} \mathcal{L}(\sigma_n, T_n)$, i.e. (σ, T) is optimal. □

For any pair (σ, T) such that $\mathcal{L}(\sigma, T) < \infty$ let us pone

$$i(\sigma, T) = (id \times X)_\sharp \sigma,$$

where $T = \sigma \wedge X$ as in Lemma 3.1.1. Then the following theorem holds.

Theorem 3.1.5. *For every homology class $[h]$ Mather's problem $\beta([h])$ defined in (2.19) and problem $\mathcal{P}$ are equivalent in the sense that:*

1. *$\beta([h]) = \mathcal{P}$,*
2. *for each (σ, T) such that $\mathcal{L}(\sigma, T) < \infty$ it results: $p(i(\sigma, T)) = T$ and $[T] = [i(\sigma, T)]$. Moreover $\mathcal{A}(i(\sigma, T)) = \mathcal{L}(\sigma, T)$.*

Proof. Taking into account (3.15), the Jensen inequality implies the following

$$\begin{aligned} \mathcal{A}(\mu) = \int_{TM} L(x, v) d\mu = \int_M \int_{T_x M} L(x, v) d\nu_x(v) d\sigma \geq \\ \geq \int_M L(x, \int_{T_x M} v d\nu_x(v)) d\sigma = \mathcal{L}(\sigma, p(\mu)). \end{aligned} \tag{3.17}$$

If μ is a minimal measure, then

$$\mathcal{A}(\mu) \geq \mathcal{L}(\sigma, p(\mu)) = \mathcal{A}(i(p(\mu))) \geq \mathcal{A}(\mu)$$

where the last inequality depends on the minimality of μ. Therefore point 1 is proved. A direct computation gives point 2. □

Going back to inequality (3.17), actually for a minimal measures μ (still denoting by $\sigma \otimes \nu_x$ the disintegration of μ) we have that for σ-a.e. x holds the equality

$$\int_{T_x} L(x, v) d\nu_x = L(x, \int_{T_x} v d\nu_x),$$

and this last equality implies $\nu_x = \delta_{\varphi(x)}$. This is an easy consequence of the convexity, for sake of completeness we report a proof of this fact in Lemma 3.1.7.

Remark 3.1.6. A consequence of the previous theorem is that each minimal measure is supported on a graph.

Lemma 3.1.7. *$g : \mathbb{R}^N \to \mathbb{R}^+$ strictly convex, $\nu \in \mathcal{P}_1(\mathbb{R}^N)$. If*

$$\int_{\mathbb{R}^N} g(x) d\nu = g(\int_{\mathbb{R}^N} x d\nu)$$

then $\nu = \delta_{x_0}$ for some x_0.

Proof. Fix the value $K = \int_{\mathbb{R}^N} g(x)d\nu$ and consider the sets: $A := \{x \in \mathbb{R}^N \ : \ g(x) > K\}$ and $B := \{x \in \mathbb{R}^N \ : \ g(x) \leq K\}$ If $\nu(A) = 0$ then ν is concentrated on the set $\{g(x) = K\}$ and then by the strict convexity of g either $\nu = \delta_{x_0}$ for some x_0 or $\int_{\mathbb{R}^N} xd\nu \in \{x \in \mathbb{R}^N \ : \ g(x) < K\}$ which concludes the proof. If both A and B are both of positive measure then we can consider on them the measures $\frac{\nu \llcorner A}{\nu(A)}$ and $\frac{\nu \llcorner B}{\nu(B)}$. The following facts are true: $\nu(A) + \nu(B) = 1$, $0 < \nu(A), \nu(B) < 1$, $\int g \frac{d\nu \llcorner A}{\nu(A)} > K$, B is convex. Then we have the following:

$$
\begin{aligned}
g(\int_{\mathbb{R}^N} xd\nu) = g\left(\nu(A)\int_A x\frac{d\nu\llcorner A}{\nu(A)} + \nu(B)\int_B x\frac{d\nu\llcorner B}{\nu(B)}\right) \leq \\
\leq \nu(A)g\left(\int_A x\frac{d\nu\llcorner A}{\nu(A)}\right) + \nu(B)g\left(\int_B x\frac{d\nu\llcorner B}{\nu(B)}\right) \leq \\
\leq \int_A g(x)d\nu + \int_B g(x)d\nu = \int_{\mathbb{R}^N} g(x)d\nu.
\end{aligned}
$$

The first inequality is the convexity inequality for g while the second is due to the application of the Jensen inequality to both integrals. Then the first inequality is strict for the strict convexity of g unless $\int_A x\frac{d\nu\llcorner A}{\nu(A)} = \int_B x\frac{d\nu\llcorner B}{\nu(B)}$, but by the convexity of B this would imply $\int_A x\frac{d\nu\llcorner A}{\nu(A)} \in B$ and then the second inequality since

$$
g\left(\int_A x\frac{d\nu\llcorner A}{\nu(A)}\right) \leq K < \int_A g(x)\frac{d\nu\llcorner A}{\nu(A)}.
$$

Then it cannot be that $\nu(A) > 0$ and then the conclusion follows as in the first part of the proof. □

We conclude this section with a proposition which connects the equivalence just proved with a classical geometric problem in the space of currents and to the results of [7].

For each $T \in \mathcal{N}_1^c(M)$ we denote by $\sigma_s \wedge X_s$ the special parametrization introduced by Theorem A.0.7 and we recall that σ_s is a probability measure such that $g(X_s, X_s) = M(T)^2$ σ_s-a.e.

Proposition 3.1.8. *Let $L(x,v) = g(v,v)^p$ for some $p > \frac{1}{2}$, let $T \in \mathcal{N}_1^c(M)$ and let $\sigma \wedge X$ be any parametrization of T then*

$$
\mathcal{L}(\sigma, T) \geq \mathcal{L}(\sigma_s, T) = M(T)^{2p}.
$$

Proof. By Lemma A.0.8 we have $\sigma_s = f(x)\sigma$ and $X = f(x)X_s$ for a positive function $f \in L^1_\sigma$ such that $\int_M f(x)d\sigma = 1$. Let us pone $c(x) = 0$ if $f(x) = 0$ and $c(x) = \frac{1}{f(x)}$ otherwise. Then

$$
\begin{aligned}
\mathcal{L}(\sigma, T) = \int_M g(X,X)^p d\sigma = \int_M f(x)^{2p} g(X_s, X_s)^p d\sigma = \\
= M(T)^{2p}\int_M c(x)^{1-2p} d\sigma_s \geq M(T)^{2p}\left(\int_M c(x)d\sigma_s\right)^{1-2p} = M(T)^{2p},
\end{aligned}
$$

where the last inequality follows from the Jensen inequality and the convexity of the function $y \mapsto y^{1-2p}$. □

The previous result in the case $p = 1$ was proved by Bangert in [7]. The interpretation of Proposition 3.1.8 is that whenever $L(x,v) = g(v,v)^p$ for some $p > \frac{1}{2}$ our problem is equivalent to the classical problem of minimizing the mass in a given homology class. As one expects the minimizers do not depend on p and the minimal value is homogeneous with respect to p.

Remark 3.1.9. The probability measure σ in a minimal pair (σ, T) has a counterpart in the theory of optimal transport which is the so called transport density, see (1.29). The transport density is a very useful tool in the optimal transport theory and in some of its applications ([11, 12, 13, 34]).

The weak KAM theory of A. Fathi [26, 42, 43], suggests that for a minimal pair (σ, T) the supremum in remark 3.1.2 is achieved at a pair $\alpha(x) \equiv c$ and $\omega(x) = du(x)$ where u is a viscosity solution of the equation $H(x, du(x)) = c$. Let us discuss the case of some special Hamiltonians.

First observe that if we denote by $\omega(x)$ the Legendre transform of $X(x)$, then the following equality is satisfied

$$\mathcal{L}(\sigma, T) = \int_M [-H(x, \omega(x)) + \langle \omega(x), X(x) \rangle] d\sigma,$$

since the supremum in (3.9) is in fact attained at the pair $(-H(x, \omega), \omega)$. Since $X(x) \in L^1_\sigma$, in general we cannot say too much about the regularity of $\omega(x)$.

Thus, assuming that (α, ω) is a maximizing pair in (3.5) we have

$$\mathcal{L}(\sigma, T) = \int_M L(x, X) d\sigma = \int_M \alpha(x) d\sigma + \int_M \langle \omega, X \rangle d\sigma \leq$$

$$\leq \int_M [-H(x, \omega) + \langle \omega, X \rangle] d\sigma \leq \int_M L(x, X) d\sigma.$$

Then, in fact, equality holds in all the previous estimates. Thus, it follows that

$$\alpha(x) + H(x, \omega(x)) = 0 \quad \sigma - a.e.$$

As we have seen in in section 2.3, the weak KAM theory deals with weak solutions of the equation

$$c_0 + H(x, du(x)) = 0.$$

Let us consider the Mañè critical value with the opposite sign (see section 2.4)

$$c_0 = \min \left\{ \int_M L(x, v) d\mu \mid \mu \in \mathcal{M}^c \right\}. \tag{3.18}$$

The minimum in (3.18) will be attained in some homology class $[\mu]$. By the existence result stated in Proposition 3.1.4, $c_0 = \int_M L(x, X) d\sigma$ for a suitable optimal pairs $(\sigma, \sigma \wedge X)$ in such homology class. We will now use this optimal pair to prove a well known formula in a particular case.

Lemma 3.1.10. *For symmetric Lagrangians $L(x, v) = L(x, -v)$ one has:*

$$c_0 = \sup_{u \in \mathcal{C}^1} \min_{x \in M} -H(x, du).$$

Proof. Let $u \in \mathcal{C}^1$. We have:

$$c_0 = \int_M L(x.X) d\sigma \geq \int_M \langle du, X \rangle d\sigma - \int_M H(x, du) d\sigma =$$

$$= \int_M -H(x, du) d\sigma \geq \min_{x \in M} -H(x, du),$$

where the second equality depends on the fact that $\sigma \wedge X$ is closed. Taking the supremum we obtain

$$c_0 \geq \sup_{u \in \mathcal{C}^1} \min_{x \in M} -H(x, du).$$

Vice versa, consider a point $x_0 \in M$ such that $L(x_0, 0) = \min_{x \in M} L(x, 0)$, and the closed measure $\delta_{(x_0, 0)}$ on TM. Observe that also the Hamiltonian H is strictly convex, superlinear and symmetric. Then

$$\forall (x, \omega) \in T^*M : H(x, 0) \leq H(x, \omega).$$

So we can evaluate

$$c_0 \le \int_{TM} L(x,v)d\delta_{(x_0,0)} = L(x_0,0) \le L(x,0) = -H(x,0) = -H(x,dc),$$

where c denotes any constant function on the manifold M. But this holds for every $x \in M$, hence

$$c_0 \le \min_{x\in M} -H(x,dc) \le \sup_{u\in\mathcal{C}^1} \min_{x\in M} -H(x,du).$$

$\square$

Remark 3.1.11. This min-max formula is well known (see [26, 52]) and also holds for non-symmetric Lagrangians. We have chosen symmetric Lagrangians in order to have a simpler proof. Observe that in the symmetric case the above supremum in the min-max formula is in fact attained, at least on constant functions.

Let u be a function such that $c_0 = min_{x\in M} - H(x,du)$. Then the pair (c_0, du) is admissible for $\mathcal{L}(\sigma,T)$. But we have also

$$0 = \int_M \langle du, X\rangle d\sigma \le \int_M [L(x,X) + H(x,du)]d\sigma = \int_M [c_0 + H(x,du)]d\sigma \le 0,$$

where the first equality follows from the fact that $T = \sigma \wedge X$ is closed. Hence

$$c_0 + H(x,du) = 0 \quad \sigma - a.e.$$

The above equation is true for every minimizing measure μ in (3.18), so if we consider $\mathcal{M}_0 \subset M$ as the projected Mather's set $\pi(\overline{\bigcup_\mu \mathrm{spt}(\mu)})$ (see [26, 43]) we can state the following:

Theorem 3.1.12. *For symmetric Lagrangians the minimal measures are concentrated in the set of points where the function $x \mapsto H(x,0)$ achieves its maximum.*

Remark 3.1.13. In particular, considering Lagrangians of the type $\frac{1}{2}|v|^2 - V(x)$, the previous theorem says that the minimal measures concentrate on the set of maximum points of the potential $-V$. Moreover, this result explains in a particular case the duality between the weak KAM theory and the Mather theory, as discussed at the beginning of this chapter in the case of the torus. For another duality approach see [51].

3.2 Transport equation and Eulerian representations

In this section we show that Mather's problem is equivalent to two different minimization problems on some space of curves on $\mathcal{P}(M)$ which solve a continuity equation.

The first problem we consider is the following: minimize the convex and l.s.c. functional

$$\mathcal{K}(\rho,v) = \int_0^1 \int_M L(x,v(x,t))d\rho_t dt, \tag{3.19}$$

among all pairs $\rho : [0,1] \to \mathcal{P}(M)$ and $v : M \times [0,1] \to TM$ which solve

$$\begin{cases} \dfrac{\partial\rho}{\partial t} + \nabla\cdot(v\rho) = 0 \quad \text{in} \quad (0,1)\times M, \\ \\ \rho_0 = \rho_1 \end{cases} \tag{3.20}$$

and satisfy an homological constraint in a sense that will be explained right below. Observe that the vectorial measure $v(x,t)\rho_t$ corresponds to the measure $E_t \in \mathcal{M}^n(M)$ of Brenier problem as discussed in section 1.5. To each solution (ρ, v) of (3.20) we can associate a closed, normal 1-current

$$T_{\rho,v} : \omega \mapsto \int_0^1 \int_M \langle \omega(x), v(x,t)\rangle d\rho_t dt.$$

By (1.35) it follows that $T_{\rho,v}$ is closed. Indeed, using (1.35) we have

$$T_{\rho,v}(df) = \int_0^1 \int_M \langle df(x), v(x,t)\rangle d\rho_t dt = \int_0^1 \frac{d}{dt}[\int_M f d\rho_t] dt = \langle \rho_1, f\rangle - \langle \rho_0, f\rangle = 0.$$

Therefore $T_{\rho,v}$ has a well defined homology class. Then the problem is formulated as follows

$$\mathcal{B} := \min \{\mathcal{K}(\rho, v) \mid (3.20) \text{ holds, and } [T_{\rho,v}] = h\}. \tag{3.21}$$

Each pair (ρ, v) will be called *parametrization* of the current $T_{\rho,v}$. Notice that different parametrizations can be associated to the same current. In the following we will shortly write v_t to denote the vector field $v(\cdot, t)$.

To each pair (ρ, v) solution of (3.20) we can associate a pair (σ, T) admissible for problem $\mathcal{P}$ by the map

$$(\rho, v) \xrightarrow{\overline{c}} (\int_0^1 \rho_t dt, T_{\rho,v}).$$

Vice versa, for each pair (σ, T) which is admissible for problem $\mathcal{P}$ we define the map $\underline{c}(\sigma, T) = (\rho, v)$ where $\rho_t = \sigma$ and $v(\cdot, t) = X$ for all t. It is immediate to check that $\underline{c}(\sigma, T) = (\rho, v)$ is solution of (3.20).

Theorem 3.2.1. *Problem $\mathcal{B}$ defined in (3.21) and problem $\mathcal{P}$ are equivalent in the sense that $\mathcal{P} = \mathcal{B}$.*

Moreover the maps $\overline{c}$ and $\underline{c}$ defined above associate minimizers of one problem to minimizers of the other.

Proof. If (σ, T) is such that $\mathcal{L}(\sigma, T) < \infty$ then $\rho_t \equiv \sigma$ and $v_t \equiv X$ is a solution of (3.20) and

$$\mathcal{K}(\underline{c}(\sigma, T)) = \mathcal{K}(\rho, v) = \int_0^1 \int_M L(x, X(x)) d\sigma dt = \mathcal{L}(\sigma, T). \tag{3.22}$$

Moreover $T_{\rho,v}$ and T are in the same homology class. Hence $\mathcal{B} \le \mathcal{P}$.

On the other hand, let (ρ, v) be a solution of (3.20) such that $\mathcal{K}(\rho, v) < \infty$ then $\overline{c}(\rho, v)$ is such that $\mathcal{L}(\overline{c}(\rho, v)) < \infty$. Let us define $E_t = v_t \rho_t$. By the inequality

$$|\int_0^1 E_t(B) dt| \le \int_0^1 \int_B |v_t| d\rho_t dt,$$

We deduce that $\int_0^1 E_t dt \ll \sigma = \int_0^1 \rho_t dt$.

As already done in Lemma 3.1.3, it is possible to check that

$$\mathcal{K}(\rho, v) = \mathcal{K}(\rho, E) := \sup\left\{\int_0^1 \int_M \alpha(t,x) d\rho_t dt + \int_0^1 \int_M \omega(t,x) \cdot dE_t dt, \right\} \tag{3.23}$$

where the supremum runs among continuous functions $\alpha(t,x), \omega(t,x)$ which satisfy the constraint $\alpha(t,x) + H(x, \omega(t,x)) \le 0$ (see pag. 93 of [19]). Let us now compare the values of $\mathcal{K}$ and $\mathcal{L} \circ \overline{c}$.

$$\mathcal{L}(\sigma, T_{\rho,v}) = \sup_{\alpha(x)+H(x,\omega(x))\le 0} \left\{\int_M \alpha(x)\, d\sigma + T_{\rho,v}(\omega)\right\} =$$

$$= \sup_{\alpha(x)+H(x,\omega(x))\le 0} \left\{\int_0^1 \int_M \alpha(x) d\rho_t dt + \int_0^1 \int_M \omega(x) \cdot dE_t dt\right\} \le$$

$$\le \sup_{\alpha(t,x)+H(x,\omega(t,x))\le 0} \left\{\int_0^1 \int_M \alpha(t,x) d\rho_t dt + \int_0^1 \int_M \omega(t,x) \cdot dE_t dt\right\} = \mathcal{K}(\rho, E).$$

Therefore $\mathcal{P} \le \mathcal{B}$. The second part of the statement follows directly from the equality (3.22). $\square$

Remark 3.2.2. By the strict convexity of the Lagrangian L, once we fix the curve ρ_t there exists essentially only one vector field v_t such that (ρ, v) is solution of (3.20) and $\mathcal{K}(\rho, v)$ is minimal.

It may happen that for some currents $T \in \mathcal{N}_1^c(M)$ the only periodic parametrization is the parametrization constant with respect to t. We now prove the equivalence of problem $\mathcal{P}$ with another problem $\mathcal{B}_1$ defined below, which includes problem $\mathcal{B}$ in a suitable sense and guarantees the existence of non trivial minimal parametrizations for all minimal measures.

Here we consider the pairs $\rho : \mathbb{R} \to \mathcal{P}(M)$ and $v : M \times \mathbb{R} \to TM$ which solves

$$\frac{\partial \rho}{\partial t} + \nabla \cdot v\rho = 0 \quad on\ \mathbb{R} \times M, \tag{3.24}$$

and such that the sequence of currents T_k given by

$$T_k(\omega) = \frac{1}{2k} \int_{-k}^{k} \int_M \langle \omega(x), v(x,t) \rangle d\rho_t dt \tag{3.25}$$

converges toward a closed, normal 1-current $T_{\rho,v}$. If (ρ, v) is a periodic solution of (3.20) then the limit of the sequence in (3.25) coincides with $T_{\rho,v}$ already defined so that there is no abuse of notation here. As already noticed $T_{\rho,v}$ has a well defined homology class. The cost to be minimized will be given by

$$\mathcal{K}_1(\rho, v) = \liminf_{k \to \infty} \frac{1}{2k} \int_{-k}^{k} \int_M L(x, v(x,t)) d\rho_t dt. \tag{3.26}$$

Then the new problem is written as

$$\mathcal{B}_1 := \min \left\{ \mathcal{K}_1(\rho, v) \mid (3.24) \text{ holds, and } [T_{\rho,v}] = h \right\}. \tag{3.27}$$

The equivalence between problems $\mathcal{P}$ and $\mathcal{B}_1$ can be proved exactly as in Theorem 3.2.1 and then we will skip the proof. We will focus instead on an explicit example of non trivial minimal parametrizations for each minimal current T. The example we present shows also some relation between the PDE (3.20) and another formulation of the Mather's problem presented in [7] in which Mather's minimal measures are described in terms of probability measures on the space of Lipschitz curves. The results of [7] are derived for $L(x, v) = g(v, v)$, but the results we need also hold for Lagrangians of the form $L(x, v) = \psi(g(v, v))$. The following theorem provides non-trivial parametrizations of minimal measures composed by elementary solenoids using a decomposition theorem for normal 1-currents due to Smirnov. For the statement and precise definitions we refer to the appendix A.

Theorem 3.2.3. *Assume that the Lagrangian L is of the form $L(x,v) = \psi(g(v,v))$ (and still satisfies our standard assumptions). Then for each minimal measure μ there exists a non-trivial minimal parametrization* composed *of elementary solenoids.*

Proof. Let μ be a minimal measure for Mather's problem $\beta([h])$ and let T be the associated minimal current. Let ν be the probability measure on the space of elementary solenoids $\mathcal{S}$ associated to the Smirnov decomposition of T given by Theorem A.0.13. It is possible to check that this measure ν is also minimal in the sense of [7]. Moreover each solenoid is identified by a curve γ, and by Theorem 3.11 of [7] two curves in the support of ν never intersect. Then we define: $\rho_t = \int_{\mathcal{S}} \delta_{\gamma(t)} d\nu(\gamma)$ and $v(x,t) = \dot{\gamma}(t)$ if $\gamma(t) = x$ for some $\gamma \in \operatorname{spt}(\nu)$ and $v(x,t) = 0$ otherwise. It is immediate to check that (ρ, v) is a solution of (3.24), moreover

$$\lim_{k \to \infty} \left(\frac{1}{2k} \int_{-k}^{k} \rho_t dt, T_k \right) = (\sigma_s, T), \tag{3.28}$$

where σ_s is the probability measure associated to T by point 3 of Theorem A.0.7. Indeed

$$\sigma_s(B) = \frac{1}{M(T)} \sup \left\{ T(\omega) \mid \omega \in \Gamma^\infty(U),\ \|\omega\|_\infty \leq 1, spt(\omega) \subset B \right\} =$$

$$= \frac{1}{M(T)} \sup \left\{ \int_{\mathcal{S}} \lim_{k\to\infty} \frac{1}{2k} \int_{-k}^{k} \langle \omega(\gamma(t)), \dot\gamma(t) \rangle dt d\nu(\gamma), Im(\gamma) \subset B \right\} =$$

$$= \frac{1}{M(T)} \lim_{k\to\infty} \frac{1}{2k} \int_{\mathcal{S}\cap\{Im(\gamma)\subset B\}} \int_{-k}^{k} |\dot\gamma(t)| dt d\nu(\gamma) =$$

$$= \lim_{k\to\infty} \frac{1}{2k} \int_{\mathcal{S}} \int_{-k}^{k} \delta_\gamma(t)(B) dt d\nu(\gamma) = \lim_{k\to\infty} \frac{1}{2k} \int_{-k}^{k} \rho_t(B) dt.$$

By Proposition 3.1.8 the pair on the right hand side of (3.28) is minimal. □

Remark 3.2.4. Using a variation of the technique we used in the previous proof it is sometimes possible to build many explicit parametrizations of a minimal measure μ. We will not go further in details on this but much will be clear from the following examples.

Some minimal parametrizations on $\mathbb{T}^2$

In this subsection we work out in detail an example of explicit minimal measure on the flat 2-torus.

This example will be better described with the help of the figure 3.1. By $\{\gamma_i\}_{i=1,\dots,4}$ we denote both the curves and their supports. The distinction will be clear from the context.

Let $\{\gamma_i\}_{i=1,\dots,4}$ be as in Figure 3.1.

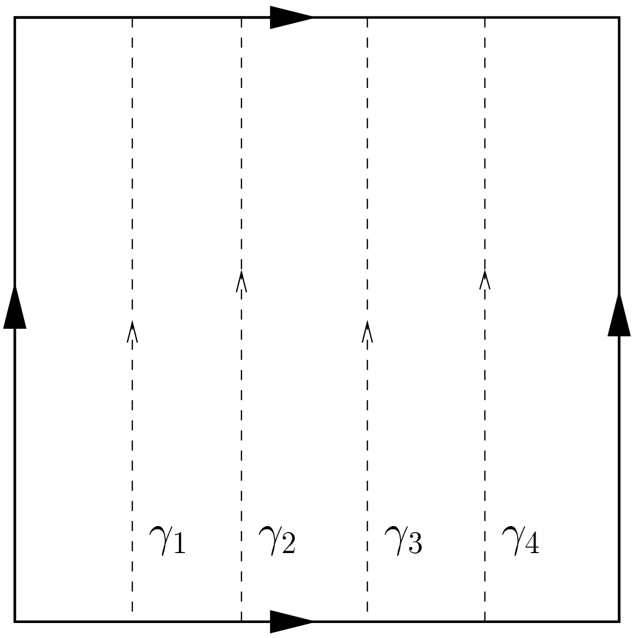


Figure 3.1: An example of path on the torus.

Consider $\mu \in \mathcal{P}(T\mathbb{T}^2)$ defined by

$$\mu = \frac{1}{4}((\mathcal{H}^1 \llcorner \gamma_1) \otimes \delta_{e_2} + (\mathcal{H}^1 \llcorner \gamma_2) \otimes \delta_{e_2} + (\mathcal{H}^1 \llcorner \gamma_3) \otimes \delta_{e_2} + (\mathcal{H}^1 \llcorner \gamma_4) \otimes \delta_{e_2}).$$

The measure μ is invariant for the geodesic flow, the homology class of μ is $[\mu] = (0,1) \in \mathbb{R} \times \mathbb{R} = H_1(\mathbb{T}^2)$, and μ is minimal being its cost equal to the length of the minimal geodesic in this class.

The following are minimal parametrizations of μ:

1. $\rho_1(t) = \frac{1}{4}(\delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)} + \delta_{\gamma_4(t)})$, $\quad v_1(x,t) = e_2$,
2. $\rho_2(t) = \frac{1}{4}(\delta_{\gamma_1(t)} + \delta_{\gamma_2(t)} + \delta_{\gamma_3(t)}) + \frac{1}{4}\mathcal{H}^1 \llcorner \gamma_4$, $\quad v_2(x,t) = e_2$,
3. $\rho_3(t) = \frac{1}{4}(\delta_{\gamma_1(t)} + \delta_{\gamma_2(t)}) + \frac{1}{4}(\mathcal{H}^1 \llcorner \gamma_3 + \mathcal{H}^1 \llcorner \gamma_4)$, $\quad v_3(x,t) = e_2$,
4. $\rho_4(t) = \frac{1}{4}\delta_{\gamma_1(t)} + \frac{1}{4}(\mathcal{H}^1 \llcorner \gamma_2 + \mathcal{H}^1 \llcorner \gamma_3 + \mathcal{H}^1 \llcorner \gamma_4)$, $\quad v_4(x,t) = e_2$,
5. Constant parametrization,
6. Parametrizations obtained by commuting the index i in the previous examples,
7. Compositions of the previous examples with time shift.

Chapter 4

Action minimizing measures for optimal transportation problems

In chapter 3 we have derived a PDE formulation for Mather's problem. This formulation plays the same role of Brenier's formulation of the mass transportation problem. In this chapter we would like to formulate the Monge-Kantorovich problem in terms of minimal probability measures on the phase space. We start by investigating the problem:

$$\mathbf{M} := \min \left\{ \int_{TM} L(x,v) d\mu \mid \mu \in \mathcal{P}(TM),\ \partial p(\mu) = f^+ - f^- \right\}, \tag{4.1}$$

where $f^+, f^- \in \mathcal{P}(M)$, and $p(\mu)$ is the current associated to μ by (3.14). As in section 3.1 we can formulate the following problem on the space of currents:

$$\mathbb{P} := \min \left\{ \mathcal{L}(\sigma, T) \mid \sigma \in \mathcal{M}(M),\ T \in N_1(M),\ \partial T = f^+ - f^- \right\}. \tag{4.2}$$

We point out that in Lemma 3.1.1 and Lemma 3.1.3 the boundary terms do not play any role. Therefore we can state the following

Theorem 4.0.5. *$\boldsymbol{M} = \mathbb{P}$.*

Comparisons between M and MK

The PDE's formulation for the Monge-Kantorovich problem is: minimize the functional

$$K(\rho, E) = \int_0^1 \int_M L(x, v(t,x)) d\rho_t dt$$

among all maps $\rho : [0,1] \to \mathcal{P}(M)$ and $E : [0,1] \to \mathcal{M}(M, TM)$ which solve

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \nabla \cdot E = 0 & \text{on } (0,1) \times M, \\ \rho_0 = f^+, \quad \rho_1 = f^-. \end{cases} \tag{4.3}$$

Where (4.3) is satisfied in the sense of distributions in $(0,1) \times M$ and $v(t,x)$ is the Radon-Nikodym derivative of E_t with respect to ρ_t. We set

$$\mathbf{B} := \min\{K(\rho, E) \mid (\rho, E) \text{ solves } (4.3)\}. \tag{4.4}$$

As in Lemma 3.1.1 and Lemma 3.1.3, one can check that whenever $K(\rho, E) < +\infty$, then E is absolutely continuous with respect to ρ and, as discussed in the proof of Theorem 3.2.1, it results:

$$K(\rho, E) = \sup \left\{ \int_0^1 \int_M \alpha(t,x) d\rho_t dt + \int_0^1 \int_M \beta(t,x) dE_t dt \right\}.$$

Where the supremum is taken for $\alpha(t,x), \beta(t,x)$ continuous functions which satisfy the constraint $\alpha(t,x) + H(x, \beta(t,x)) \leq 0$. Consider the "'length cost"'

$$c(x,y) = \inf_{\substack{\gamma \in AC([0,1],M) \\ \gamma(0)=x, \gamma(1)=y}} \int_0^1 L(\gamma(t), \dot{\gamma}(t)) dt, \tag{4.5}$$

and denote with **MK** the Monge-Kantorovich minimum associated to the cost $c(x,y)$ above. Then the following theorem gives a first comparison.

Theorem 4.0.6. $\boldsymbol{M} \leq \boldsymbol{B} \leq \boldsymbol{MK}$.

Proof. Let η be a Tonelli minimizer for the cost $c(a,b)$, and γ an optimal plan for **MK**. Then we put:

$$\rho_t = \int_{M \times M} \delta_{x - \eta(t)} d\gamma(a,b) \ , \ E_t = \int_{M \times M} \dot{\eta}(t) \delta_{x-\eta(t)} d\gamma(a,b). \tag{4.6}$$

As in the proof of Theorem 1.5.5 we have that the couple (ρ, E) is admissible for **B**.

We can evaluate:

$$\mathbf{B} \leq \sup_{\alpha, \beta} \int_0^1 \int_M [\alpha(t,x) d\rho_t + \beta(t,x) dE_t] dt \leq$$

$$\leq \sup_{\beta} \int_0^1 \int_M [-H(x, \beta(t,x)) d\rho_t + \beta(t,x) dE_t] dt =$$

$$= \sup_{\beta} \int_{M \times M} \int_0^1 [-H(\eta(t), \beta(t, \eta(t))) dt d\gamma(a,b) + \langle \beta(t, \eta(t)), \dot{\eta}(t) \rangle dt d\gamma(a,b)] \leq$$

$$\leq \int_{M \times M} \int_0^1 L(\eta(t), \dot{\eta}(t)) dt d\gamma(a,b) = \int_{M \times M} c(a,b) d\gamma(a,b) = \mathbf{MK}.$$

For every (ρ, E) admissible for **B**, we consider the couple (σ, T) defined by

$$T(\omega) = \int_0^1 \int_M \omega dE_t dt \ , \ \sigma = \int_0^1 d\rho_t dt.$$

By (1.35) we have

$$\partial T(\varphi) = T(d\varphi) = \int_0^1 \int_M d\varphi(x) dE_t dt = \int_0^1 \frac{d}{dt} \int_M \varphi(x) d\rho_t dt =$$

$$= \langle \rho_1, \varphi \rangle - \langle \rho_0, \varphi \rangle = \langle f^+ - f^-, \varphi \rangle.$$

Hence (σ, T) is admissible for problem $\mathbb{P}$.

Then, by Theorem 4.0.5 and the inequality $\mathcal{L}(\sigma, T) \leq K(\rho, E)$, we obtain that $\mathbf{M} \leq \mathbf{B}$. ☐

We remark that the arguments used in section 3.2 to prove the full equivalence of the PDE formulation $\mathcal{B}$ of Mather's problem do not work in this contest. Let us denote by $\mathbf{M}_p$, $\mathbf{MK}_p$, $\mathbf{B}_p$, respectively problems (4.1), **MK**, (4.4) for the Lagrangian $L(x,v) = g(v,v)^p$. With the same proof of Proposition 3.1.8 we can state the following

Proposition 4.0.7. *For $L(x,v)=g(v,v)^p$ with $p>\frac{1}{2}$ it results*

$$\boldsymbol{M}_p=\min\left\{M(T)^{2p}|\ T\in\mathcal{N}_1\ ,\ \partial T=f^+-f^-\right\}.$$

A consequence of Proposition 4.0.7 is that $\mathbf{M}_p$ is homogeneous with respect to p, i.e. for every $\alpha>1$ it results $\mathbf{M}_{\alpha p}=(\mathbf{M}_p)^\alpha$. Let us show by a counterexample that this is not the case for $\mathbf{MK}_p$. It will follows that the strict inequality $\mathbf{M}_p<\mathbf{MK}_p$ holds for every $p>\frac{1}{2}$. Indeed, let T a minimal current for $\mathbf{M}_{1/2}$, the homogeneity of $\mathbf{M}_p$ with respect to p and the equality $\mathbf{MK}_p=\mathbf{M}_p$ would imply the following:

$$\mathbf{MK}_p=\mathbf{M}_p=M(T)^{2p}=\mathbf{MK}_{1/2}^{2p}.$$

However, this equality in general does not hold as we show by the following example.

Example 4.0.8. *Consider the measures $f^+=\frac{1}{2}\delta_x+\frac{1}{2}\delta_{x^*}$, $f^-=\frac{1}{2}\delta_y+\frac{1}{2}\delta_{y^*}$ in $\mathbb{R}^2$ where the points x,x^*,y,y^* are disposed as in the figure 4.1 and consider for example the case $p=1$.*

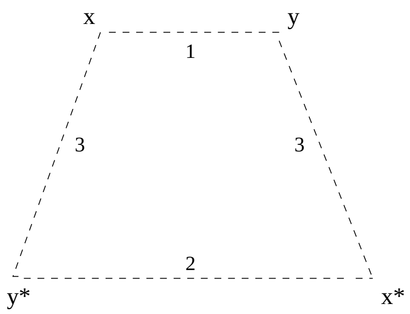


Figure 4.1: The integer numbers denote the distances between the points.

An optimal transport plan for both the distance and the distance to the power 2 is given by $\gamma=\frac{1}{2}\delta_{(x,y)}+\frac{1}{2}\delta_{(x^*,y^*)}$.

In this case we have

$$\boldsymbol{MK}_1=\int_{M\times M}d(x,y)^2d\gamma(x,y)=\frac{1}{2}d(x,y)^2+\frac{1}{2}d(x^*,y^*)^2=\frac{5}{2}.$$

On the other hand

$$\boldsymbol{MK}_{1/2}^2=\left(\int_{M\times M}d(x,y)d\gamma(x,y)\right)^2=\left(\frac{1}{2}d(x,y)+\frac{1}{2}d(x^*,y^*)\right)^2=\frac{9}{4}.$$

Hence $\boldsymbol{MK}_{1/2}^2<\boldsymbol{MK}_1$.

Let γ be an optimal plan between two probability measures f^+,f^- for the Monge-Kantorovich problem with the geodesic distance d on M as cost. Consider the following 1-current T_γ :

$$T_\gamma(\omega)=\int_{M\times M}\int_0^1\langle\omega(\sigma_{x,y}(t)),\dot\sigma_{x,y}(t)\rangle dtd\gamma(x,y).$$

Where $\sigma_{x,y}$ is a geodesic connecting the points x,y. Observe that

$$|T_\gamma(\omega)|\le\int_{M\times M}\int_0^1|\dot\sigma_{x,y}(t)|dtd\gamma(x,y)\|\omega\|_\infty=\int_{M\times M}d(x,y)d\gamma(x,y)\|\omega\|_\infty,$$

from which we obtain $M(T_\gamma) \leq \mathbf{MK}_{1/2}$. By Proposition 1.5.5 and Proposition 4.0.7, we have that T_γ is also optimal for $\mathbf{M}_p$ for every $p > \frac{1}{2}$. Furthermore, it results that

$$M(T_\gamma) = \int_{M\times M} d(x,y)d\gamma(x,y) = \mathbf{MK}_{1/2}.$$

Let γ_p be an optimal plan for $\mathbf{MK}_p$ with $p > \frac{1}{2}$ and let T_{γ_p} be defined as above. It can happen that the 1-current T_{γ_p} is not optimal for problem $\mathbf{M}_p$ in this case. Indeed, the minimality of T_{γ_p} also implies that $M(T_{\gamma_p}) = \mathbf{MK}_{1/2}$, and then γ_p would be an optimal plan also for $\mathbf{MK}_{1/2}$. However, it is known that $\mathbf{MK}_p$ admits a unique optimal plan induced by a transport map which in general does not minimize $\mathbf{MK}_{1/2}$ (see [48] for more details and explicit examples).

Refining M

In this section we prove that the Monge-Kantorovich problem is in fact equivalent to a refinement of problem $\mathbf{M}$ in which another constraint is added on the admissible measures. Actually, the constraint will define a strict subfamily of the admissible measures for $\mathbf{M}$. The result will be achieved in different steps.

Duality and Brenier's formulation

As a first step we establish another duality formula for $\mathbf{MK}$. We have found explicitly this formula in the quadratic case in [70], see also [86]. Furthermore, the strategy to establish this formula is delineated in [19] but in order to include general Lagrangian $L(x,v)$ some additional work is necessary.

Theorem 4.0.9. *The minimum of the Monge-Kantorovich problem* $\boldsymbol{MK}$ *for* f^+ *and* f^- *is equal to*

$$\boldsymbol{D} := \sup_{\phi\in Lip([0,1]\times M)} \left\{ \int_M \phi(1,x)df^- - \int_M \phi(0,x)df^+ \mid \partial_t\phi + H(x,d_x\phi) \leq 0 \ a.e. \right\}.$$

Proof. Fix ϕ admissible for $\mathbf{D}$. For each (x,y) let η be an optimal curve for the cost $c(x,y)$ and γ an optimal plan for $\mathbf{MK}$. Observe that $\phi(t,\eta(t))$ is absolutely continuous. We have

$$\int_M \phi(1,x)df^- - \int_M \phi(0,x)df^+ = \int_{M\times M} \phi(1,y)d\gamma(x,y) - \int_{M\times M} \phi(0,x)d\gamma(x,y) =$$

$$= \int_{M\times M}\int_0^1 \frac{d}{dt}[\phi(t,\eta(t))]dtd\gamma(x,y) =$$

$$= \int_{M\times M}\int_0^1 [\partial_t\phi(t,\eta(t)) + \langle d_x\phi(t,\eta(t)),\dot\eta(t)\rangle]dtd\gamma(x,y) \leq$$

$$\leq \int_{M\times M}\int_0^1 [-H(\eta(t),d_x\phi(t,\eta(t))) + \langle d_x\phi(t,\eta(t)),\dot\eta(t)\rangle]dtd\gamma(x,y) \leq$$

$$\leq \int_{M\times M}\int_0^1 L(\eta(t),\dot\eta(t))dtd\gamma(x,y) = \int_{M\times M} c(x,y)d\gamma(x,y) = \mathbf{MK}.$$

Taking the supremum with respect to ϕ we have the inequality $\mathbf{D} \leq \mathbf{MK}$.

Vice versa, using Theorem 1.2.2, $\forall\varepsilon > 0$ we can find continuous functions h_ε , g_ε such that

$$\forall(x,y) \in M\times M : h_\varepsilon(x) + g_\varepsilon(y) \leq c(x,y) \tag{4.7}$$

and

$$\mathbf{MK} \leq \int_M h_\varepsilon df^+ + \int_M g_\varepsilon df^- + \varepsilon. \tag{4.8}$$

Consider the Cauchy problem:

$$\begin{cases} \partial_t \phi(t,x) + H(x, d_x\phi(t,x)) = 0 \\ \phi(0,x) = -h_\varepsilon(x). \end{cases}$$

We claim that the function $\phi(t,x)$ defined through the Lax-Oleinik semigroup by

$$\phi(t,x) = T_t^-(-h_\varepsilon)(x) = \inf_{\eta \in AC([0,t],M)} \left\{ -h_\varepsilon(\eta(0)) + \int_0^t L(\eta,\dot\eta)ds \mid \eta(t) = x \right\},$$

solves this Cauchy problem a.e. Indeed, the general properties of the Lax-Oleinik semigroup (see Corollary 2.2.6) ensure that $\phi(t,x)$ is a locally Lipschitz function.

If (t_0, x_0) is a point of differentiability for $\phi(t,x)$, by Lemma 2.2.3 there exists an extremal curve η with $\eta(t_0) = x_0$ such that

$$\phi(t_0,x_0) = -h_\varepsilon(\eta(0)) + \int_0^{t_0} L(\eta,\dot\eta)ds.$$

Furthermore, by Lemma 2.2.7, we have that $d_x\phi(t_0,x_0) = \frac{\partial L}{\partial v}(\eta(t_0),\dot\eta(t_0))$. Observe that the real function

$$u(t) := \phi(t,\eta(t)) + h_\varepsilon(\eta(0)) - \int_0^t L(\eta,\dot\eta)ds$$

has a maximum point at t_0. Hence

$$\begin{aligned} 0 = u'(t_0) &= \partial_t\phi(t_0,\eta(t_0)) + \langle d_x\phi(t_0,x_0),\dot\eta(t_0))\rangle - L(\eta(t_0),\dot\eta(t_0)) = \\ &= \partial_t\phi(t_0,\eta(t_0)) + \langle \frac{\partial L}{\partial v}(\eta(t_0),\dot\eta(t_0)),\dot\eta(t_0)\rangle - L(\eta(t_0),\dot\eta(t_0)) = \\ &= \partial_t\phi(t_0,x_0) + H(x_0, d_x\phi(t_0,x_0)). \end{aligned}$$

Therefore ϕ is admissible for $\mathbf{D}$. Furthermore, we remark that $\phi(1,x) = h_\varepsilon^c(x)$, where h_ε^c is the c-transform of h_ε defined as in (1.24).

In fact if $y \in M$, and η is a Tonelli minimizer for $c(y,x)$, we have

$$\phi(1,x) \le -h_\varepsilon(y) + \int_0^1 L(\eta,\dot\eta)dt = -h_\varepsilon(y) + c(y,x).$$

Taking the infimum with respect to $y \in M$ we obtain $\phi(1,x) \le h_\varepsilon^c(x)$. Vice versa, by Lemma 2.2.3 there exists a Tonelli minimizer η such that

$$\phi(1,x) = -h_\varepsilon(\eta(0)) + \int_0^1 L(\eta,\dot\eta)dt = -h_\varepsilon(\eta(0)) + c(\eta(0),x),$$

from which we obtain

$$h_\varepsilon^c(x) \le \phi(1,x).$$

Observe that by (4.7) it results $g_\varepsilon(x) \le h_\varepsilon^c$, and then by (4.8) we have

$$\begin{aligned} \mathbf{MK} &\le \int_M h_\varepsilon df^+ + \int_M g_\varepsilon df^- + \varepsilon \le \int_M h_\varepsilon df^+ + \int_M h_\varepsilon^c df^- + \varepsilon = \\ &= \int_M \phi(1,x)df^- - \int_M \phi(0,x)df^+ + \varepsilon \le \mathbf{D} + \varepsilon. \end{aligned}$$

Letting $\varepsilon \to 0^+$ we obtain the desired inequality. $\square$

Corollary 4.0.10. $\boldsymbol{B} = \boldsymbol{MK}$.

Proof. Since $\mathbf{B} \leq \mathbf{MK}$ by Theorem 4.0.6, it remains to check the opposite inequality. Then fix ϕ admissible for $\mathbf{D}$ and (ρ, E) optimal for $\mathbf{B}$. As shown in Proposition 1.5.3, the curve $t \mapsto \rho_t$ is absolutely continuous with respect to the 1-Wasserstein distance. Therefore, the real function $t \mapsto \int_M \phi(t,x)d\rho_t$ is absolutely continuous too. Furthermore, since $\mathcal{C}^1$ maps are dense in $\mathcal{C}(M,\mathbb{R})$ with respect to the topology of uniform convergence, we can suppose without loss of generality that ϕ is $\mathcal{C}^1$. We have

$$\int_M \phi(1,x)df^- - \int_M \phi(0,x)df^+ =$$

$$= \int_M \phi(1,x)d\rho_1 - \int_M \phi(0,x)d\rho_0 = \int_0^1 \frac{d}{dt}[\int_M \phi(t,x)d\rho_t]dt =$$

$$= \int_0^1 \int_M \partial_t\phi(t,x)d\rho_t dt + \int_0^1 \int_M d_x\phi(t,x)dE_t dt \leq \tag{4.9}$$

$$\leq K(\rho, E) = \mathbf{B}.$$

Where the last equality is obtained as in (1.39) using that (ρ, E) solves the continuity equation (4.3). Taking the supremum with respect to ϕ we obtain $\mathbf{MK} \leq \mathbf{B}$. □

Remark 4.0.11. For the case $L = \varphi(|v|)$ see [19] and [1].

Time dependent problem

As second step we now consider the space $[0,1] \times M$, and indicate with the superscript ˜ all the quantity in this product space. In particular $\tilde{v} = (v_0, v) \in \mathbb{R} \times \mathbb{R}^n$, and $\tilde{\omega} = (\omega_0, \omega)$. We introduce the new Hamiltonian:

$$\widetilde{H}(t,x,\tilde{\omega}) := \omega_0 + H(x,\omega) \quad \forall\ (t,x,\tilde{\omega}) \in [0,1] \times M \times T^*_{(t,x)}\left([0,1] \times M\right).$$

The corresponding Lagrangian is

$$\tilde{L}(t,x,\tilde{v}) = \sup_{\tilde{\omega}}\{\langle\tilde{\omega},\tilde{v}\rangle - \widetilde{H}(t,x,\tilde{\omega})\} = \sup_{\tilde{\omega}}\{\langle\omega_0,v_0\rangle + \langle\omega,v\rangle - \omega_0 - H(x,\omega)\}.$$

From the above definition it follows the relation:

$$\tilde{L}(t,x,\tilde{v}) < +\infty \Leftrightarrow v_0 = 1 \Rightarrow \tilde{L}(t,x,\tilde{v}) = L(x,v). \tag{4.10}$$

Define the functional:

$$\widetilde{\mathcal{L}}(\tilde{\sigma},\widetilde{T}) = \sup_{\tilde{\alpha},\tilde{\omega}}\left\{\int_0^1 \int_M \tilde{\alpha}(t,x)d\tilde{\sigma} + \widetilde{T}(\tilde{\omega})\right\}. \tag{4.11}$$

Where $\tilde{\sigma} \in \mathcal{P}([0,1] \times M)$, $\widetilde{T} \in \mathcal{N}_1([0,1] \times M)$ and the supremum is taken for continuous $(\tilde{\alpha},\tilde{\omega})$ which satisfy the constraint $\tilde{\alpha}(t,x) + \widetilde{H}(t,x,\tilde{\omega}) \leq 0$.

By (4.10), Lemma 3.1.1 and Lemma 3.1.3 we have the following

Lemma 4.0.12. *If $\widetilde{\mathcal{L}}(\tilde{\sigma},\widetilde{T}) < +\infty$ then there exists $\widetilde{X} \in L^1_{\tilde{\sigma}}$ such that*

$$\widetilde{T} = \tilde{\sigma} \wedge \widetilde{X},\ X_0(t,x) = 1 \quad \tilde{\sigma} - a.e.$$

Furthermore, in such case we have that $\widetilde{\mathcal{L}}(\tilde{\sigma},\widetilde{T}) = \int_0^1 \int_M L(x, X(t,x))d\tilde{\sigma}(t,x)$.

Proof. If $\widetilde{\mathcal{L}}(\tilde{\sigma},\widetilde{T}) < +\infty$ we use the same arguments of Lemma 3.1.1 to find $\widetilde{X} \in L^1_{\tilde{\sigma}}$ such that $\widetilde{T} = \tilde{\sigma} \wedge \widetilde{X}$. Let us split $\widetilde{X} = (X_0, X)$.

Let $\omega_0 \in \mathbb{R}$ and consider the form $\tilde{\omega} = (\omega_0, \omega)$, where $\omega = \mathcal{L}(X(t,x))$ and $\mathcal{L}$ is the Legendre transform on M. Then we have

$$\widetilde{\mathcal{L}}(\tilde{\sigma}, \widetilde{T}) \geq \int_0^1 \int_M -\widetilde{H}(t,x,\tilde{\omega}) d\tilde{\sigma} + \widetilde{T}(\tilde{\omega}) =$$

$$= -\omega_0 - \int_0^1 \int_M H(x,\omega) d\tilde{\sigma} + \int_0^1 \int_M \langle \omega_0, X_0 \rangle d\tilde{\sigma} + \int_0^1 \int_M \langle \omega, X \rangle d\tilde{\sigma} =$$

$$= \int_0^1 \int_M L(x, X(t,x)) d\tilde{\sigma} + \int_0^1 \int_M \langle \omega_0, (X_0(t,x)) - 1) \rangle d\tilde{\sigma}.$$

However, by the assumption $\widetilde{\mathcal{L}}(\tilde{\sigma}, \widetilde{T}) < +\infty$ we infer that $X_0(t,x) = 1\ \tilde{\sigma} - a.e.$

The problem in the above argument is that in general ω is merely Borel. This problem can be solved considering a truncation $\omega_k(x) = \omega(x)$ if $|\omega(x)| \leq k$, $\omega_k(x) = 0$ otherwise, and then passing to the limit as $k \to +\infty$ as done in the proof of Lemma 3.1.3. $\square$

Define $F^+ = \delta_0 \otimes f^+$ and $F^- = \delta_1 \otimes f^-$ and consider the problem:

$$\widetilde{\mathcal{P}} := \min \left\{ \widetilde{\mathcal{L}}(\tilde{\sigma}, \widetilde{T}) \mid \partial \widetilde{T} = F^- - F^+ \right\},$$

then the analogous of problem $\mathbf{M}$ on $[0,1] \times M$ is

$$\widetilde{\mathbf{M}} := \min \left\{ \int_{T([0,1]\times M)} \tilde{L}(t,x,\tilde{v}) d\tilde{\mu} \mid \tilde{\mu} \in \mathcal{P}(T([0,1] \times M)),\ \partial p(\tilde{\mu}) = F^- - F^+ \right\}.$$

We have the following relationship with the Monge-Kantorovich problem.

Theorem 4.0.13. $\boldsymbol{MK} = \widetilde{\mathcal{P}} = \widetilde{\boldsymbol{M}}$.

Proof. If $\widetilde{T}$ is admissible for $\widetilde{\mathcal{P}}$, by Theorem 4.0.9 it results

$$\mathbf{MK} = \sup_{\phi \in Lip([0,1]\times M)} \left\{ \widetilde{T}(d\phi) \mid \partial_t \phi + H(x, d_x \phi) \leq 0 \right\}.$$

Hence, for any $\mathcal{C}^1$ function $\phi(t,x)$ the couple $(0, d\phi)$ is admissible for (4.11) and then $\mathbf{MK} \leq \widetilde{\mathcal{P}}$.

Vice versa, consider $\tilde{\gamma}$ optimal plan between F^+ and F^-, $\tilde{\eta}$ optimal curve for the cost $\tilde{c}(a,b)$ associated to $\tilde{L}$ as in (4.5), and denote with $\widetilde{\mathbf{MK}}$ the corresponding mass transportation problem. Furthermore consider the measure $\tilde{\sigma}$ defined as follows:

$$\int_0^1 \int_M f(t,x) d\tilde{\sigma} = \int_{([0,1]\times M)^2} \int_0^1 f(\tilde{\eta}(t)) dt d\tilde{\gamma}(a,b).$$

Also define:

$$\widetilde{T}(\tilde{\omega}) = \int_{([0,1]\times M)^2} \int_0^1 \langle \tilde{\omega}(\tilde{\eta}(t)), \dot{\tilde{\eta}}(t) \rangle dt d\tilde{\gamma}(a,b).$$

Then

$$\widetilde{T}(d\phi) = \int_{([0,1]\times M)^2} \int_0^1 \langle d\phi(\tilde{\eta}(t)), \dot{\tilde{\eta}}(t) \rangle dt d\tilde{\gamma}(a,b) =$$

$$= \int_{([0,1]\times M)^2} \int_0^1 \frac{d}{dt} [\phi(\tilde{\eta}(t))] dt d\tilde{\gamma}(a,b) =$$

$$= \int_{([0,1]\times M)^2} [\phi(b) - \phi(a)] d\tilde{\gamma}(a,b) = \int_{[0,1]\times M} \phi(b) dF^- - \int_{[0,1]\times M} \phi(a) dF^+.$$

And then $\partial\widetilde{T} = F^- - F^+$.

Hence, if $(\tilde{\alpha}, \tilde{\omega})$ is admissible for (4.11), we have

$$\int_0^1 \int_M \tilde{\alpha}(t,x)d\tilde{\sigma} + \tilde{T}(\tilde{\omega}) = \int_{([0,1]\times M)^2} \int_0^1 [\tilde{\alpha}(\tilde{\eta}(t)) + \langle\tilde{\omega}(\tilde{\eta}(t)), \dot{\tilde{\eta}}(t)\rangle]dtd\tilde{\gamma}(a,b) \leq$$

$$\leq \int_{([0,1]\times M)^2} \int_0^1 [-\widetilde{H}(\tilde{\eta}(t), \tilde{\omega}(\tilde{\eta}(t))) + \langle\tilde{\omega}(\tilde{\eta}(t)), \dot{\tilde{\eta}}(t)\rangle]dtd\tilde{\gamma}(a,b) \leq$$

$$\leq \int_{([0,1]\times M)^2} \int_0^1 \tilde{L}(\tilde{\eta}, \dot{\tilde{\eta}})dtd\tilde{\gamma}(a,b) = \int_{([0,1]\times M)^2} \tilde{c}(a,b)d\tilde{\gamma}(a,b) = \widetilde{\mathbf{MK}}.$$

Taking the supremum with respect to $(\tilde{\alpha}, \tilde{\omega})$ we obtain $\widetilde{\mathcal{P}} \leq \widetilde{\mathbf{MK}}$. Now, let us take γ an optimal plan on $M \times M$ between f^+ and f^- and consider the plan between F^+ and F^- defined as follows:

$$\tilde{\gamma} = \delta_0(t_1) \otimes \delta_1(t_2) \otimes \gamma.$$

Let η be an optimal curve for the cost $c(x_1, x_2)$ on $[0,1] \times M$, and the curve $\tilde{\eta}(t) = (t, \eta(t))$ on $[0,1] \times M$ we can evaluate

$$\widetilde{\mathcal{P}} \leq \widetilde{\mathbf{MK}} \leq \int_{([0,1]\times M)^2} \tilde{c}(t_1, x_1, t_2, x_2)d\tilde{\gamma}(t_1, x_1, t_2, x_2) =$$

$$= \int_{M\times M} \tilde{c}(0, x_1, 1, x_2)d\gamma(x_1, x_2) \leq$$

$$\leq \int_{M\times M} \int_0^1 \tilde{L}(\tilde{\eta}, \dot{\tilde{\eta}})dtd\gamma(x_1, x_2) = \int_{M\times M} \int_0^1 L(\eta, \dot{\eta})dtd\gamma(x_1, x_2) =$$

$$= \int_{M\times M} c(x_1, x_2)d\gamma(x_1, x_2) = \mathbf{MK},$$

which conclude the proof of the first equivalence.

To prove the second equivalence let us consider $\tilde{\mu}$ admissible for $\widetilde{\mathbf{M}}$. By disintegration Theorem 1.4.3 we can write $\tilde{\mu} = \tilde{\nu}_{t,x} \otimes \tilde{\sigma}$, with

$$\tilde{\nu}_{t,x} \in \mathcal{P}(T_{(t,x)}([0,1] \times M)), \ \tilde{\sigma} \in \mathcal{P}([0,1] \times M).$$

Using Jensen's inequality we have

$$\widetilde{\mathcal{L}}(\tilde{\sigma}, p(\tilde{\mu})) \leq \int_{[0,1]\times M} \tilde{L}\left(t, x, \int_{T_{(t,x)}([0,1]\times M)} \tilde{v}d\tilde{\nu}_{t,x}\right) d\tilde{\sigma} \leq$$

$$\leq \int_{[0,1]\times M} \int_{T_{(t,x)}([0,1]\times M)} \tilde{L}(t, x, \tilde{v})d\tilde{\nu}_{t,x}d\tilde{\sigma} =$$

$$= \int_{T([0,1]\times M)} \tilde{L}(t, x, \tilde{v})d\tilde{\mu} = \mathcal{A}(\tilde{\mu}).$$

Then $\widetilde{\mathcal{P}} \leq \widetilde{\mathbf{M}}$. Vice versa, let us take $(\tilde{\sigma}, \widetilde{T})$ such that $\tilde{\mathcal{L}}(\tilde{\sigma}, \widetilde{T}) < +\infty$. By Lemma 4.0.12 we have that $\widetilde{T} = \tilde{\sigma} \wedge \widetilde{X}$ with $X_0(t,x) = 1$ $\tilde{\sigma}$ - a.e.

Considering $i(\tilde{\sigma}, \widetilde{T}) = (id \times \widetilde{X})_\sharp \tilde{\sigma}$ we obtain

$$\mathcal{A}(i(\tilde{\sigma}, \widetilde{T})) = \int_{[0,1]\times M} \tilde{L}(t, x, \widetilde{X})d\tilde{\sigma} = \int_{[0,1]\times M} L(x, X(t,x))d\tilde{\sigma} = \widetilde{\mathcal{L}}(\tilde{\sigma}, \widetilde{T}).$$

Then $\widetilde{\mathbf{M}} \leq \widetilde{\mathcal{P}}$.

□

Remark 4.0.14. As in Proposition 3.1.4 one can check that $\widetilde{\mathcal{P}}$ admits minimizers. Then the arguments above tell us that $\widetilde{\mathbf{M}}$ admits minimizers too.

The Monge-Kantorovich problem is then equivalent to the problem

$$\widetilde{\mathbf{M}} = \min\left\{\int_{T([0,1]\times M)} \tilde{L}(t,x,\tilde{v})d\tilde{\mu} \mid \tilde{\mu} \in \mathcal{P}(T([0,1]\times M)),\ \partial p(\tilde{\mu}) = F^- - F^+\right\}.$$

Moreover, by (4.10) we know that $\tilde{L}(t,x,\tilde{v}) < +\infty \Leftrightarrow v_0 = 1$. Hence, the value of the minimum does not change if we minimize on the set of measures of the form $\tilde{\mu} = \delta_1 \otimes \bar{\mu}$, with $\bar{\mu} \in \mathcal{M}([0,1]\times TM)$.

Considering the projection on the second factor $\pi_{(x,v)} : [0,1]\times TM \to TM$ and $\mu = (\pi_{(x,v)})_\sharp \bar{\mu}$, by disintegration Theorem 1.4.3 there exists $\nu_{x,v} \in \mathcal{P}([0,1])$ such that $\bar{\mu} = \nu_{x,v}\otimes\mu$. Hence we have

$$\int_{T([0,1]\times M)} \tilde{L}(t,x,\tilde{v})d\tilde{\mu} = \int_0^1\int_{TM} L(x,v)d\bar{\mu} =$$

$$= \int_{TM}\left[\int_0^1 L(x,v)d\nu_{x,v}(t)\right]d\mu(x,v) = \int_{TM} L(x,v)d\mu(x,v).$$

Now we compute the boundary of $p(\mu)$. Observe that

$$d_{(t,x)}\varphi(\pi_{(x,v)}(t,x,v)) =$$
$$= \partial_t\varphi(\pi_{(x,v)}(t,x,v))\partial_t(\pi_{(x,v)}(t,x,v)) + d_x\varphi(\pi_{(x,v)}(t,x,v))\cdot\partial_x(\pi_{(x,v)}(t,x,v)) =$$
$$= d_x\varphi(\pi_{(x,v)}(t,x,v)).$$

Then we can evaluate

$$\int_{TM}\langle d\varphi, v\rangle d\mu = \int_0^1\int_{TM}\langle d_x\varphi(\pi_{(x,v)}(t,x,v)), v\rangle d\bar{\mu} =$$

$$= \int_{T([0,1]\times M)}\langle d_{(t,x)}\varphi(\pi_{(x,v)}(t,x,v)),\tilde{v}\rangle d\tilde{\mu} = \int_0^1\int_M \varphi(\pi_{(x,v)}(t,x,v))d(F^- - F^+) =$$

$$= \int_M \varphi(x)df^- - \int_M\varphi(x)df^+ \Leftrightarrow \partial p(\mu) = f^- - f^+. \tag{4.12}$$

Then define:

$$\Lambda = \left\{\mu\in\mathcal{P}(TM) \mid \text{ for } \mu - a.e.(x,v)\ \exists\ \nu_{x,v}\in\mathcal{P}([0,1])\ s.t.\ \partial p(\delta_1\otimes\nu_{x,v}\otimes\mu) = F^- - F^+\right\},$$

where writing $\delta_1\otimes\nu_{x,v}\otimes\mu$ we assume that $\nu_{x,v}$ satisfy the properties (i) and (ii) of Theorem 1.4.3, and consider

$$\mathbf{M}_c := \min\{\mathcal{A}(\mu) \mid \mu\in\Lambda\}. \tag{4.13}$$

By uniqueness of disintegration of measures and Theorem 4.0.13 we can state the following

Theorem 4.0.15. $\boldsymbol{MK} = \boldsymbol{M}_c$.

Therefore, the Monge-Kantorovich problem can be obtained from problem $\mathbf{M}$ on the phase space defined in (4.1) by adding a constraint on the set of the admissible measures. This additional constraint actually define a strict subfamily of admissible measures as we are going to prove.

By definition of F^- and F^+, a measure $\mu\in\Lambda$ if and only if

$$\int_{TM}\int_0^1\left[\partial_t\varphi(t,x) + \langle d_x\varphi(t,x), v\rangle\right]d\nu_{x,v}(t)d\mu(x,v) = \int_M\varphi(1,x)df^- - \int_M\varphi(0,x)df^+, \tag{4.14}$$

for every $\mathcal{C}^1$ function $\varphi(t,x)$ on $[0,1]\times M$.

Lemma 4.0.16. *Let $f : [0,1] \to \mathbb{R}$ be a continuous function. Then*

$$\int_{TM} \left[\int_0^1 f(t) d\nu_{x,v}(t) \right] d\mu(x,v) = \int_0^1 f(t)dt.$$

Proof. It is sufficient to take $F(t) = \int_0^t f(s)ds$ and to put $\varphi(t,x) = F(t)$ into (4.14). □

Observe that if $\mu \in \Lambda$ is such that the disintegration $\nu_{x,v}$ is independent from (x,v), then by Lemma 4.0.16 $\nu_{x,v} = \mathcal{L}^1_{\lfloor [0,1]}$.

The following example shows that there are admissible measures μ for $\mathbf{M}$ which do not belongs to Λ. Furthermore, we will prove that in the case of $f^+ = f^-$ every such measure μ belongs to Λ. Actually, it turns out that the Lebesgue measure on $[0,1]$ always satisfy the condition (4.14) in place of $\nu_{x,v}$ in the case of $f^+ = f^-$, while it in general fails for $f^+ \neq f^-$. Note that in our case the choice $f^+ = f^-$ is non-trivial and corresponds to minimize the action of closed measures, see section 2.4 and chapter 3.

For sake of simplicity, in the following examples we consider the manifold $M = [0,1]$. However, the same arguments can be repeated on S^1 without further difficulty.

Example 4.0.17. *Let $M = [0,1]$, $\mu = 2x\mathcal{L}^1_{\lfloor [0,1]} \otimes \delta_{\frac{1}{2}}$ and suppose that $\mu \in \Lambda$. Observe that*

$$\int_{TM} \langle d\varphi, v \rangle d\mu = \int_0^1 \varphi'(x) x dx =$$

$$= \varphi(1) - \int_0^1 \varphi(x)dx \Rightarrow \partial p(\mu) = \delta_1 - \mathcal{L}^1_{\lfloor [0,1]}.$$

Choosing $\varphi(t,x) = tf(x)$ into (4.14) we obtain

$$\int_0^1 2xf(x)dx + \int_0^1 [\int_0^1 t d\nu_{x,\frac{1}{2}}(t)] f'(x) x dx = f(1). \tag{4.15}$$

Taking $f(x) = 2x$, by Lemma 4.0.16 the left side of (4.15) becomes

$$\int_0^1 4x^2 dx + \int_0^1 t dt = \frac{11}{6}.$$

However, this is a contradiction since the right side of (4.15) is equal to 2.

The previous example can be builded on S^1 with the same technique taking the measures $\mu_\varepsilon = \frac{x-\varepsilon}{c_\varepsilon} \mathcal{L}^1_{\lfloor [\varepsilon, 1-\varepsilon]} \otimes \delta_{c_\varepsilon}$, with $c_\varepsilon = \frac{1}{2} + 2\varepsilon(\varepsilon - 1)$, and then letting $\varepsilon \to 0^+$.

We also remark that in the case $L(x,v) = g(v,v)^{\frac{p}{2}}$ for $p > 1$, the values of the functionals in $\mathbf{M}_{2p}$ and $\mathbf{MK}_{2p}$ satisfy the relation

$$\mathbf{M}_{2p} = \int_{TM} |v|^p d\mu = \frac{1}{2^p} < \frac{1}{p+1} = \int_M |x|^p dx = W_p(\mathcal{L}^1_{\lfloor [0,1]}, \delta_1)^p = \mathbf{MK}_{2p},$$

showing the necessity of the additional constraint Λ in (4.13) to establish the equality $\mathbf{MK} = \mathbf{M}_c$.

However, if $f^+ = f^-$ then the additional constraint Λ in $\mathbf{M}_c$ does not play any role, as results by the following

Lemma 4.0.18. *Every closed measures $\mu \in \mathcal{P}(TM)$ belongs to Λ.*

Proof. Let $\mu \in \mathcal{P}(TM)$ be a closed measure. By (3.14) and (2.18) this is equivalent to have $\partial p(\mu) = 0 = \sigma - \sigma$, with σ the projection of μ on M. Taking $\tilde{\mu} = \delta_1 \otimes \mathcal{L}^1_{\lfloor [0,1]} \otimes \mu$, for every $\mathcal{C}^1$ function $\varphi(t,x)$ on $[0,1] \times M$ we have

$$\int_{T([0,1]\times M)} \langle d\varphi(t,x), \tilde{v}\rangle d\tilde{\mu} =$$

$$= \int_{TM} \left[\int_0^1 \partial_t \varphi(t,x)dt\right] d\mu(x,v) + \int_{TM} \left[\int_0^1 \langle d_x\varphi(t,x), v\rangle dt\right] d\mu(x,v) =$$

$$= \int_{TM} [\varphi(1,x) - \varphi(0,x)]\, d\mu(x,v) = \int_M \left[\int_{T_xM} \varphi(1,x) - \varphi(0,x) d\nu_x(v)\right] d\sigma =$$

$$= \int_M \varphi(1,x)d\sigma - \int_M \varphi(0,x)d\sigma \Rightarrow \partial p(\tilde{\mu}) = F^- - F^+,$$

where the second addendum in the second equality is zero since $\partial p(\mu) = 0$, and we have again used the disintegration Theorem 1.4.3. Hence $\mu \in \Lambda$. $\square$

Denoted with $\mathbf{M}_c^0$ the value of $\mathbf{M}_c$ in the case of $f^+ = f^-$, as consequence of Lemma 4.0.18 we have the following

Proposition 4.0.19. $-c_0 := \min \{\mathcal{A}(\mu) \mid \partial p(\mu) = 0\} = \boldsymbol{M}_c^0$.

Here c_0 is the Mañé critical value (see section 2.4). We have the same result if we also consider an homological constraint and then recovering Mather's problem $\beta([h])$.

We end this section showing that in the case of $f^+ \neq f^-$ the Lebesgue measure in general does not satisfy the condition (4.14) in place of $\nu_{x,v}$.

Example 4.0.20. *Let again $M = [0,1]$, and $\mu = \mathcal{L}^1_{\lfloor[0,1]} \otimes \delta_1$. Observe that*

$$\int_{TM} \langle d\varphi, v\rangle d\mu = \int_0^1 \varphi'(x)dx =$$

$$= \varphi(1) - \varphi(0) \Rightarrow \partial p(\mu) = \delta_1 - \delta_0.$$

Now, putting $\varphi(t,x) = tf(x)$ and $\nu_{x,v}(t) = \mathcal{L}^1_{\lfloor[0,1]}$ into (4.14) we obtain

$$\int_0^1 f(x)dx + \int_0^1 [\int_0^1 tdt] f'(x)dx = f(1) \Leftrightarrow \tag{4.16}$$

$$\Leftrightarrow \int_0^1 f(x)dx + \frac{1}{2}\int_0^1 f'(x)dx = f(1) \Leftrightarrow \tag{4.17}$$

$$\Leftrightarrow \int_0^1 f(x)dx = \frac{1}{2}f(0) + \frac{1}{2}f(1). \tag{4.18}$$

Any regular function f with zero boundary values and non-vanishing integral is a counterexample to (4.16).

In this particular case we can take $\nu_x(t) = \delta_x(t)$ to obtain

$$\int_0^1\int_0^1 \partial_t\varphi(t,x)d\nu_x(t)dx + \int_0^1\int_0^1 \partial_x\varphi(t,x)d\nu_x(t)dx =$$

$$= \int_0^1 \partial_t\varphi(x,x)dx + \int_0^1 \partial_x\varphi(x,x)dx = \int_0^1 \frac{d}{dx}\varphi(x,x)dx = \varphi(1,1) - \varphi(0,0),$$

and then satisfying the condition (4.14).

The previous example can be builded on S^1 taking the measure $\mu_\varepsilon = \frac{1}{1-2\varepsilon}\mathcal{L}^1 \llcorner [\varepsilon, 1-\varepsilon] \otimes \delta_1$.

Chapter 5

Geometry of $(\mathcal{P}_p(M), W_p)$ and a counterexample to the extension of Lipschitz maps

In this chapter we discuss some geometric properties of $(\mathcal{P}_p(M), W_p)$ and we use these properties to give a counterexample to the extension of Lipschitz map between metric spaces with bounded curvature, see example 5.3.4.

Since we are interested to study the space $\mathcal{P}(M)$ with M a compact Riemannian manifold, throughout this chapter (M, d) is assumed to be a complete, separable and geodesic space. A metric space is said geodesic if for each pair of points $x, y \in M$ the distance $d(x, y)$ is given by

$$d(x,y) = \min\left\{ l(\gamma) := \int_0^1 |\dot{\gamma}|(t)dt \mid \gamma \in \mathrm{Lip}([0,1], M), \gamma(0) = x, \gamma(1) = y \right\}. \tag{5.1}$$

Where $|\dot{\gamma}|$ is the metric derivative of γ defined by (1.36). If $x, y \in M$, we say that a Lipschitz curve $\gamma : [0,1] \to M$ is a geodesic between x and y if γ is a minimizing curve in (5.1). In this chapter we will always assume that any geodesic is parametrized proportionally to arclength, i.e. $d(\sigma(t), \sigma(s)) = v|t - s|$, where v is the constant speed. We say that the metric space (M, d) admits a measurable choice of geodesics if there exists a measurable function $f : M \times M \times [0,1] \to M$ such that the map $t \mapsto f(x, y, t)$ is a geodesic between x and y. If for instance M is a Banach space, the function

$$f(x, y, t) = x + t(y - x)$$

is a measurable choice of geodesics. Actually, f is a continuous function. This is also true if M is a compact Riemannian manifold replacing the segments with geodesics and using the fact that geodesics depend continuously with respect to their endpoints. Observe that all this notions are inherited by $\mathcal{P}_p(M)$ as stated in the following

Proposition 5.0.21. *If M admits a measurable choice of geodesics, then for every $1 \leq p$ the space $(\mathcal{P}_p(M), W_p)$ is complete, separable and geodesic.*

Proof. Using Banach-Alaoglu and Krein-Milman theorem we find that every probability measure can be approximated in the weak$*$ convergence by measures of the form $\sum_{i=1}^n t_i \delta_{x_i}$, with $t_i \geq 0$ and $\sum_{i=1}^n t_i = 1$. It costs no generality to assume that t_i are rational and then to restrict ourselves to measures of the form $\frac{1}{n}\sum_{i=1}^n \delta_{x_i}$. Let us denote by $\mathcal{P}_A(M)$ the set of probability measures of this last form. By a truncation argument and using Theorem 1.4.9 on bounded sets, it follows that $\mathcal{P}_A(M)$ is dense in $(\mathcal{P}_p(M), W_p)$. If M is separable the points x_i can be taken to be countable and therefore $\mathcal{P}_p(M)$ is separable too. To prove the completeness we can assume that M is a Banach space. Indeed, every separable metric space can be isometrically embedded into l^∞, see Theorem 4.4.3 of [6], taking the function

$$j(x) := (d(x, x_1) - d(x_0, x_1), d(x, x_2) - d(x_0, x_2), \cdots),$$

where $\{x_n\} \subset M$ is a dense subset, so that $d(x, y) = \sup_{h \geq 0} |d(x, x_h) - d(y, x_h)| = \|j(x) - j(y)\|_\infty$. It is then sufficient to check that any sequence $(\mu_n) \subset \mathcal{P}_p(M)$ such that

$$\sum_{n=1}^{+\infty} W_p(\mu_n, \mu_{n+1}) < +\infty,$$

converges. Let $\gamma_{n,n+1}$ be optimal plans between μ_n and μ_{n+1}. Set $\widetilde{M} = M^{\mathbb{N}}$. By Lemma 1.4.6 there exists $\tilde{\mu} \in \mathcal{P}(\widetilde{M})$ such that $(\pi_n, \pi_{n+1})_\sharp \tilde{\mu} = \gamma_{n,n+1}$ and $(\pi_n)_\sharp \tilde{\mu} = \mu_n$. Observe that

$$\int_{\widetilde{M}} \|\pi_n - \pi_{n+1}\|^p d\tilde{\mu} = \int_{M \times M} \|x - y\|^p d\gamma_{n,n+1} = W_p(\mu_n, \mu_{n+1})^p,$$

so that

$$\sum_{n=1}^{+\infty} \|\pi_n - \pi_{n+1}\|_{L^p_{\tilde{\mu}}(\widetilde{M}, M)} < +\infty.$$

Therefore the series $\pi_1 + \sum_{n=1}^{+\infty} (\pi_n - \pi_{n+1})$ is converging and defines a map $\tilde{\pi} \in L^p_{\tilde{\mu}}(\widetilde{M}, M)$. Set $\mu = \tilde{\pi}_\sharp \tilde{\mu}$. Since the measure $\eta = (\pi_n, \tilde{\pi})_\sharp \tilde{\mu}$ is a transport plan between μ_n and μ we have

$$\limsup_{n \to +\infty} W_p(\mu_n, \mu) \leq \limsup_{n \to +\infty} \left(\int_{M \times M} \|x - y\|^p d\eta \right)^{\frac{1}{p}} = \limsup_{n \to +\infty} \|\pi_n - \tilde{\pi}\|_{L^p_{\tilde{\mu}}(\widetilde{M}, M)} \leq$$

$$\leq \limsup_{n \to +\infty} \sum_{j=n}^{+\infty} \|\pi_j - \pi_{j+1}\|_{L^p_{\tilde{\mu}}(\widetilde{M}, M)} = 0.$$

Let us now fix $\mu, \nu \in \mathcal{P}_p(M)$ and take γ an optimal transport plan between μ, ν. For $0 \leq t \leq 1$ consider the function $f_t(x, y) = \sigma_{x,y}(t)$, where $\sigma_{x,y}$ is a geodesic connecting x to y. By hypothesis the function $f_t : M \times M \to M$ is measurable.

Now, we claim that the curve $\rho : [0, 1] \to M$ defined by $\rho(t) = (f_t)_\sharp \gamma$ is a geodesic, parametrized proportionally to arclength, between μ, ν. Indeed, observe that $\eta = (f_t \times f_s)_\sharp \gamma$ is a transport plan between $\rho(t)$ and $\rho(s)$. Then

$$W_p(\rho(t), \rho(s)) \leq \left(\int_{M \times M} d(x, y)^p d\eta \right)^{\frac{1}{p}} = \left(\int_{M \times M} d(\sigma_{x,y}(t), \sigma_{x,y}(s))^p d\gamma \right)^{\frac{1}{p}} =$$

$$= |t - s| \left(\int_{M \times M} d(x, y)^p d\gamma \right)^{\frac{1}{p}} = |t - s| W_p(\mu, \nu).$$

Actually, the above inequality holds with the equality sign. In fact, if there is a strict inequality for some $s < t$ we would have

$$W_p(\mu, \nu) \leq W_p(\mu, \rho(s)) + W_p(\rho(s), \rho(t)) + W_p(\rho(t), \nu) <$$

$$< sW_p(\mu, \nu) + (t - s)W_p(\mu, \nu) + (1 - t)W_p(\mu, \nu) = W_p(\mu, \nu).$$

□

Lemma 5.0.22. *$(\mathcal{P}_p(M), W_p)$ is contractible and locally contractible.*

Proof. Let μ_t a continuous curve in $(\mathcal{P}_p(M), W_p)$. Consider the function

$$F : [0,1] \times [0,1] \to \mathcal{P}_p(M),$$

defined by

$$F(t,s) = s\mu_0 + (1-s)\mu_t. \tag{5.2}$$

We have that $F(t,0) = \mu_t$, while $F(t,1) = \mu_0$. By Proposition 1.4.12 we have

$$W_p(F(t,s), F(t_0,s_0)) = W_p(s\mu_0 + (1-s)\mu_t, s_0\mu_0 + (1-s_0)\mu_{t_0}) \leq$$

$$\leq W_p(s\mu_0 + (1-s)\mu_t, s\mu_0 + (1-s)\mu_{t_0}) + W_p(s\mu_0 + (1-s)\mu_{t_0}, s_0\mu_0 + (1-s_0)\mu_{t_0}) \leq$$

$$\leq (1-s)^{\frac{1}{p}} W_p(\mu_t, \mu_{t_0}) + W_p(s\mu_0 + (1-s)\mu_{t_0}, s_0\mu_0 + (1-s_0)\mu_{t_0}) \leq \varepsilon,$$

provided that $|s-s_0|, |t-t_0| \leq \delta$. To check the second claim consider the ball $B(\nu,1) \subset \mathcal{P}_p(M)$. If μ_t is a curve on the ball such that $\mu_0 = \nu$, by (1.4.12) and (5.2) we have

$$W_p(\nu, F(s,t)) = W_p(\mu_0, s\mu_0 + (1-s)\mu_t) \leq (1-s)^{\frac{1}{p}} W_p(\mu_0, \mu_t) \leq 1.$$

□

5.1 Curvature in metric spaces

For a Riemannian manifold with constant sectional curvature there exist several characterizations of the curvature in terms of the geodesic distance. One of these characterizations is stated in the following

Theorem 5.1.1. *A Riemannian manifold (M,g) has sectional curvature ≥ 0 (resp. ≤ 0) if and only if for every point $p \in M$ there exists $r_p > 0$ such that if $\sigma : [0,1] \to M$ is a geodesic (parametrized proportionally to arclength) contained in $B(p,r_p)$, for $0 \leq t \leq 1$ it results:*

$$d(p,\sigma(t))^2 \geq (1-t)d(p,\sigma(0))^2 + td(p,\sigma(1))^2 - t(1-t)d(\sigma(0),\sigma(1))^2 \ \textit{(resp. } \leq\textit{)}. \tag{5.3}$$

Since it involves only metric quantities, the inequality (**??**) can be used to define positive or negative curvature also on metric spaces.

If $x_0, x_1 \in M$, we say that $m \in M$ is a midpoint between x_0, x_1 if it results:

$$d(x_0,m) = \frac{1}{2}d(x_0,x_1) = d(x_1,m).$$

Observe that if $\sigma : [0,1] \to M$ is a geodesic between x_0, x_1 then $\sigma(\frac{1}{2})$ is a midpoint between x_0, x_1. The following definition introduce the concept of curvature in the sense of Alexandrov.

Definition 5.1.2. *A metric space (M,d) has local curvature ≥ 0 (resp. ≤ 0) if for every $p \in M$ there exists $r_p > 0$ such that for all $x_0, x_1, z \in B(p,r_p)$ and for every midpoint $m \in M$ between x_0, x_1 it results:*

$$4d(z,m)^2 - 2d(z,x_0)^2 - 2d(z,x_1)^2 + d(x_0,x_1) \geq 0 \ \textit{(resp. } \leq 0\textit{)}. \tag{5.4}$$

(M,d) has global curvature ≥ 0 if condition (5.4) holds in the whole space.

The condition (5.4) can be stated in terms of geodesics.

Proposition 5.1.3. *(M,d) has local curvature ≥ 0 (resp. ≤ 0) if and only if for every $p \in M$ there exists $r_p > 0$ such that for every $x_0, x_1, z \in B(p,r)$ and for any geodesic $\sigma : [0,1] \to M$ connecting x_0 to x_1 contained in $B(p,r_p)$, for $0 \leq t \leq 1$ it results:*

$$d(z,\sigma(t))^2 \geq (1-t)d(z,\sigma(0))^2 + td(z,\sigma(1))^2 - t(1-t)d(\sigma(0),\sigma(1))^2 \ \textit{(resp. } \leq\textit{)}. \tag{5.5}$$

Proof. Let us suppose that (5.4) holds. The inequality (5.5) obviously holds for $t = 0$ and $t = 1$. Fix $n \in \mathbb{N}$ and suppose that (5.5) holds for $t = \frac{k}{2^n}$ with $k = 0, 1, \ldots, 2^n$. We want to prove that (5.5) also holds for $t = \frac{k}{2^{n+1}}$ with $k = 0, 1, \ldots, 2^{n+1}$. For even k this is true by assumption. Fixed $t = \frac{k}{2^{n+1}}$ with an odd k, if we set $\Delta t = \frac{1}{2^{n+1}}$, applying (5.5) at the points $t \pm \Delta t$ we obtain

$$d(z, \sigma(t \pm \Delta t))^2 \geq$$
$$\geq (1 - t \mp \Delta t) d(z, \sigma(0))^2 + (t \pm \Delta t) d(z, \sigma(1))^2 - (t \pm \Delta t)(1 - t \mp \Delta t) d(\sigma(0), \sigma(1))^2. \tag{5.6}$$

By (5.4) we have

$$d(z, \sigma(t))^2 \geq \frac{1}{2} d(z, \sigma(t - \Delta t))^2 + \frac{1}{2} d(z, \sigma(t + \Delta t))^2 - \frac{1}{4} d(\sigma(t - \Delta t), \sigma(t + \Delta t))^2. \tag{5.7}$$

Putting together (5.6) and (5.7) we obtain

$$\begin{aligned}
d(z, \sigma(t))^2 &\geq \frac{1}{2}(1 - t + \Delta t) d(z, \sigma(0))^2 + \frac{1}{2}(t - \Delta t) d(z, \sigma(1))^2 - \\
&- \frac{1}{2}(1 - t + \Delta t)(t - \Delta t) d(\sigma(0), \sigma(1))^2 + \frac{1}{2}(1 - t - \Delta t) d(z, \sigma(0))^2 + \\
&+ \frac{1}{2}(t + \Delta t) d(z, \sigma(1))^2 - \frac{1}{2}(1 - t - \Delta t)(t + \Delta t) d(\sigma(0), \sigma(1))^2 = \\
&= (1 - t) d(z, \sigma(0))^2 + t d(z, \sigma(1))^2 + \\
&+ d(\sigma(0), \sigma(1))^2 [-\Delta t^2 - \frac{1}{2}(1 - t + \Delta t)(t - \Delta t) - \frac{1}{2}(1 - t - \Delta t)(t + \Delta t)] = \\
&= (1 - t) d(z, \sigma(0))^2 + t d(z, \sigma(1))^2 - t(1 - t) d(\sigma(0), \sigma(1))^2.
\end{aligned}$$

By continuity the inequality (5.5) holds for every $0 \leq t \leq 1$. Vice versa, let $x_0, x_1, z \in B(p, r_p)$ and $m \in M$ a midpoint between x_0, x_1. Set $x_{\frac{1}{2}} = m$. Then we can obtain the points $x_{\frac{1}{4}}$ and $x_{\frac{3}{4}}$ as midpoints between $x_0, x_{\frac{1}{2}}$ and $x_{\frac{1}{2}}, x_1$ respectively. Iterating this procedure we obtain the points $\sigma(t) := x_t$ for $t = \frac{k}{2^n}$ with $k = 0, 1, \ldots, 2^n$. Observe that $d(\sigma(t), \sigma(s)) = d(x_0, x_1)|t - s|$. Since the space is complete we can extend σ by continuity in order to obtain a geodesic $\sigma : [0, 1] \to M$. Applying the inequality (5.5) for $t = \frac{1}{2}$ we obtain the desired condition (5.4). $\square$

We remark that by the proof of Proposition 5.1.3 it follows that a complete metric space is geodesic if and only if every pair of points admits a midpoint.

By Theorem 5.1.1 and Proposition 5.1.3 we see that these two notions of curvature agree on Riemannian manifolds.

Proposition 5.1.4. *A complete Riemannian manifold has sectional curvature ≥ 0 (resp. ≤ 0) if and only if it has local curvature ≥ 0 (resp. ≤ 0) as metric space.*

In order to study the geometry of the space of probability measures we have to consider some Wasserstein distance. In fact Banach spaces have no curvature properties.

Proposition 5.1.5. *A Banach space X has global curvature ≥ 0 or ≤ 0 if and only if it is an Hilbert space.*

Proof. If X is an Hilbert space then (5.4) holds with the equality sign as an immediate consequence of the parallelogram equality. Let us now suppose that X has global curvature ≥ 0. Let $x_0, x_1 \in X$ and take the midpoint $\frac{x_0 + x_1}{2}$. Choosing $z = 0$ in (5.4) we obtain

$$\|x_0 + x_1\|^2 + \|x_0 - x_1\|^2 \geq 2\|x_0\|^2 + 2\|x_1\|^2. \tag{5.8}$$

Replacing x_0 and x_1 with $x_0 + x_1$ and $x_0 - x_1$ respectively in (5.8) we obtain

$$\|x_0 - x_1\|^2 + \|x_0 + x_1\|^2 \leq 2\|x_0\|^2 + 2\|x_1\|^2.$$

A similar argument works if X has global curvature ≤ 0. $\square$

Observe that the condition (5.5) holds with the equality sign for a triangle in $\mathbb{R}^2$. Hence the geometric meaning of (5.5) is that considered a comparison triangle $\bar{\Delta}(\bar{x}_0, \bar{x}_1, \bar{z})$ in $\mathbb{R}^2$, i.e. such that

$$d(x_0, x_1) = |\bar{x}_0 - \bar{x}_1|, \ d(x_0, z) = |\bar{x}_0 - \bar{z}|, \ d(z, x_1) = |\bar{z} - \bar{x}_1|,$$

and denoted with $\bar{\sigma}$ the segment joining $\bar{x}_0, \bar{x}_1$ it results:

$$\forall t \in [0,1] : d(z, \sigma(t))^2 \geq |z - \bar{\sigma}(t)|^2. \tag{5.9}$$

The condition (5.9) can be taken as an equivalent definition of curvature as shown in [77]. Furthermore, this condition on comparison triangles can be also stated in terms of "angles". Finally it turns out that for curvature bounded from below the local curvature is in fact a global condition. We summarize this assertions in the following theorem the proof of which is in [77].

Theorem 5.1.6. *The following statements are equivalent:*

1. *M has local curvature ≥ 0.*
2. *M has global curvature ≥ 0.*
3. *All triangles defined through geodesics $\Delta(\sigma_1, \sigma_2, \sigma_3)$ satisfy:*

$$d(\sigma_i, (s)\sigma_j(t)) \geq |\bar{\sigma}_i(s), \bar{\sigma}_j(t)|,$$

for all $s, t \in [0,1]$ and $i, j = 1, 2, 3$ with $\bar{\Delta}(\bar{\sigma}_1, \bar{\sigma}_2, \bar{\sigma}_3)$ a comparison triangle.

4. *For all $k \in \mathbb{N}$ and for all $z, y_1, \ldots, y_k \in M$ it results:*

$$2k \sum_{i=1}^{k} d(z, y_i)^2 \geq \sum_{i,j=1}^{k} d(y_i, y_j)^2.$$

The following result states a curvature property of the space of probabilities endowed with the 2−Wasserstein distance.

Proposition 5.1.7. *The space $(\mathcal{P}_2(M), W_2)$ has curvature ≥ 0 if and only if M has curvature ≥ 0.*

Proof. As in the proof of Proposition 5.0.21, the set $\mathcal{P}_A(M)$ is dense in $(\mathcal{P}_2(M), W_2)$. Hence, let $\mu, \nu_1, \ldots \nu_k \in \mathcal{P}_2(M)$ and γ_i optimal plans between μ and ν_i; we can assume without restriction that $\mu = \frac{1}{n}\sum_{\alpha=1}^{n} \delta_{z_\alpha}$, $\nu_i = \frac{1}{n}\sum_{\alpha=1}^{n} \delta_{y_i^\alpha}$ and $\gamma_i = \frac{1}{n}\sum_{\alpha=1}^{n} \delta_{(z_\alpha, y_i^\alpha)}$. By 4 of Theorem 5.1.6 we have

$$\sum_{i,j=1}^{k} W_2(\nu_i, \nu_j)^2 \leq \frac{1}{n} \sum_{i,j=1}^{k} \sum_{\alpha=1}^{n} d(y_i^\alpha, y_j^\alpha)^2 \leq$$

$$\leq \frac{2k}{n} \sum_{i=1}^{k} \sum_{\alpha=1}^{n} d(z_\alpha, y_i^\alpha)^2 = 2k \sum_{i=1}^{k} W_2(\mu, \nu_i)^2.$$

For the converse it suffices to note that the function $x \in M \mapsto \delta_x \in \mathcal{P}_2(M)$ is an isometry. □

We refer to [79] for this and other curvature properties of the space of probabilities endowed with the 2−Wasserstein distance.

For spaces of curvature ≤ 0 there is a huge literature rich of monographs devoted exclusively to this kind of spaces, see for example [59, 78] and the reference therein. Here we briefly present some particular features of these spaces.

Proposition 5.1.8. *If M has local curvature ≤ 0 then the geodesics σ are unique and depend continuously with respect to their endpoints.*

Proof. Let $\sigma, \eta : [0,1] \to M$ be two geodesic. Applying (5.5) twice we obtain

$$d(\eta(t),\sigma(t))^2 \leq (1-t)d(\eta(t),\sigma(0))^2 + td(\eta(t),\sigma(1))^2 - t(1-t)d(\sigma(0),\sigma(1))^2 \leq$$
$$\leq (1-t)[(1-t)d(\eta(0),\sigma(0))^2 + td(\eta(1),\sigma(0))^2 - t(1-t)d(\eta(0),\eta(1))^2]+$$
$$+t[(1-t)d(\eta(0),\sigma(1))^2 + td(\eta(1),\sigma(1))^2 - t(1-t)d(\eta(0),\eta(1))^2]-$$
$$-t(1-t)d(\sigma(0),\sigma(1))^2 = (1-t)^2 d(\eta(0),\sigma(0))^2 + t^2 d(\eta(1),\sigma(1))^2+$$
$$+t(1-t)[d(\eta(1),\sigma(0))^2 - d(\eta(0),\eta(1))^2 + d(\eta(0),\sigma(1))^2 - d(\sigma(0),\sigma(1))^2],$$

from which we deduce all the assertions. □

Corollary 5.1.9. *Every metric space of curvature ≤ 0 is locally contractible.*

Proof. If $x \in B(p, r_p)$, by Proposition 5.1.8 the geodesic $\sigma_{x,p}$ connecting x to p is unique and depends continuously with respect to x. Hence we can retract the ball $B(p, r_p)$ to its center p along such geodesic considering the homotopy $H(x,t) = \sigma_{x,p}(t)$. □

The following is a useful comparison with quadrilaterals.

Theorem 5.1.10. *Let (M,d) be a space of curvature ≤ 0 and let $\sigma_1, \sigma_2 : [0,1] \to M$ be geodesics such that*

$$d(\sigma_1(\tau), \sigma_2(t)) \leq \min(r_{\sigma_1(\tau)}, r_{\sigma_2(t)}).$$

Then for every $0 \leq t \leq 1$ and every parameter $0 \leq s \leq 1$ it results:

$$d(\sigma_1(0),\sigma_2(t))^2 + d(\sigma_1(1),\sigma_2(1-t))^2 \leq d(\sigma_1(0),\sigma_2(0))^2 + d(\sigma_1(1),\sigma_2(1))^2+ \tag{5.10}$$
$$+2t^2 d(\sigma_2(0),\sigma_2(1))^2 + t[d(\sigma_1(0),\sigma_1(1))^2 - d(\sigma_2(0),\sigma_2(1))^2]-$$
$$-ts[d(\sigma_1(0),\sigma_1(1)) - d(\sigma_2(0),\sigma_2(1))]^2 - t(1-s)[d(\sigma_1(0),\sigma_2(0)) - d(\sigma_1(1),\sigma_2(1))]^2.$$

Proof. We first consider the case $t = 1$ and $s = 0$. Let us define:

$$a_i := d(\sigma_i(0), \sigma_i(1)),\ i = 1,2$$
$$b_1 := d(\sigma_1(0),\sigma_2(0)),\ b_2 := d(\sigma_1(1),\sigma_2(1)),$$
$$d_1 := d(\sigma_2(0),\sigma_1(1)),\ d_2 := d(\sigma_1(0),\sigma_2(1)).$$

Let also $\delta : [0,1] \to M$ be the geodesic between $\sigma_1(0)$ and $\sigma_2(1)$. For $0 < \lambda < 1$ we put

$$d'_\lambda := d(\sigma_2(0), \delta(\lambda)),\ d''_\lambda := d(\sigma_1(1), \delta(\lambda)).$$

By condition (5.5) we have

$$(d'_\lambda)^2 \leq (1-\lambda)b_1^2 + \lambda a_2^2 - \lambda(1-\lambda)d_2^2,$$
$$(d''_\lambda)^2 \leq \lambda b_2^2 + (1-\lambda)a_1^2 - \lambda(1-\lambda)d_2^2.$$

For $\varepsilon > 0$, since the curvature is ≤ 0 and using the elementary inequality

$$(a+b)^2 \leq (1+\varepsilon)a^2 + (1+\frac{1}{\varepsilon})b^2,$$

we deduce

$$d_1^2 \leq (d'_\lambda + d''_\lambda)^2 \leq (1+\varepsilon)(d'_\lambda)^2 + (1+\frac{1}{\varepsilon})(d''_\lambda)^2 \leq$$

$$\leq (1+\varepsilon)(1-\lambda)b_1^2 + (1+\varepsilon)\lambda a_2^2 + (1+\frac{1}{\varepsilon})\lambda b_2^2 +$$
$$+(1+\frac{1}{\varepsilon})(1-\lambda)a_1^2 - (2+\varepsilon+\frac{1}{\varepsilon})\lambda(1-\lambda)d_2^2.$$

Choosing $\varepsilon = \frac{1-\lambda}{\lambda}$ we obtain

$$d_2^2 + d_1^2 \leq a_1^2 + a_2^2 + \frac{1-\lambda}{\lambda}b_1^2 + \frac{\lambda}{1-\lambda}b_2^2.$$

Taking $\lambda = \frac{b_1}{b_1+b_2}$ we have

$$d_1^2 + d_2^2 \leq a_1^2 + a_2^2 + b_1^2 + b_2^2 - (b_1 - b_2)^2, \tag{5.11}$$

which is the required inequality. Changing the role of a_i and b_i we also obtain the inequality for $t = 1, s = 1$, namely

$$d_1^2 + d_2^2 \leq a_1^2 + a_2^2 + b_1^2 + b_2^2 - (a_1 - a_2)^2.$$

Taking convex combination we deduce the inequality for $t = 1, 0 \leq s \leq 1$:

$$d_1^2 + d_2^2 \leq a_1^2 + a_2^2 + b_1^2 + b_2^2 - s(a_1 - a_2)^2 - (1-s)(b_1 - b_2)^2.$$

Finally, if $0 \leq t \leq 1$ by (5.5) we have

$$d(\sigma_1(0), \sigma_2(t))^2 + d(\sigma_1(1), \sigma_2(1-t))^2 \leq$$
$$\leq (1-t)b_1^2 + td_2^2 - t(1-t)a_2^2 + (1-t)b_2^2 + td_1^2 - t(1-t)a_2^2 =$$
$$= b_1^2 + b_2^2 + 2t^2a_2^2 - 2ta_2^2 + td_1^2 + td_2^2 - tb_1^2 - tb_2^2 \leq$$
$$\leq b_1^2 + b_2^2 + 2t^2a_2^2 - t(a_2^2 - a_1^2) - ts(a_1 - a_2)^2 - t(1-s)(b_1 - b_2)^2.$$

□

Corollary 5.1.11. *Let (M,d) a space with global curvature ≤ 0. Then we have the following quadruple comparison:*

$$\forall x_1, x_2, x_3, x_4 \in M : d(x_1,x_2)^2 + d(x_3,x_4)^2 \leq d(x_1,x_3)^2 + d(x_2,x_4)^2 + d(x_1,x_4)^2 + d(x_2,x_3)^2.$$

Proof. Take $\sigma_2(0) = x_1,\ \sigma_1(1) = x_2,\ \sigma_2(0) = x_1,\ \sigma_1(0) = x_3,\ \sigma_2(1) = x_4$, in (5.11). □

Theorem 5.1.12. *In the assumptions of Theorem 5.1.10 it results:*

$$d(\sigma_1(t), \sigma_2(t))^2 \leq (1-t)d(\sigma_1(0), \sigma_2(0))^2 + d(\sigma_1(1), \sigma_2(1))^2 - \tag{5.12}$$
$$-t(1-t)\{s[d(\sigma_1(0), \sigma_1(1)) - d(\sigma_2(0), \sigma_2(1))]^2 + (1-s)[d(\sigma_1(0), \sigma_2(0)) - d(\sigma_1(1), \sigma_2(1))]^2\},$$

and furthermore

$$d(\sigma_1(t), \sigma_2(t)) \leq (1-t)d(\sigma_1(0), \sigma_2(0)) + td(\sigma_1(1), \sigma_2(1)). \tag{5.13}$$

Proof. As in the proof of (5.5) it is sufficient to show the inequality for $t = \frac{1}{2}$. Set

$$e_1 := d(\sigma_1(0), \sigma_2(\frac{1}{2})),\ e_2 := d(\sigma_1(1), \sigma_2(\frac{1}{2})).$$

With the notation of the preceding proof and by (5.5) we have

$$d(\sigma_1(\frac{1}{2}), \sigma_2(\frac{1}{2}))^2 \leq \frac{1}{2}e_1^2 + \frac{1}{2}e_2^2 - \frac{1}{4}a_1^2.$$

By (5.10) we deduce

$$e_1^2 + e_2^2 \leq b_1^2 + b_2^2 + \frac{1}{2}a_1^2 - \frac{1}{2}s(a_1 - a_2)^2 - \frac{1}{2}(1-s)(b_1 - b_2)^2.$$

Hence

$$d(\sigma_1(\frac{1}{2}), \sigma_2(\frac{1}{2}))^2 \leq \frac{1}{2}b_1^2 + \frac{1}{2}b_2^2 - \frac{1}{4}s(a_1 - a_2)^2 - \frac{1}{4}(1-s)(b_1 - b_2)^2.$$

Taking $s = 1$ in (5.12) we obtain

$$d(\sigma_1(t), \sigma_2(t))^2 \leq$$

$$\leq (1-t)d(\sigma_1(0), \sigma_2(0))^2 + td(\sigma_1(1), \sigma_2(1))^2 - t(1-t)[d(\sigma_1(0), \sigma_2(0)) - d(\sigma_1(1), \sigma_2(1))]^2 =$$

$$= [(1-t)d(\sigma_1(0), \sigma_2(0)) + td(\sigma_1(1), \sigma_2(1))]^2.$$

□

Definition 5.1.13. *A subset $K \subset M$ is called convex if for every geodesic $\sigma : [0,1] \to M$ such that $\sigma(0), \sigma(1) \in K$ it results $\sigma([0,1]) \subset K$.*

A function $f : M \to \mathbb{R}$ is called convex if it is convex along geodesics, i.e. for $0 \leq t \leq 1$ it results:

$$f(\sigma(t)) \leq (1-t)f(\sigma(0)) + tf(\sigma(1)).$$

The function f is said to be uniformly convex if there exists a strictly increasing function α : $[0, +\infty[\to [0, +\infty[$ such that

$$f(\sigma(t)) \leq (1-t)f(\sigma(0)) + tf(\sigma(1)) - t(1-t)\alpha(d(\sigma(0), \sigma(1))).$$

If $(M_1, d_1), (M_2, d_2)$ are two metric space, we define the product metric space $(M_1 \times M_2, d)$ with the distance defined by

$$d((x_1, y_1), (x_2, y_2)) := \left(d_1(x_1, x_2)^2 + d_2(y_1, y_2)^2\right)^{\frac{1}{2}}.$$

If we define

$$epi(f) := \{(x, r) \in M \times \mathbb{R} \mid f(x) \leq r\} \subset M \times \mathbb{R},$$

as usual one can check that

1. f is convex if and only if $epi(f)$ is convex,
2. f is l.s.c. if and only if $epi(f)$ is closed.

According to these definitions by Theorem 5.1.12 it follows that if M has curvature ≤ 0 the distance function $d : M \times M \to \mathbb{R}$ is convex. The convexity of the geodesic distance on Riemannian manifold is another equivalent definition for the constant sectional curvature ≤ 0. Metric spaces with convex metric are also called Busemann spaces. By Theorem 5.1.12 spaces of curvature ≤ 0 in the sense of Alexandrov are Busemann spaces. However, the Busemann condition is a weaker condition on metric spaces, see for example [59].

Proposition 5.1.14. *Let $f : M \to \mathbb{R}$ a uniformly convex, l.s.c. function bounded from below. Then f admits a unique minimum point $x \in M$, which we will denote by*

$$x = argmin_{z \in M} f(z).$$

Proof. Let x_n be a minimizing sequence and take $x_{n,k}$ a midpoint between x_n and x_k. By uniformly convexity we have

$$\inf_M f := c \leq f(x_{n,k}) \leq \frac{1}{2}f(x_n) + \frac{1}{2}f(x_k) - \frac{1}{4}\alpha(d(x_n, x_k)).$$

Hence (x_n) is a Cauchy sequence. Therefore, there exists a point $x = \lim_{n\to+\infty} x_n$. By semicontinuity x is a minimum point. If x, y are two minimum points, if we suppose that $x \neq y$ then taking a midpoint z we would have

$$f(x) \leq f(z) < \frac{1}{2}f(x) + \frac{1}{2}f(y) = f(x).$$

Hence $x = y$. □

Proposition 5.1.15. *Let (M, d) a space with global curvature ≤ 0. If $K \subset M$ is a closed and convex subset then there exists a projection map $\pi_K : M \to K$ such that*

$$\forall z \in M : d(\pi_K(z), z) = \inf_{w\in K} d(w, z).$$

Moreover π_K is "orthogonal":

$$\forall x \in M, \forall y \in K : d(\pi_K(x), x)^2 + d(\pi_K(x), y)^2 \leq d(x, y)^2,$$

and it is a contraction:

$$\forall x, y \in M : d(\pi_K(x), \pi_K(y))^2 \leq d(x, y)^2.$$

Proof. As a subset, K has global curvature ≤ 0. By (5.5) it results that the function $d(z, \cdot)^2$ is uniformly convex on K and therefore it admits a unique minimum point $\pi_K(z)$. If σ is a geodesic connecting $\pi_K(x)$ to y, by (5.5) and since $\sigma(t) \in K$ by convexity we deduce

$$d(\pi_K(x), x)^2 \leq d(\sigma(t), x)^2 \leq (1-t)d(\pi_K(x), x)^2 + td(x, y)^2 - t(1-t)d(\pi_K(x), y)^2 \Rightarrow$$

$$\Rightarrow d(\pi_K(x), x)^2 + (1-t)d(\pi_K(x), y)^2 \leq d(x, y)^2,$$

From which we obtain the "orthogonality". By corollary 5.1.11 and using "orthogonality" we have

$$d(x, y)^2 + d(\pi_K(x), \pi_K(y))^2 + d(\pi_K(x), x)^2 + d(\pi_K(y), y)^2 \geq d(\pi_K(y), x)^2 + d(\pi_K(x), y)^2 \geq$$

$$\geq d(\pi_K(x), x)^2 + d(\pi_K(x), \pi_K(y))^2 + d(\pi_K(y), y)^2 + d(\pi_K(x), \pi_K(y))^2 =$$

$$= d(\pi_K(x), x)^2 + d(\pi_K(y), y)^2 + 2d(\pi_K(x), \pi_K(y))^2 \Rightarrow$$

$$\Rightarrow d(\pi_K(x), \pi_K(y))^2 \leq d(x, y)^2.$$

□

5.2 Barycenter maps

In this section we describe a canonical construction of barycenter maps for metric space M of global curvature ≤ 0. Here for barycenter map we mean a function

$$\beta : \mathcal{P}_1(M) \to M,$$

which satisfy the conditions:

1. $\forall x \in M : \beta(\delta_x) = x$,
2. $\beta \in \mathrm{Lip}_1(\mathcal{P}_1(M), M)$ with respect to the 1-Wasserstein distance.

The attempt to find barycenter maps on more general spaces leads to the counterexample discussed in section 5.3.

Theorem 5.2.1 (Barycenters). *Let (M, d) a space with global curvature ≤ 0. If $\mu \in \mathcal{P}_1(M)$ and $y \in M$ are fixed, then there exists a unique point $b(\mu) \in M$, independent of the choice of y, such that*

$$b(\mu) = argmin_{z \in M} \int_M [d(z,x)^2 - d(y,x)^2] d\mu(x).$$

Proof. Let $F_y(z) = \int_M [d(z,x)^2 - d(y,x)^2] d\mu(x)$. Then

$$|F_y(z)| = |\int_M [d(z,x) - d(y,x)][d(z,x) + d(y,x)] d\mu(x)| \leq$$

$$\leq d(z,y)[\int_M d(z,x) d\mu(x) + \int_M d(y,x) d\mu(x)] < +\infty,$$

and F_y is certainly bounded from below. If $\sigma : [0,1] \to M$ is a geodesic, by (5.5) we have

$$F_y(\sigma(t)) = \int_M [d(\sigma(t),x)^2 - d(y,x)^2] d\mu(x) \leq$$

$$\leq \int_M [(1-t)d(\sigma(0),x)^2 + t d(\sigma(1),x)^2 - t(1-t)d(\sigma(0),\sigma(1))^2 - d(y,x)^2] d\mu(x) \leq$$

$$\leq (1-t)\int_M [d(\sigma(0),x)^2 - d(y,x)^2] d\mu(x) + t\int_M [d(\sigma(1),x)^2 - d(y,x)^2] d\mu(x) - t(1-t)d(\sigma(0),\sigma(1))^2 =$$

$$= (1-t)F_y(\sigma(0)) + tF_y(\sigma(1)) - t(1-t)d(\sigma(0),\sigma(1))^2.$$

Hence F_y is uniformly convex. Furthermore, since $|F_y(z) - F_y(z')| \leq \int_M |d(z,x)^2 - d(z',x)^2| d\mu(x)$, it follows that F_y is continuous. Thus, by Proposition 5.1.14 F_y admits a unique minimum point $b(\mu)$. Observe that $F_y(z) - F_{y'}(z) = \int_M [d(y',x)^2 - d(y,x)^2] d\mu(x) = c$, so that the minimum points of F_y and $F_{y'}$ are the same. □

Proposition 5.2.2. *If $M = \mathbb{R}^n$, or in general an Hilbert space, then for each $\mu \in \mathcal{P}_1(\mathbb{R}^n)$ it results:*

$$b(\mu) = \int_{\mathbb{R}^n} x d\mu.$$

Proof. $v := b(\mu)$ is the unique minimizer of $F(z) = \int_{\mathbb{R}^n} [|z-x|^2 - |x|^2] d\mu(x)$. Taking the variation $f(\varepsilon) = F(v + \varepsilon y)$, f has a minimum point at 0. Hence

$$f'(0) = \lim_{\varepsilon \to 0} \frac{f(\varepsilon) - f(0)}{\varepsilon} = \lim_{\varepsilon \to 0} \int_{\mathbb{R}^n} \frac{1}{\varepsilon} \left[|v - x + \varepsilon y|^2 - |v-x|^2 \right] d\mu(x) = 2\int_{\mathbb{R}^n} \langle y, v - x \rangle d\mu(x) = 0.$$

By the arbitrariness of y we deduce that $b(\mu) = \int_{\mathbb{R}^n} x d\mu$. □

Proposition 5.2.3. *Let (M, d) a space of global curvature ≤ 0. If $\mu \in \mathcal{P}_1(M)$ is supported on a closed and convex subset K then $b(\mu) \in K$.*

Proof. By Proposition 5.1.15 we have

$$\int_M [d(b(\mu),x)^2 - d(y,x)^2] d\mu(x) \geq \int_M [d(\pi_K(b(\mu)),x)^2 - d(y,x)^2] d\mu(x).$$

By the uniqueness of the barycenter we deduce that $b(\mu) = \pi_K(b(\mu)) \in K$. □

Theorem 5.2.4 (Jensen's inequality). *Let (M,d) be a space of global curvature ≤ 0 and let $f : M \to \mathbb{R}$ be a convex and l.s.c. function. Then for every $\mu \in \mathcal{P}_1(M)$ such that $\int_M |f| d\mu < +\infty$ we have*

$$f(b(\mu)) \leq \int_M f(x) d\mu.$$

Proof. Consider the product space $\widetilde{M} = M \times \mathbb{R}$. Observe that $\widetilde{M}$ is a complete, geodesic space of global curvature ≤ 0 as well. Let $K = epi(f)$ which is a closed and convex subset of $\widetilde{M}$. Since $\int_M |f| d\mu < +\infty$, it follows that the measure $\tilde{\mu} = (Id \times f)_\sharp \mu \in \mathcal{P}_1(\widetilde{M})$. Indeed, if $\tilde{y} = (y, f(y))$ we have

$$\int_{\widetilde{M}} \tilde{d}(\tilde{x}, \tilde{y}) d\tilde{\mu} \leq \int_M [d(x,y) + |f(x) - f(y)|] \, d\mu(x) < +\infty.$$

We claim that $b(\tilde{\mu}) = (b(\mu), \int_M f d\mu)$. Indeed, if we take $\tilde{y} = ((b(\mu), \int_M f d\mu)$, $b(\tilde{\mu})$ can be obtained minimizing the function

$$\int_{\widetilde{M}} [\tilde{d}(\tilde{z}, \tilde{x})^2 - \tilde{d}(\tilde{y}, \tilde{x})^2] \tilde{\mu}(\tilde{x}) =$$

$$= \int_M [d(z,x)^2 - d(b(\mu)), x)^2] d\mu(x) + \int_M \left[|t - f(x)|^2 - | \int_M f d\mu - f(x)|^2 \right] d\mu(x) \geq 0,$$

since $\int_M f d\mu$ is the barycenter of the measure $f_\sharp \mu$ by Proposition 5.2.2. Hence the minimum is attained just at the point $\tilde{y}$. Since $(x, f(x)) \in epi(f)$ then $supp(\mu) \subset (Id \times f)^{-1} epi(f)$. Hence

$$\tilde{\mu}(\widetilde{M} \setminus epi(f)) = \mu \left((Id \times f)^{-1} (\widetilde{M} \setminus epi(f)) \right) = \mu(M \setminus (Id \times f)^{-1} epi(f)) \leq \mu(M \setminus supp(\mu)) = 0.$$

Therefore we have that

$$supp(\tilde{\mu}) \subset epi(f) = K.$$

By Proposition 5.2.3 it follows that $b(\tilde{\mu}) \in K$ and hence $f(b(\mu)) \leq \int_M f d\mu$. □

Proposition 5.2.5 (Contraction property). *Let (M,d) a space of global curvature ≤ 0. Then it results:*

$$\forall \mu, \nu \in \mathcal{P}_1(M) : d(b(\mu), b(\nu)) \leq W_1(\mu, \nu).$$

Proof. Let γ an optimal transport plan between μ and ν. As in the proof of Theorem 5.2.4 one can check that $b(\gamma) = (b(\mu), b(\nu))$. Applying Jensen's inequality to the convex function $d : M \times M \to \mathbb{R}$ we obtain

$$d(b(\mu), b(\nu)) = d(b(\gamma)) \leq \int_{M \times M} d(x,y) d\gamma = W_1(\mu, \nu).$$

□

Putting together Theorem 5.2.1 with Proposition 5.2.5 we can conclude that the function $\beta : \mathcal{P}_1(M) \to M$ defined by $\beta(\mu) = b(\mu)$ is a barycenter map on spaces of global curvature ≤ 0.

Actually, we point out that the existence of barycenter maps produces some restrictions on the geometry of the space M.

- If $x, y \in M$ we can define the curve $\sigma(t) := \beta((1-t)\delta_x + t\delta_y)$ which turns out to be a geodesic because of Proposition 5.2.5. Furthermore the distance is convex along such geodesics. Hence, if geodesics on M are unique then M is a Busemann space.

- If M is a Riemannian manifold then M must have sectional curvature ≤ 0. Indeed, if $M_0 \subset M$ it also admits a barycenter map. Choosing M_0 sufficiently small, geodesics on M_0 are unique and then the distance is convex on M_0. This implies that M has sectional curvature ≤ 0.

We also remark that by Proposition 5.1.9 Riemannian manifolds with global curvature ≤ 0 are contractible and then with trivial homology. Hence, to study non trivial homology we have to look for manifolds with local curvature ≤ 0. However, there is no canonical way to define barycenters on such spaces. Since we can always define barycenters on the set Δ of Dirac deltas by $\beta(\delta_x) = x$, certainly satisfying the contraction property of Proposition 5.2.5, a possibility is to try to extend the map β to the whole space $\mathcal{P}_1(M)$ in a Lipschitz way.

5.3 A counterexample to the extension of Lipschitz maps

The Lipschitz extension of maps into metric spaces is a difficult matter. Of course, if the target space is $\mathbb{R}$ the Lipschitz extension is always possible. Indeed, a function $f \in \mathrm{Lip}_l(A, \mathbb{R})$ can be extended to the whole space (E, d), with the same Lipschitz constant, taking the Mc Shane's formula:

$$\tilde{f} = \sup_{y \in A} \{f(y) - ld(x, y)\}.$$

The function $\tilde{f}$ defined above is in fact the smallest possible extension; the largest one is given by

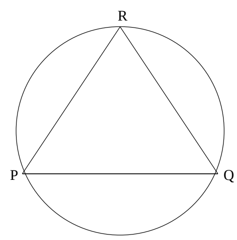
$$\hat{f} = \inf_{y \in A} \{f(y) + ld(x, y)\}.$$

If $f \in \mathrm{Lip}_l(A, \mathbb{R}^m)$ one can argue componentwise, but in general enlarging the Lipschitz constant. Preserving the Lipschitz constant is a more difficult question. In the case of $f : A \subset E \to F$ with $E = \mathbb{R}^k$, $F = \mathbb{R}^m$ this can be done by the classical Kirszbraun theorem, see for example [45]. More generally, by a projection argument, the same is true if F is an Hilbert space. Kirszbraun theorem also holds for certain metric spaces with bounded curvature. The following statement, adapted to the case of curvature ≤ 0 and ≥ 0, is taken from [60].

Theorem 5.3.1 (General Kirszbraun theorem). *Let E, F geodesics metric spaces such that E has curvature ≥ 0, while F is complete with global curvature ≤ 0. If $A \subset E$, every $f \in \mathrm{Lip}(A, F)$ admits a Lipschitz extension with the same Lipschitz constant.*

We remark that this statement actually holds whenever the curvatures are bounded by a constant $k \in \mathbb{R}$ instead of 0, provided that $diam(f(A)) \leq \frac{1}{2}D_k$, where $D_k = \frac{\pi}{\sqrt{k}}$ if $k > 0$, $D_k = +\infty$ otherwise. The necessity of the bound on the diameter of the image for positive curvature can be illustrated by the following example:

Example 5.3.2. *Consider $A = \{P, Q, R\}$ where the points P, Q, R are the vertices of a regular triangle inscribed into S^1, see figure 5.3.2. We can consider S^1 as the intersection of S^2 with the*



equatorial plane. Let $f : A \subset S^2 \to S^1$ be the identity map. Observe that $diam f(A) = \frac{2}{3}\pi > \frac{1}{2}D_1 = \frac{\pi}{2}$. Let us now suppose to have a 1-Lipschitz extension $\tilde{f}$ to $A \cup \{N\}$, with N the north pole of S^2. Then there exists a point $S \in A$ such that

$$\frac{2}{3}\pi < d(\tilde{f}(N), S) = d(\tilde{f}(N), \tilde{f}(S)) \leq d(N, S) = \frac{\pi}{2}.$$

Hence there is no 1-Lipschitz extension to $A \cup \{N\}$.

For related results of extension of Lipschitz maps on Banach spaces see [4, 30, 55, 66].

Theorem 5.3.3. *Let (M, d) be a non-contractible metric space and $\Delta \subset \mathcal{P}_p(M)$ be the set of Dirac deltas. Then the 1-Lipschitz map*

$$\beta : \Delta \to M$$

defined by $\beta(\delta_x) = x$ cannot be extended in a Lipschitz way to the whole $\mathcal{P}_p(M)$.

Proof. Suppose by contradiction that there exists a continuous extension $\beta : \mathcal{P}_p(M) \to M$. Consider a continuous curve $\gamma : [0,1] \to M$ not homotopic to a constant. Let $i : M \to \mathcal{P}_p(M)$ be the canonical isometry defined by $i(x) = \delta_x$. Take the continuous function $F : [0,1] \times [0,1] \to \mathcal{P}_p(M)$ defined by (5.2) with $\mu_t = i(\gamma(t))$ and consider the continuous composition $\beta \circ F : [0,1] \times [0,1] \to M$. We have

$$\beta \circ F(t, 0) = \beta(i(\gamma(t)) = \gamma(t),$$

$$\beta \circ F(t, 1) = \beta(\mu_0),$$

which is a contradiction since the curve γ is not homotopic to a constant. $\square$

The following example shows that in Theorem 5.3.1 the assumption of target space of global curvature ≤ 0 cannot be reduced to a local one.

Example 5.3.4. *Consider the flat torus $\mathbb{T}^n$. By Proposition 5.1.7 the space $(\mathcal{P}(\mathbb{T}^n), W_2)$ has curvature ≥ 0. Of course, $\mathbb{T}^n$ is non-contractible and it has local curvature ≤ 0. Therefore, by Theorem 5.3.3 we infer that the 1-Lipschitz map*

$$\beta : \Delta \to \mathbb{T}^n$$

defined by $\beta(\delta_x) = x$ cannot be extended in a Lipschitz way to the whole $\mathcal{P}(\mathbb{T}^n)$.

If M is a manifold the same example can be done using the isometric embedding of M in $\mathbb{R}^k$ for k big enough. The example 5.3.4 however holds for general metric space M. Moreover M is embedded in the space of probability measures equipped with the Wasserstein distance which may have non trivial curvature (see [79]).

Chapter 6

Further directions

The concave case

Recently, currents revealed useful also to study transportation with concave costs. This kind of problems come out from the so-called "irrigation problems". Actually, on Monge-Kantorovich problem the masses can split only at the beginning, while in many situation of interest the configurations of the transport are "tree" shaped unlike that "V" shaped, as for instance in the delta's of rivers or in the circulation of the blood.

In [83] this problem is treated defining a "concave" mass for currents. More precisely, let $K \subset \mathbb{R}^n$ be a compact set. We define the flat norm for the 1-current T as

$$F(T) := \sup \{T(\omega) \mid \omega \in \Gamma^\infty(K) \text{ such that } \|\omega\|_\infty \leq 1, \|d\omega\|_\infty \leq 1\}.$$

The space of 1-flat chains, which we denote by $\mathcal{F}_1(K)$, is the closure of $\mathcal{N}_1(K)$ with respect to the flat norm.

It is known, see for example [45, 68], that flat chains can be approximated by 1-polyhedral chains in the flat norm. A 1-polyhedral chains is a linear combination of Lipschitz curves. Given $P = \sum_{i=1}^n a_i \gamma_i$ a polyhedral chain we can define:

$$M_\alpha(P) := \sum_{i=1}^n (a_i)^\alpha length(\gamma_i),$$

and by semicontinuity we can define for any $T \in \mathcal{F}_1(K)$

$$M_\alpha(T) := \inf_{P_i \xrightarrow{F} T} \liminf_{i \to +\infty} M_\alpha(P_i).$$

Flat chains with $M_\alpha(T) < +\infty$ are in fact rectifiable, in the sense that they can be represented by

$$T(\omega) = \int_A \theta(x) \langle \omega(x), \zeta(x) \rangle d\mathcal{H}^1,$$

where A is a 1-rectifiable subset of K, while θ is a $\mathcal{H}^1 \llcorner A$ integrable positive function and $\zeta(x)$ is a unit tangent vector field on A. See [84] for details. Furthermore it results

$$M_\alpha(T) := \int_A \theta^\alpha d\mathcal{H}^1.$$

In this setting, given $\mu, \nu \in \mathcal{P}(K)$, the transportation problem amounts to

$$\min \{M_\alpha(T) \mid T \in \mathcal{F}_1(K),\ \partial T = \mu - \nu\}. \tag{6.1}$$

If $\alpha > 1 - \frac{1}{n}$, problem (6.1) admits always a solution with $M_\alpha(T) < +\infty$.

Another approach to irrigation is given in [14] where the problem is formulated in terms of curves in the Wasserstein space $(\mathcal{P}(K), W_p), p \geq 1$. For $\mu \in \mathcal{P}(K)$ we set

$$g_\alpha(\mu) := \begin{cases} \sum_{n\in\mathbb{N}} (a_n)^\alpha & \text{if } \mu = \sum_{n\in\mathbb{N}} a_n \delta_{x_n}, \\ +\infty & \text{otherwise.} \end{cases}$$

Then, for a Lipschitz curve $\rho : [0,1] \to (\mathcal{P}(K), W_p)$ we can define

$$M_\alpha(\rho) := \int_0^1 g_\alpha(\rho(t)) |\dot{\rho}|(t) dt.$$

Hence we are considering a sort of $M_\alpha(T_\rho)$ for the metric current T_ρ associated to the Lipschitz curve ρ (see [4] for the theory of currents in a metric space setting). Given $\mu, \nu \in \mathcal{P}(K)$, the corresponding transport problem is the following

$$\min \{ M_\alpha(\rho) \mid \rho \in \mathrm{Lip}([0,1], (\mathcal{P}(K), W_p)),\ \rho(0) = \mu,\ \rho(1) = \nu \}. \tag{6.2}$$

In [14] it is shown that problem (6.2) admits always a solution with $M_\alpha(\rho) < +\infty$ if $\alpha > 1 - \frac{1}{n}$, while if $\alpha \leq 1 - \frac{1}{n}$ there exists a measure $\mu \in \mathcal{P}(K)$ such that for every non-constant Lipschitz curve $\rho : [0,1] \to (\mathcal{P}(K), W_p)$ such that $\rho(0) = \mu$ it results $M_\alpha(\rho) = +\infty$.

For an alternative approach to irrigation problems we refer to [61] and the references therein. Actually, at the moment the connections among these different approaches are not completely clear.

Appendix A

Normal euclidean 1-currents

The concept of current was introduced by Federer and Fleming in 1960. We refer to [45, 49, 68, 75] for an exthensive study of the subject. Here we restrict ourselves to introduce some basic properties of one-dimensional currents.

Let $U \subset \mathbb{R}^n$ be an open set. We denote by $\Gamma^\infty(U)$ the set of compactly supported smooth 1-forms on U, endowed with the norm $\|\omega\|_\infty = \sup_U |\omega(x)|$. The space of 1-dimensional currents, which we denote by $\mathcal{D}_1(U)$, is the space of linear functionals on $\Gamma^\infty(U)$ which are continuous with respect to the $\mathcal{C}^\infty$ topology. Analougously, for any $k \geq 0$ a k-current will be a linear and continuous functional on the k-forms.

The boundary of a current is defined by duality with the differential through the formula:

$$\partial T(\omega) = T(d\omega).$$

The subspace of the currents $T \in \mathcal{D}_1(U)$ such that $T(\omega) = 0$ whenever ω is an exact form is the space of closed 1-currents and it is denoted by $\mathcal{D}_1^c(U)$. Then for a current T, being closed is equivalent to $\partial T = 0$.

Definition A.0.5. *The support of a current T is defined by*

$$spt(T) = \bigcap \{K \subset U \mid T(\omega) = 0 \ \forall \omega \in \Gamma^\infty(U) \text{ with } spt(\omega) \subset U \setminus K\}.$$

We say that a sequence $\{T_h\} \subset \mathcal{D}_1(U)$ weakly converges to $T \in \mathcal{D}_1(U)$, and we denote by $T_h \rightharpoonup T$, if T_h pointwise converge to T, i.e.

$$\lim_{h \to +\infty} T_h(\omega) = T(\omega) \quad \forall \omega \in \Gamma^\infty(U).$$

The mass of a current is defined as follows:

$$M(T) = \sup\{T(\omega) \mid \|\omega\|_\infty \leq 1\}.$$

We define the set of normal 1-currents as

$$\mathcal{N}_1(U) := \{T \in \mathcal{D}_1(U) \mid M(T) < +\infty, M(\partial T) < +\infty\},$$

and by $\mathcal{N}_1^c(U)$ the set of normal 1-currents which are closed.

Proposition A.0.6 (Lower semicontinuity of the mass). *Let $T_h, T \in \mathcal{D}_1(U)$ such that $T_h \rightharpoonup T$. Then*

$$M(T) \leq \liminf_{h \to +\infty} M(T_h).$$

Proof. Let $\omega \in \Gamma^\infty(U)$ with $\|\omega\|_\infty \leq 1$. We have

$$T(\omega) = \lim_{h\to+\infty} T_h(\omega) \leq \liminf_{h\to+\infty} M(T_h).$$

Taking the supremum with respect to ω we are done. □

Currents with finite mass are representable by integration as showed in the following

Theorem A.0.7 (Representation by integration). *Let $T \in \mathcal{D}_1(U)$ such that $M(T) < +\infty$. Then there exists a probability measure $\sigma \in \mathcal{P}(U)$ and a Borel vector field $X : U \to \mathbb{R}^n$ such that*

1. $|X(x)| = M(T) \quad \sigma - a.e.\ x \in U$;
2. $\forall \omega \in \Gamma^\infty(U) : \ T(\omega) = \int_U \langle \omega, X\rangle d\sigma$;
3. $\forall B \in \mathcal{B}(U) : \sigma(B) = \frac{1}{M(T)} \sup\{T(\omega) \mid \omega \in \Gamma^\infty(U),\ \|\omega\|_\infty \leq 1, spt(\omega) \subset B\}$.

Proof. The current T can be seen as a linear continuous functional on $\mathcal{C}_c(U, \mathbb{R}^n)$ with respect to the norm $\|\omega\|_\infty$. Hence, by the Riesz representation theorem (see for instance [2, 6]) we infer the existence of a measure $\mu = (\mu_1, \dots, \mu_n) \in \mathcal{M}^n(U)$ such that $T(\omega) = \sum_{i=1}^n \int_U \omega_i d\mu_i$. By polar decomposition we can write $\mu = f|\mu|$. with $f \in L^1_{|\mu|}(U, \mathbb{R}^n)$ with $|f| = 1$ $|\mu|$- a.e. Therefore we have

$$T(\omega) = \int_U \langle \omega, f\rangle d|\mu|.$$

Observe that $M(T) = \|\mu\|$. Moreover it results

$$|\mu|(B) = \sup\{T(\omega)\ :\ \|\omega\|_\infty \leq 1, spt(\omega) \subset B\}.$$

Indeed, a first inequality follows from the estimate:

$$T(\omega) = \int_U \langle \omega, f\rangle d|\mu| \leq |\mu|(B).$$

On the other hand, using Lusin's theorem, we can approximate $f\chi_B$ in $L^1_{|\mu|}$ by $w_h \in \Gamma^\infty(U)$ such that $\|\omega_h\|_\infty \leq 1$. We have

$$|\mu|(B) = \int_B |f|^2 d|\mu| = \lim_{h\to+\infty} \int_U \langle \omega_h, f\rangle d|\mu|,$$

from which we obtain the other inequality. Taking $\sigma = \frac{1}{M(T)}|\mu|$ and $X = M(T)f$ all the claims follow. □

Whenever a current satisfy the condition

$$T(\omega) = \int_M \langle \omega, X\rangle d\sigma,$$

we briefly write $T = \sigma \wedge X$. We denote the representation given in Theorem A.0.7 by the underscript s which stands for "special" so that, for every 1-currents T with finite mass we have $T = X_s \wedge \sigma_s$.

By approximation, every 1-current $T = \sigma \wedge X$ of finite mass can be extended to 1-forms with Borel bounded coefficients, again satisfying

$$T(\omega) = \int_U \langle \omega, X\rangle d\sigma.$$

In particular, if $B \subset U$ is any Borel set we can define the restriction of T on B, $T \llcorner B$, by

$$T \llcorner B(\omega) = \int_B \langle \omega, X\rangle d\sigma.$$

We point out that there are many possible choices for the representation as in Theorem A.0.7. For example, given a representation $\sigma \wedge X$ of a current T, we can obtain another representation (which in special case will coincide with the first one) as follows: choose a function $c \in L^1_\sigma$ such that $\int_M c d\sigma = 1$ and $c(x) \geq \delta > 0$, then $\sigma_1 = c(x)\sigma$ is a probability measure and $X_1 = \frac{X}{c}$ is a vector field such that $\sigma_1 \wedge X_1 = T$.

Actually, the following lemma gives a more general vision of this last fact:

Lemma A.0.8. *Let $T \in \mathcal{D}_1(U)$ with finite mass such that $T = \sigma \wedge X$. Then $\sigma_s << \sigma$ and $X = \frac{\partial\sigma_s}{\partial\sigma} X_s$.*

Proof. Let $B \subset U$ a Borel set. We have

$$\sigma(B) = 0 \Rightarrow T \llcorner B = 0 \Rightarrow \int_B |X_s| d\sigma_s = 0 \Rightarrow \sigma_s(B) = 0.$$

Hence $\sigma_s = \frac{\partial\sigma_s}{\partial\sigma}\sigma$. Since $T = \sigma \wedge X = \sigma_s \wedge X_s$, we infer that

$$X_s\sigma_s = X\sigma \Rightarrow X_s \frac{\partial\sigma_s}{\partial\sigma}\sigma = X\sigma$$

and then $X = \frac{\partial\sigma_s}{\partial\sigma}\sigma X_s$. □

Proposition A.0.9 (Compactness). *Let $T_h \in \mathcal{D}_1(U)$ such that $M(T_h) \leq C$. Then, up to subsequences, there exists $T \in \mathcal{D}_1(U)$ such that $T_h \rightharpoonup T$.*

Proof. By Theorem A.0.7 we can write $T_h = X_h \wedge \sigma_h$. In particular we have that $\|X_h\sigma_h\| \leq c$. Since $\mathcal{M}^n(U)$ is the dual space of $\mathcal{C}_o(U, \mathbb{R}^n)$, by Banach-Alaoglu theorem it follows that there exists a measure $\mu \in \mathcal{M}^n(U)$ such that$X_h\sigma_h \overset{*}{\rightharpoonup} \mu$. Hence if we consider the current $T(\omega) := \int_U \omega\mu$, by weak* convergence we have that $T_h \rightharpoonup T$. □

If M is a smooth manifold of finite dimension, by Whitney embedding theorem we may assume that M is a smooth submanifold of some $\mathbb{R}^n$. Moreover we can assume that M has a tubular neighborhood, i.e. there exists an open set $U \subset \mathbb{R}^n$ and a smooth retraction map $\pi : U \to M$ such that

$$\pi(U) = M, \quad \forall y \in M : \pi(y) = y.$$

If $i : M \to \mathbb{R}^n$ is the immersion map we set

$$Null_M(U) := \{\omega \in \Gamma^\infty(U) \mid i^\sharp\omega = 0\},$$

where $i^\sharp\omega$ is the pull-back towards the map i of ω. The space of 1-currents on the manifold M is defined by

$$\mathcal{D}_1(M) := \{T \in \mathcal{D}_1(U) \mid T(\omega) = 0 \ \forall\omega \in Null_M(U)\}.$$

Analogously

$$\mathcal{N}_1(M) := \{T \in \mathcal{N}_1(U) \mid T(\omega) = 0 \ \forall\omega \in Null_M(U)\}.$$

It is possible to check that if $T \in \mathcal{D}_1(M)$ then *spt*$(T) \subset \overline{M}$. But if $T \in \mathcal{D}_1(U)$ and *spt*$(T) \subset M$ it is not true in general that $T \in \mathcal{D}_1(M)$. However we have the following

Theorem A.0.10. *Let M a closed smooth submanifold of $\mathbb{R}^n$. If $T \in \mathcal{N}_1(U)$ and spt$(T) \subset M$, then $T \in \mathcal{N}_1(M)$.*

Therefore normal 1-currents on M may be identified to currents in the ambient space $\mathbb{R}^n$ with support on M. We refer to section 5.3.1 of [49] for more details.

The first example of closed normal 1-current is given by the integration along a closed, Lipschitz curve on the manifold M. However, periodicity is not the only way for a curve to produce a closed current.

Example A.0.11. *Let v be an irrational direction on the torus $\mathbb{T}^n$ and consider the curve $\gamma(t) = x_0 + tv\,(mod.\ \mathbb{Z}^n)$ defined for $t \in \mathbb{R}$ and $x_0 \in \mathbb{T}^n$ fixed. For $k \in \mathbb{N}$, define the normal 1-current associated to the curve $\gamma_{|(-k,k)}$ as follows:*

$$T_k(\omega) = \frac{1}{2k}\int_{-k}^{k} \langle \omega(\gamma(t)), \dot\gamma(t)\rangle dt.$$

Up to a subsequence there exists $T_\infty(\omega) = \lim_{k\to+\infty} T_k(\omega)$ for any 1-form ω on the torus. Since

$$T_\infty(df) = \lim_{k\to+\infty} \frac{f(\gamma(k))) - f(\gamma(-k))}{2k} = 0,$$

T_∞ is a closed current. Moreover, if γ is $[0,1]$-periodic then

$$T_\infty(\omega) = \int_0^1 \langle \omega(\gamma(t)), \dot\gamma(t)\rangle dt.$$

Indeed, for any $k > 0$

$$\begin{aligned} T_k(\omega) &= \frac{1}{2k}\int_{-k}^{k} \langle \omega(\gamma(t)), \dot\gamma(t)\rangle dt = \frac{1}{2k}\sum_{h=1}^{2k}\int_{-k+h-1}^{-k+h} \langle \omega(\gamma(t)), \dot\gamma(t)\rangle dt = \\ &= \frac{1}{2k}\sum_{h=1}^{2k}\int_0^1 \langle \omega(\gamma(t)), \dot\gamma(t)\rangle dt = \int_0^1 \langle \omega(\gamma(t)), \dot\gamma(t)\rangle dt. \end{aligned}$$

It turns out that this example contains the building block of normal closed 1-currents. Proofs of what follows here are contained in [7] and [76], although our exposition is closer to the second reference.

Definition A.0.12. *A current $T \in \mathcal{N}_1^c(M)$ is said an elementary solenoid if there exists a Lipschitz curve $\gamma : \mathbb{R} \to M$ enjoying the following properties:*

1. *$Lip(\gamma) \le M(T)$;*
2. *$T(\omega) = \lim_{k\to\infty} \frac{1}{2k}\int_{-k}^{k} \langle \omega(\gamma(t)), \dot\gamma(t)\rangle dt$, for any 1-form ω;*
3. *$\gamma(\mathbb{R}) \subset \operatorname{spt}(T)$.*

In particular the semicontinuity of the mass together with (1), (2), (3) implies that $|\dot\gamma(t)| = M(T)$ a.e. We denote by $\mathcal{S}$ the set of all elementary solenoids.

The following is Theorem B in [76].

Theorem A.0.13. *Every $T \in \mathcal{N}_1^c(M)$ can be decomposed into elements of $\mathcal{S}$ in the sense that there exists a probability measure ν on $\mathcal{S}$ such that*

$$T = \int_{\mathcal{S}} R d\nu(R) \qquad and \qquad M(T) = \int_{\mathcal{S}} M(R) d\nu(R).$$

Bibliography

[1] L. Ambrosio, Lecture Notes on Transport Problems, in "Mathematical Aspects of Evolving Interfaces", Lecture Notes in Mathematics **1812**, Springer, Berlin, 2003, 1-52.

[2] L. Ambrosio, N. Fusco, D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, Oxford University Press, New York, 2000.

[3] L. Ambrosio, N. Gigli, G. Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures, Lecture Notes in Mathematics, ETH Zürich, Birkhäuser, 2005.

[4] L. Ambrosio, B. Kirchheim, Currents in metric spaces, Acta Mathematica **185** (2000), no. 1, 1-80.

[5] L. Ambrosio, B. Kirchheim, Rectifiable sets in metric and Banach spaces, Mathematische Annalen **318** (2000), no. 3, 527-555.

[6] L. Ambrosio, P. Tilli, Topics on Analysis in Metric Spaces, Oxford Lectures Series in Mathematics and its Applications **25**, Oxford University Press, Oxford, 2004.

[7] V. Bangert, Minimal measures and minimizing closed normal one-currents, GAFA Geom. Funct. Anal. **9** (1999), no. 3, 413–427.

[8] J. Benamou, Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, Numer. Math. **84** (2000), no. 3, 375–393.

[9] P. Bernard, B. Buffoni, Optimal mass transportation and Mather theory, preprint.

[10] T. Bhattacharya, E. Di Benedetto, J. Manfredi, Limits as $p \to +\infty$ of $\Delta_p u_p = f$ and related extremals problems, Rend. Sem. Mat. Univ. Politec. Torino, Nonlinear PDE's, Special Issue (1989), 15-68.

[11] G. Bouchitté, G. Buttazzo, Characterization of optimal shapes and masses through Monge–Kantorovich equation, Journal European Math. Soc. **3** (2001), no. 2, 139–168.

[12] G. Bouchitté, G. Buttazzo, P. Seppecher, Shape optimization solutions via Monge–Kantorovich equation, C. R. Acad. Sci. Paris **324-I** (1997), no. 10, 1185–1191.

[13] G. Bouchitté, G. Buttazzo, L. De Pascale, A p-Laplacian approximation for some mass optimization problems, Journal of Optimization Theory and Applications **118** (2003), no. 1, 1–25.

[14] A. Brancolini, G. Buttazzo, F. Santambrogio, Path functionals over Wasserstein spaces, preprint, available on http://cvgmt.sns.it.

[15] Y. Brenier, Décomposition polaire et réarrangement des champs des vecteurs, C. R. Acad. Sci. Paris. Sér. I Math. **305** (1987), no. 19, 805-808.

[16] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, Comm. Pure Appl. Math. **44** (1991), no. 4, 375-417.

[17] Y. Brenier, A homogenized model for vector sheets, Arch. Rat. Mech. Anal. **138** (1997), 319-353.

[18] Y. Brenier, A Monge-Kantorovich Approach to the Maxwell Equations, Hyperbolic Problems: Theory, Numerics, Applications, Vol. I, II, Internat. Ser. Numer. Math. **140, 141**, Birkhäuser, Basel, 2001, 179-186.

[19] Y. Brenier, Extended Monge-Kantorovich Theory, in "Optimal Transportation and Applications", Lecture Notes in Mathematics **1813**, Springer, Berlin, 2003, 91-121.

[20] Y. Brenier, Extension of the Monge-Kantorovich Theory to Classical Electrodynamics, Recent advances in the theory and applications of mass transport, Contemp. Math., **353**, AMS, Providence, RI, 2004, 19-41.

[21] H. Brezis, Analyse Fonctionelle, Théorie at Applications, Colléction mathématiques appliquées pour la maîtrise, Masson, Paris, 1983.

[22] G. Buttazzo, M. Giaquinta, S. Hildebrandt, One-dimensional Variational Problems, An Introduction, Oxford Lecture Series in Mathematics and its Applications, Vol. **15**, Oxford University Press, New York, 1998.

[23] L. Caffarelli, M. Feldman, R. J. McCann, Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs, J. Amer. Math. Soc. **15** (2002), no. 1, 1-26.

[24] P. Cannarsa, P. Cardaliaguet, Representation of equilibrium solutions to the table problem of growing sand piles, J. Eur. Math. Soc. **6** (2004), no.4, 435-464.

[25] G. Contreras, Action potential and weak KAM solutions. Calc. Var. Partial Differential Equations **13** (2001), no. 4, 427–458.

[26] G. Contreras, R. Iturriaga, Global Minimizers of Autonomous Lagrangians. IMPA, Rio de Janeiro, 1999.

[27] G. Contreras, R. Iturriaga, G. P. Paternain, M. Paternain, Lagrangian graphs, minimizing measures and Mañé's critical values. Geom. Funct. Anal. **8** (1998), no. 5, 788–809.

[28] G. Contreras, J. Delgado, R. Iturriaga, Lagrangian flows: the dynamics of globally minimizing orbits. II. Bol. Soc. Brasil. Mat. (N.S.) **28** (1997), no. 2, 155–196.

[29] B. Dacorogna, Direct Methods in the Calculus of Variations, Applied Mathematical Sciences, **78**. Springer-Verlag, Berlin, 1989.

[30] B. Dacorogna, W. Gangbo, Extension theorems for vector valued maps, Jour. Math Pure Appliquees, to appear.

[31] C. Dellacherie, P. A. Meyer, Probabilities and Potential, North-Holland Mathematics Studies, Vol. **29**, North-Holland Publishing Co., Amsterdam-New York, 1978.

[32] L. De Pascale, On a constrained shape optimization problem, Communications in Applied Analysis, Vol. **7** no. 4 (2003)

[33] L. De Pascale, M. Gelli, L. Granieri, Minimal measures, one-dimensional currents and the Monge-Kantorovich problem, preprint. available on http://cvgmt.sns.it.

[34] L. De Pascale, A. Pratelli, Regularity properties for Monge transport density and for solutions of some shape optimization problems, Calc. Var. Partial Differential Equations **14** (2002), no. 3, 249–274.

[35] J. Dieudonné, Foundations of Modern Analysis, Academic Press, New York and London, 1969.

[36] L. C. Evans, Some new PDE methods for weak KAM theory, Calc. Var. Partial Differential Equations **17** (2003), no. 2, 159–177.

[37] L. C. Evans, Partial differential equations and Monge-Kantorovich mass transfer (survey paper), *Current Developments in Mathematics, 1997*, International Press (1999), edited by S. T. Yau.

[38] L. C. Evans, Partial Differential Equations, Graduate Studies in Mathematics **19**, AMS, Providence, RI, 1998.

[39] L. C. Evans, W. Gangbo, Differential Equation Methods for the Monge-Kantorovich Mass Transfer Problem, Memoirs AMS, Vol. **137**, no. 653, 1999.

[40] L. C. Evans, D. Gomes, Linear programming interpretation of Mather's variational principle, ESAIM Control Optim.Calc.Var. **8** (2002), 693-702.

[41] L. C. Evans, D. Effective Hamiltonians and averaging for Hamiltonian dynamics. I. Arch. Ration. Mech. Anal. **157** (2001), no. 1, 1–33.

[42] A. Fathi, Thèorème KAM faible et thèorie de Mather sur les systèmes lagrangiens, C. R. Acad. Sci. Paris Sèr. I Math. **324** (1997), no. 9, 1043–1046.

[43] A. Fathi, Weak Kam Theorem in Lagrangian Dynamics, Fifth Preliminary Version, forthcoming book, Cambridge University Press.

[44] A. Fathi, A. Siconolfi, Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation, Invent. Math. **155** (2004), no. 2, 363-388.

[45] H. Federer, Geometric Measure Theory, Springer, Berlin, 1969.

[46] M. Feldman, R. J. McCann, Monge's transport problem on a Riemannian manifold, Transaction of the AMS **354** (2002), no. 4, 1667-1697.

[47] M. Feldman, R. J. McCann, Uniqueness and transport density in Monge's mass transportation problem, Calc. Var. Partial Differential Equations **15** (2002), no. 1, 81–113.

[48] W. Gangbo, R. J. McCann, The geometry of optimal transportation, Acta Math. **177** (1996), no. 2, 113-161.

[49] M. Giaquinta, G. Modica, J. Souv̌ček, Cartesian Currents in the Calculus of Variations I, Cartesian currents, Ergebnisse der Mathematik und ihrer Grenzgebiete 3, Folge, A Series of Modern Surveys in Mathematics **37**, Springer-Verlag, Berlin, 1998.

[50] E. Giorgieri, A boundary value problem for a PDE model in mass transfer theory: representation and regularity results, Ph.D. Thesis in Mathematics, Roma Tor Vergata University, Rome, 2004.

[51] D. Gomes, Viscosity solutions methods and the discrete Aubry-Mather problem, preprint.

[52] D. Gomes, Viscosity solutions of Hamilton-Jacobi equations, and asymptotics for Hamiltonian systems. Calc. Var. Partial Differential Equations **14** (2002), no. 3, 345–357.

[53] L. Granieri, On action minimizing measures for the Monge-Kantorovich problem, preprint, available on http://cvgmt.sns.it.

[54] S. Hohloch, Optimale Massebewegung im Monge-Kantorovich Transportproblem, Degree Thesis, University of Freiburg, Germany, 2002.

[55] W. B. Johnson, J. Lindenstrauss, G. Schechtman, Extensions of Lipschitz maps into Banach spaces. Israel J. Math. **54** (1986), no. 2, 129–138.

[56] R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker-Plank equation, Siam J. Math. Anal. **29** (1998), no. 1, 1-17.

[57] J. Jost, X. Li-Jost, Calculus of Variations, Cambridge Studies in Advanced Mathematics **64**, Cambridge University Press, Cambridge, 1998.

[58] J. Jost, Riemannian Geometry and Geometric Analysis, Univesitext, Springer-Verlag, Berlin, 2002.

[59] J. Jost, Nonpositive Curvature: Geometric and Analytic Aspects, Lectures in Math. ETH Zurich, Birkhauser Verlag, Basel, 1997.

[60] U. Lang, V. Schroeder, Kirszbraun's theorem and metric spaces of bounded curvature, GAFA Geom. Funct. Anal. **7** (1997), no. 3, 535-560.

[61] F. Maddalena, J. M. Morel, S. Solimini, A variational model of irrigation patterns, Interfaces Free Bound. **5** (2003), no. 4, 391-415.

[62] R. Mañé, Generic properties and problems of minimizing measures of Lagrangian systems, Nonlinearity **9** (1996), no. 2, 273-310.

[63] R. Mañé, On the minimizing measures of Lagrangian dynamical systems. Nonlinearity **5** (1992), no. 3, 623–638.

[64] J. Mather, Action minimizing invariant measures for positive definite Lagrangian systems, Math. Z. **207** (1991), no. 2, 169–207.

[65] J. Mather, Minimal measures, Comment. Math. Helv. **64** (1989), no. 3, 375–394.

[66] E. Matoušková, Extensions of continuous and Lipschitz functions. Canad. Math. Bull. **43** (2000), no. 2, 208–217.

[67] R. J. McCann, Polar factorization of maps on Riemannian manifolds, GAFA Geom. Funct. Anal. **11** (2001), no.3, 589-608.

[68] F. Morgan, Geometric Measure Theory, A beginner's guide, Academic Press, Inc., Boston, MA, 1988.

[69] F. Otto, The geometry of dissipative evolution equations: the porus medium equation, Comm. Partial Differential Equations **26** (2001) no. 1-2, 101-174.

[70] F. Otto, C. Villani, Generalization of an inequality by Talagrand and links with the logaritmic Sobolev inequality, Journal of Functional Analysis **173** (2000), 361-400.

[71] A. Pratelli, Equivalence between some definitions for the optimal mass transport problem and for the transport density on manifolds, Ann. Mat. Pura Appl., to appear.

[72] A. Pratelli, Existence of optimal transport maps and regularity of the transport density in mass transportation problems, Ph.D. Thesis, Scuola Normale Superiore, Pisa, 2003.

[73] S. T. Rachev, L. Ruschendorf, Mass Transportation Problems Vol. I, Probab. Appl. Springer-Verlag, New York, 1998.

[74] W. Rudin, Functional Analysis, Second Edition, International Series in Pure and Applied Mathematics, McGraw-Hill, Inc., New York, 1991.

[75] L. Simon, Lectures on Geometric Measure Theory, Proceedings Centre Math. Analysis Australian National University, Vol. **3**, 1983.

[76] S. K. Smirnov, Decomposition of solenoidal vector charges into elementary solenoids and the structure of normal one-dimensional currents, St. Petersburg Math. J. **5** (1994), no. 4, 841-867.

[77] K. T. Sturm, Metric spaces of lower bounded curvature, Exposition. Math. **17** (1999), no. 1, 35–47.

[78] K. T. Sturm, Probability Measures on Metric Spaces of Nonpositive Curvature, Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces (Paris, 2002), 357–390, Contemp. Math. **338**, AMS, Providence, RI, 2003.

[79] K. T. Sturm, On the geometry of metric measure spaces, preprint.

[80] V. Sudakov, Geometric problems in the theory of infinite-dimensional probability distributions, Proc. Steklov Inst. Math. **141** (1979), no. 2, 1-178.

[81] N. S. Trudinger, X. J. Wang, On the Monge mass transfer problem, Calc. Var. Partial Differential Equations, **13** (2001), no. 1, 19-31.

[82] C. Villani, Topics in Optimal Transportation, Graduate Studies in Mathematics **58**, AMS, Providence, RI, 2003.

[83] Q. Xia, Optimal paths related to transport problems, Commun. Contemp. Math. **5** (2003), no. 2, 251-279.

[84] Q. Xia, Interior regularity of optimal transport paths, Calc. Var. Partial Differential Equations **20** (2004), no. 3, 283-299.

[85] S. Wenger, Isoperimetric inequalities of euclidean type in metric spaces, preprint.

[86] G. Wolansky, Mass Transportation, the Monge-Kantorovich problem, Hamilton-Jacobi equation and metrics on measure-valued functions, preprint.