Technische Universität München
Fakultät für Mathematik

Optimal Transport: unbalanced positive measures, dissipative evolutions and Sobolev spaces

Giacomo Enrico Sodini

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Daniel Matthes

Prüfer der Dissertation:

1. Prof. Dr. Massimo Fornasier
2. Prof. Giuseppe Savaré
3. Prof. Dr. Alexander Mielke

Die Dissertation wurde am 17.06.2022 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 14.09.2022 angenommen.

ABSTRACT

In this thesis we treat different aspects of the Optimal Transport theory.

First of all, we present a new class of Optimal Transport costs for non-negative measures with possibly different masses. These are obtained by a convex relaxation procedure of a cost for non-negative Dirac masses. As a byproduct of our analysis, we show that the classical Optimal Transport cost can be obtained by the same procedure. A primal-dual formulation of the cost, optimality conditions and metric-topological properties are also presented.

Secondly, we introduce and investigate a notion of multivalued dissipative operator (called Multivalued Probability Vector Field - MPVF) in the 2-Wasserstein space of Borel probability measures on a (possibly infinite dimensional) separable Hilbert space. Taking inspiration from the theories of dissipative operators in Hilbert spaces and of Wasserstein gradient flows, we study the well-posedness for evolutions driven by such MPVFs, and we characterize them by a suitable Evolution Variational Inequality (EVI), following the Bénilan notion of integral solutions to dissipative evolutions in Banach spaces. Our approach to prove the existence of such EVI-solutions is twofold: on one side, under an abstract stability condition, we build a measure-theoretic version of the Explicit Euler scheme showing novel convergence results with optimal error estimates; on the other hand, under a suitable discrete approximation assumption on the MPVF, we recast the EVI-solution as the evolving law of the solution trajectory of an appropriate dissipative evolution in an L^2 space of random variables.

Finally, we prove a general criterium for the density in energy of subalgebras of Lipschitz functions in the metric-Sobolev space $H^{1,p}(X, \mathsf{d}, \mathfrak{m})$ associated with a Borel positive measure $\mathfrak{m}$ in a separable and complete metric space (X, d). We then provide a relevant application to the case of the algebra of cylindrical functions in the space $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_{\mathbb{M}}}, \mathfrak{m})$ arising from a probability measure $\mathfrak{m}$ on the Kantorivich-Rubinstein-Wasserstein space $(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_{\mathbb{M}}})$ of probability measures in a complete Riemannian manifold or a separable Hilbert space $\mathbb{M}$. We will show that such a Sobolev space is always Hilbertian, independently of the choice of the reference measure $\mathfrak{m}$ so that the resulting Cheeger energy is a Dirichlet form. We will eventually provide an explicit characterization for the corresponding notion of $\mathfrak{m}$-Wasserstein gradient, showing useful calculus rules and its consistency with the tangent bundle and the Γ-calculus inherited from the Dirichlet form.

PUBLICATIONS

Some of the material presented in this thesis has appeared in the following publications:

[34] Giulia Cavagnari, Giuseppe Savaré, and Giacomo E. Sodini. "Dissipative probability vector fields and generation of evolution semigroups in Wasserstein spaces." Accepted for publication in *Probability Theory and Related Fields*, May 2022.

[109] Giuseppe Savaré and Giacomo E. Sodini. "A simple relaxation approach to duality for Optimal Transport problems in completely regular spaces." In: *Journal of Convex Analysis* 29.1 (2022), pp. 1–12.

In particular, some of the material presented in Sections 3.1, 4.3 appeared in [109], while parts of Section 6.3 and Chapters 7, 8, 9 are taken from [34].

ACKNOWLEDGEMENTS

My heartfelt thanks go without doubt to Professor Giuseppe Savaré for the support, availability and infinite generosity he has shown over the years. He was a true example and the teachings he passed on to me are certainly not limited to mathematics alone.

I also thank Professor Massimo Fornasier for his support during my stay in Munich: this PhD project existed thanks to him and to the Institute of Advanced Study of the Technical University of Munich, whose support I gratefully acknowledge.

Thanks also to Giulia for the long zoom calls that led to endless refinements of our *dissipative* works.

This doctorate took place in at least three locations and in each of these environments I got to know people who, in different ways and times, made these different periods enjoyable. Thanks to the group of PhD students from Pavia: even though we spent very little time together, you immediately made me feel at home. Thanks to Luca and Hugo for the time I spent at Bocconi and for the long lunch breaks we spent together. Finally, thanks to my colleagues in Munich who have suffered all my complaints about the German bureaucracy with stentorian patience. Among these, a special thanks goes to Cristina who most of all shared with me the time spent at TUM.

Finally, thanks to Nadia, my family and my friends for the continuous support and encouragement.

CONTENTS

LIST OF SYMBOLS

General notation

$C(X;Y)$	the set of continuous functions from X to Y;
$C_b(X)$	the set of bounded continuous real valued functions on X;
$C_c(X)$	the set of continuous real valued functions with compact support;
$\mathrm{co}\,(f)$	the convex envelope of a function f as in Section 2.2;
$\overline{\mathrm{co}}\,(f)$	the closed and convex envelope of a function f as in Section 2.2;
$\mathrm{Cyl}(\mathbb{H})$	the space of cylindrical functions on $\mathbb{H}$, see Definition 2.4.5;
$\Delta_+(X)$	the sets of non-negative Dirac masses as in Section 2.1;
$\mathrm{Discr}_+(X)$	the set of discrete non-negative measures as in Section 2.1;
$f_\sharp\mu$	the push-forward of $\nu \in \mathcal{M}(X)$ through the map $f : X \to Y$;
Γf	the lower semicontinuous relaxation of a function f as in (2.2.1);
$\Gamma(\mu,\nu)$	the set of admissible couplings between μ,ν as in (2.1.1);
$\Gamma_o(\mu,\nu)$	the set of optimal couplings between μ,ν for cost=(distance)^2,
$\mathcal{L}$	the 1-dimensional Lebesgue measure;
$\mathcal{M}(X)$	the set of Radon measures on X;
$\mathcal{M}_+(X)$	the set of non-negative Radon measures on X;
$\mathcal{P}(X)$	the set of Radon probability measures on the topological space X;
$\mathcal{P}_b(X)$	the set of Radon probability measures with bounded support;
$\mathcal{P}^r(H)$	the set of regular probabilities [5, Definitions 6.2.1, 6.2.2] on H;
$\mathcal{P}_2(X)$	the subset of measures in $\mathcal{P}(X)$ with finite quadratic moments;
$\sigma(\mathcal{E})$	the smallest sigma algebra containing $\mathcal{E}$;
$\mathrm{supp}(\mu)$	the support of $\mu \in \mathcal{M}_+(X)$;
$W_2(\mu,\nu)$	the L^2-Wasserstein distance between μ and ν, see Section 2.4;
x^t	the evaluation map defined in (2.4.6);

Part I: Unbalanced Optimal Transport

$\mathfrak{C}_R[X]$	the cone with bounded radious, as in Section 3.2;
$\mathfrak{C}_R[X_1,X_2]$	the product cone with bounded radious, as in Section 3.2;
$\mathrm{co}\,(H)\,,\mathrm{cof}\,(H)$	the convex envelopes of a function H as in Definition 3.2.7 ;
$\overline{\mathrm{co}}\,(H)$	the closed convex envelope of H as in Definition 3.2.7;
$d_{\mathfrak{C}}$	the canonical distance on the cone as in (3.2.2);
$d_1 \otimes_{\mathfrak{C}} d_2$	the product distance on the product cone as in (3.2.4);

$\mathcal{D}_{H,p}$	the distance function induced by H as in Definition 5.3.1;
Φ_H	the set of admissible potentials as in (4.2.1);
$\hat{\Gamma}$	the convex cone generated by $\Gamma \subset \mathfrak{C}[X_1, X_2]$ as in (5.2.1);
H	a function on $\mathfrak{C}[X_1, X_2]$, often requested to satisfy (4.1.2);
$\mathfrak{h}^p$	the projection map from the cone as in Section 3.2;
$\mathfrak{h}_i^p$	the projection maps from the product cone as in Section 3.2;
$\mathfrak{H}^p(\mu_1, \mu_2)$	the set of p-homogeneous plans as in (3.2.6);
$\mathcal{M}_{H,p}$	the sets of measures with finite H-moment, see Definition 5.3.1;
$\mathfrak{p}, \mathfrak{q}$	the quotient map on the cone and its inverse as in Section 3.2;
$R(\mu_1, \mu_2)$	the bound on radii given in (3.2.8);
$\mathscr{S}_H$	the singular cost as in Definition 4.1.1;

Part II: Dissipative evolutions in Kantorovich-Wasserstein spaces

$\mathbf{B}^\circ$	the minimal selection of the operator $\mathbf{B}$, see Section 6.1 ;
$\mathbf{b}_\Phi$	the barycenter of $\Phi \in \mathcal{P}(\mathsf{T}\mathbb{H})$ as in Definition 7.1.1;
$B_X(x, r)$	the open ball with radius $r > 0$ centered at $x \in X$;
$\mathrm{cl}(\mathbf{F}), \mathrm{co}(\mathbf{F})[\mu]$	the sequential closure and convexification of $\mathbf{F}$, see Section 7.7;
$\overline{\mathrm{co}}(\mathbf{F})[\mu], \hat{\mathbf{F}}$	closure of convexification and extension of $\mathbf{F}$, see Section 7.7;
$\mathrm{dir}(A)$	the set of directions induced by the set A, see (6.4.4);
$\frac{d^+}{dt}\zeta, \frac{d}{dt+}\zeta$	the right upper/lower Dini derivatives of ζ, see (8.1.3);
$D(\mathbf{F})$	the domain of a set-valued function as in Definition 7.5.1;
$D_f(\mathbf{F})$	the set of measures with finite support in $D(\mathbf{F})$, see (7.8.3);
$\mathscr{E}(\mu_0, \tau, T, L)$	the sets related to the Euler scheme (EE) defined in (9.1.5);
$\mathbf{F}_N$	the dissipative operator in $\mathcal{H}_N$ compatible with $\mathbf{F}$, see (9.4.15);
$\hat{\mathbf{F}}_N$	the maximal dissipative extension in $\mathcal{H}_N$ of $\mathbf{F}_N$;
$\mathbf{F}_\infty$	the dissipative operator on $\mathcal{H}$ defined in (9.4.33) ;
$\mathbf{F}$	the dissipative operator in $\mathcal{H}$ compatible with $\mathbf{F}$, see (9.4.35) ;
$\mathbf{F}^\circ[\mu]$	the minimal selection of the MPVF $\mathbf{F}$ at μ, see Definition 9.4.10 ;
$\lvert\mathbf{F}\rvert_2(\mu)$	the 2-nd moment of $\mathbf{F}$ at μ as in (9.2.6);
$[\mathbf{F}, \boldsymbol{\mu}]_{r,t}$, $[\mathbf{F}, \boldsymbol{\mu}]_{l,t}$	the duality pairings as in Definition 7.6.1;
$[\mathbf{F}, \boldsymbol{\mu}]_{0+}, [\mathbf{F}, \boldsymbol{\mu}]_{1-}$	the limiting duality pairings as in Definition 7.6.4;
$\lvert\Phi\rvert_2$	the 2-nd moment of $\Phi \in \mathcal{P}(\mathsf{T}\mathbb{H})$ as in (7.1.3);
$[\Phi, \vartheta]_{r,t}$, $[\Phi, \vartheta]_{l,t}$	the duality pairings as in Definition 7.4.1;
$\Gamma_o^i(\mu_0, \mu_1\lvert\mathbf{F})$, $i = 0, 1$	the set of optimal couplings conditioned to $\mathbf{F}$, see (7.6.7);
$\mathcal{H}$	the Hilbert space of r.v. parametrizing $\mathcal{P}_2(\mathbb{H})$, see (9.3.1);
$\mathcal{H}_N$	the set of r.v. constant on a partition with N elements, see (9.4.3);
$\mathcal{H}_\infty$	the set of r.v. parametrizing discrete measures, see (9.4.3);

$\mathbb{H}_N$	the set of maps $\boldsymbol{x} : I_N \to \mathbb{H}$ identified with $\mathcal{H}_N$ as in Section 9.4;
I_N	the set $\{0, 1, \ldots, N-1\}$;
$\mathcal{I}$	an interval of $\mathbb{R}$;
$\boldsymbol{i}_X(\cdot)$	the identity function on a set X;
$I(\boldsymbol{\mu}\vert\mathbf{F})$	the set of time instants t s.t. $x^t_\sharp\boldsymbol{\mu}$ belongs to $D(\mathbf{F})$, see (7.6.2);
ι	the map sending a r.v. X to its law, see Section 9.4;
ι^2	the map sending a pair of r.v. (X, Y) to their law, see Section 9.4;
J_τ	the resolvent of a maximal dissipative operator, see Section 6.1;
λ_+	the positive part of $\lambda \in \mathbb{R}$, given by $\lambda_+ = \max\{\lambda, 0\}$;
Λ, Λ_o	the sets of couplings as in Definition 7.1.7 and Theorem 7.1.8;
$m_2(\nu)$	the 2-nd moment of $\nu \in \mathcal{P}(X)$ as in (7.3.1);
$\vert\dot{\mu}_t\vert$	the metric derivative at t of a curve μ;
$\mathfrak{N}$	unbounded directed subset of $\mathbb{N}$ as in Section 6.2;
$\mathcal{O}_N$	the subset of injective maps in $\mathbb{H}_N$ as in Section 9.4;
O_N	the set of injective maps from $\mathscr{I}_N$ to $\mathbb{H}$ as in Section 9.4;
$\mathcal{P}_2^{sw}(X \times Y)$	the space in Definition 6.3.2;
$\mathcal{P}_2^{sw}(\mathbb{TH})$	the space $\mathcal{P}_2^{sw}(\mathbb{H} \times \mathbb{H})$;
$\mathcal{P}_2(\mathbb{TH}\vert\mu)$	the subset of $\mathcal{P}_2(\mathbb{TH})$ with fixed first marginal μ as in (7.1.9);
$\mathcal{P}_f$	the set of probabilities on $\mathbb{H}$ with finite support as in (7.8.1);
$\mathcal{P}_c$	the set of probabilities on $\mathbb{H}$ with compact support as in (7.8.1):
$\mathcal{P}_N$	the set of discrete probabilities as in (7.8.2);
$\mathcal{P}_{\mathfrak{N}}(\mathbb{H})$	the set of discrete measures associated to $\mathfrak{N}$ as in (9.4.2);
Π_N	the projection from $\mathcal{H}$ to $\mathcal{H}_N$ as in Section 9.4;
$S(\Omega)$	the set of measure preserving maps on Ω, see Section 6.2;
S_t	the semigroup generated by an operator as in Section 6.1;
$\mathcal{S}_{\mathbf{F}}(\mu)$	the set of measures in $D_f(\mathbf{F})$ that see μ as in Definition 9.4.15 ;
$\mathrm{Tan}_\mu \mathcal{P}_2(X)$	the tangent space defined in Theorem 2.4.6;
$\mathbb{TH}$	the tangent bundle to $\mathbb{H}$;
$x^{t.\theta}$	the map defined in (6.4.1);
$[\cdot,\cdot]_r, [\cdot,\cdot]_l$	the pseudo scalar products as in Definition 7.1.4;
$\lfloor\cdot\rfloor, \lceil\cdot\rceil$	the floor and ceiling functions, see (9.1.1).
$\prec$	order relation on $\mathbb{N}$ defined in (6.2.4);
$\precnsim$	the strict order relation induced by $\prec$;

Part III: Kantorovich-Wasserstein-Sobolev spaces

$\mathscr{A}$	a unital subalgebra of $\mathrm{Lip}_b(X)$, see (10.1.3);
CE_p	the Cheeger energy, see Definition 10.1.3;

$\mathrm{Cyl}(\mathcal{P}_2(\mathbb{R}^d))$	the subalgebra of cylindrical function as in Definition 11.1.2;
$\lvert Df\rvert_\star$	the relaxed gradient of f, see Definition 10.1.1;
$\lvert Df\rvert_{\star,\mathscr{A}}$	the relaxed gradient of f w.r.t. $\mathscr{A}$, see Definition 10.1.1;
d_y	the function distance from $y \in X$, see (10.1.23);
DF	the Wasserstein gradient of F as in Definition 11.1.5;
$\lVert DF[\mu]\rVert_\mu$	the norm of DF at μ as in Definition 11.1.5 ;
$D_{\mathfrak{m}}F$	the Wasserstein differential of $F \in D^{1,2}(\mathbb{W}_2)$ as in Definition 11.2.1;
$D^{1,p}(X, d, \mathfrak{m})$	the space of functions with a relaxed gradient, see Definition 10.1.3;
$\mathbf{D}_{\mathfrak{m}}$	the multivalued gradient operator as in Definition 11.2.5;
G_0	the set of residual gradients as in Definition 11.2.6;
$G_0[\mu]$	the section of G_0 at μ defined in (11.2.40);
$\mathbf{G}$	the graph of $\mathbf{D}_{\mathfrak{m}}$ as in Definition 11.2.5;
$H^{1,p}(X, d, \mathfrak{m}; \mathscr{A})$	the metric Sobolev space with respect to $\mathscr{A}$ as in Definition 10.1.3 ;
$\mathrm{Lip}(f, A, d)$	the Lipschitz constant of f on A w.r.t. d, see (10.1.2) ;
$\mathrm{lip}_d f$	the local Lipschitz constant of f w.r.t. d as in (10.1.1);
$\ell_d(\gamma; [\alpha, \beta])$	d-length of the restriction of $\gamma : [a, b] \to X$ to $[\alpha, \beta]$ as in (10.3.2) ;
$\ell_d(\gamma)$	the d-length of $\gamma : [a, b] \to X$ equal to $\ell_d(\gamma; [a, b]$;
L_ϕ	the cylindrical function induced by $\phi \in C^1_b(\mathbb{R}^d)$ as in (11.1.3);
$L_{\boldsymbol{\phi}}$	the cylindrical function induced by $\boldsymbol{\phi} \in (C_b(\mathbb{R}^d))^N$ as in (11.1.4);
$\boldsymbol{\mathfrak{m}}$	the measure on $\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ as in (11.1.10);
pCE_p	the pre-Cheeger energy, as in Remark 10.1.4;
R_γ	canonical reparametrization of a curve γ, see (10.3.7);
$\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}})$	the tangent space to $\mathcal{P}_2(\mathbb{R}^d)$ w.r.t. $\boldsymbol{\mathfrak{m}}$ as in Definition 11.2.7;
T	the set of tangent vectors defined in (11.2.34);
$T[\mu]$	the tangent space at μ as defined in (11.2.41);
$\mathbb{W}_2$	the metric-measure space $(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ as in Section 11.1;

1 INTRODUCTION

The Optimal Transport problem was proposed for the first time by Monge [83] in 1781 and its current mathematical formulation is mainly due to Kantorovich [66]. Roughly speaking, it asks to find the best way to move a certain amount of material from one starting place to a given new configuration. Depending on the meaning assigned to the terms “best", “move" and “place", one can get many different situations: for instance allocation of resources, the physical movement of masses, evolution of particle systems, etc...

With the modern language of measure theory, we can formulate the problem as follows: suppose we are given two complete and separable metric spaces X and Y and two Borel probability measures μ on X and ν on Y; let us assume that $c : X \times Y \to [0, +\infty]$ is a cost function, meaning that the value of $c(x, y)$ tells how much we pay to move a unit mass in x to the target location y. What we want to find is a map $T : X \to Y$ “moving μ to ν" and minimizing the cost

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

The interpretation of the above quantity is straightforward: given T, we move every $x \in X$ to the assigned location $T(x)$ and we evaluate the cost of this operation as $c(x, T(x))$; integrating w.r.t. μ gives the total cost associated to T.

With the expression “T moves μ to ν", we mean that for every subset B of Y the total mass sent to B (corresponding to $\mu(T^{-1}(B))$ must coincide with the mass assigned to B (which is given by $\nu(B)$). In other words, T must satisfy

$$\nu(B) = \mu(T^{-1}(B)) \quad \text{for every Borel set } B \subset Y.$$

This is expressed in mathematical terms saying that ν is the *push forward* of μ through T, denoted by $T_\sharp\mu$. The precise formulation of the so called Monge Optimal Transport problem is then

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

Monge did a fine analysis of many properties of minimizers in case $X = Y = \mathbb{R}^d$ and the cost function is given by the Euclidean distance, in particular carrying out a deep study of the geometric properties of transport rays; existence of minimizers was only addressed later: the main issue is that in some cases the set of admissible transport maps may be empty (e.g. in case $X = Y = [0, 1]$, $\mu = \delta_0$ and ν is the Lebesgue measure on Y) and, in general, it doesn’t enjoy good compactness or closure properties.

The question of existence of minimizers was solved only many years later by Kantorovich; in his formulation, mass is allowed to split: this corresponds to the introduction of transport plans, instead of maps; a transport plan $\boldsymbol{\gamma}$ is a probability on the product space $X \times Y$ having as marginals μ and ν respectively. This means that, to each pair of Borel subsets $A \subset X$ and $B \subset Y$, the measure $\boldsymbol{\gamma}$ assigns the fraction $\gamma(A \times B)$ of mass in A that has to be moved to B. The Kantorovich Optimal Transport problem is thus to find

$$\mathsf{OT}_c(\mu, \nu) := \inf \left\{ \int_{X \times Y} c(x,y) \, d\boldsymbol{\gamma}(x,y) \mid \boldsymbol{\gamma} \in \Gamma(\mu,\nu) \right\}, \tag{1.0.1}$$

where $\Gamma(\mu, \nu)$ is the set of admissible transport plans. It is readily seen that $\Gamma(\mu, \nu)$ is always non-empty and it can be proven to enjoy very good properties, in particular it is compact in the narrow topology. Under a few regularity assumptions on the objects involved in the problem, the existence of minimizers can be easily proven (e.g. if the problem is feasible and the cost function is proper and lower semicontinuous). Moreover the nice linear structure of the problem allows to use techniques from convex optimization, and thus makes it possible to provide duality formulas, in order to characterize the optimal solutions of the problem.

After the seminal works [66, 67] of Kantorovich, the Optimal Transport problem received a great boost at the end of the XX century starting with the fundamental work of Brenier [23] (see also [52]) where he proved that, under suitable hypotheses on the measures and on the cost function, there is a unique Optimal Transport plan which moreover is concentrated on a map. Such result also provides connections with the Monge-Ampère equation (see also [28]) and was used by Brenier to prove a polar factorization theorem.

Since then, the theory of Optimal Transport has enormously grown in a number of directions, see for example the works [3, 9, 29, 47, 117] related to the existence of optimal maps, the fundamental papers [64, 90] where the connection between evolution PDEs and optimal transport problems was first noted and the works [7, 56, 79, 113, 114] linked to analysis in metric spaces.

Moreover, in more recent years, Optimal Transport has been a widely used tool in image processing or data analysis, so that it has become more and more important to come up with new ways to compute efficiently the Optimal transport cost [16–18, 45]: this is linked for example to the Entropic regularization of Optimal Transport [32] which is in turn connected with the Shrödinger problem [73, 74] and it is a fruitful ground for research. For a comprehensive introduction to the theory of Optimal Transport and for more exhaustive lists of references we refer to the classic monographs [5, 99, 100, 107, 118, 119], the more recent [4, 49] and the application-oriented [92].

In the thesis we will deal with three specific research topics related to Optimal Transport: in particular we will address the generalization of Optimal Transport to pairs of measures with different masses, the description of evolutions of probability measures under a suitable notion of velocity field and the definition of metric Sobolev spaces on the space of probability measures endowed with the Wasserstein distance. Before entering in the description of the content of this thesis, let us give a few references for these topics.

Unbalaced Optimal Transport. As it can be easily checked, the constraint $\gamma \in \Gamma(\mu, \nu)$ in (1.0.1) forces the measures μ and ν to have the same total mass. It is thus interesting to consider the situation when the two measures have different non-negative masses and define the analogue of an Optimal Transport cost in this case.

The problem of extending Optimal Transport methods to pairs of unbalanced positive measures has been considered in a large number of works with different techniques and different aims.

For what concerns dynamical formulations, many models inspired by the fluid dynamic formulation of [15] for the classical Optimal Transport problem have been proposed, see for example [69, 77, 80, 95, 96]. In such works, the authors consider source terms in the continuity equation, thus leading to gain/loss of mass during the evolution. The models proposed differ in the kind of source chosen or in the penalization of it. We refer also to [36] where a more detailed description of these models is given.

Static formulations of the unbalanced Optimal Transport problem were proposed already by Kantorovich and Rubinstein [65] and subsequently extended by Hanin [60] (see also the dual norm in [59]). These approaches can be thought as a classical Optimal Transport problem where a fraction of the mass is allowed to go (or come from) a point at infinity (see also [58]). More recent approaches are given by the so called optimal partial transport [30, 48], which was previously related to image retrieval [91, 105].

Optimal partial transport (see [36]) is in turn also related to [95, 96], since this latter works also provide a dynamic formulation of optimal partial transport. We also mention that [95, 96] are also connected to [14] where it was proposed to change the marginal constraints and to add a penalization term.

Let us finally mention the Entropic Transport approach that has been proposed independently in [37, 76]. The underlying idea is to interpolate the Wasserstein and the Hellinger metrics in order to produce a new transport cost between measures with possibly different masses. In this way, many of the above approaches can be seen as particular instances of this class of Entropy-Transport problems.

In the first part of the thesis, extending some idea already contained in [76], we define a cost between measures as the convex and lower semicontinuous relaxation of a cost defined between weighted Dirac masses. Duality formulas, optimality conditions and metric-topologial properties of this new class of costs are also presented.

Evolutions in the space of probability measures. Many relevant examples of evolutionary PDEs (describing transport and diffusion phenomena) and models for describing the interaction of agents/particles show the great importance to study the evolution in time of probability measures.

A very important class of such evolutions is provided by (a suitable adapted notion of) gradient flow in the space of probability measures. The starting point of this theory comes from the works of Jordan, Kinderlerer and Otto [64, 90] where they noted for the first time a gradient flow structure in some PDEs w.r.t. the Kantorovich-Rubinstein Wasserstein (in brief, Wasserstein) distance on probability measures (see also [1, 27, 90].

We refer to [5], where the authors develop a whole theory for the notion of gradient flow in metric spaces: in this context the Euclidean definition must be carefully adapted since neither the notion of velocity of a curve or the one of subdifferential of a functional are immediately clear. A crucial tool they employ for existence and uniqueness results is the one of geodesic convexity of a functional on a metric space which is is strictly connected to the displacement convexity introduced and studied by McCann [82].

Besides gradient flows, other kinds of evolutions describing large numbers of agents have been considered; we mention here two recent works related to evolutions of probability measures under the action of notions of vector fields. In [21] the authors aim at developing a Cauchy-Lipschitz theory for non-local continuity equations where the velocity field can be thought as a map from the space of probabilities to the space of Lipschitz vector fields on the base space. In the works of Piccoli [93, 94], which have been in part an inspiration for the second part of the thesis, it is introduced the notion of Measure Probability Vector Field: heuristically, this is a vector field on probability measures such that each point in the support of a probability measure is moved according to a probability distribution on the space of admissible velocities.

In the second part of the thesis we further generalize the theory proposed by Piccoli and we connect it to the general theory of Wasserstein gradient flows and of dissipative evolutions in Hilbert spaces. We provide a notion of evolution based on a suitable Evolution Variational Inequality and we prove existence of curves satisfying it employing measure theoretical versions of both implicit and explicit Euler schemes.

Metric Sobolev spaces and Optimal Transport. Starting with the fundamental work of Otto [90], the geometry of the Wasserstein space $(\mathcal{P}_2(X), W_2)$, where X is a Riemannain Manifold or an Hilbert space, has been deeply investigated. Besides the study of gradients for smooths functions, the definition of tangent spaces [5, 53] and in general the properties of a weak Riemmanian structure [78], in the last years there have been also proposals to define a canonical Riemannain measure on $(\mathcal{P}_2(X), W_2)$: this was done first by Sturm and Von Renesse [101] for the particular case of $X = \mathbb{S}^1$ and then generalized by Sturm [115] to the case of a closed smooth Riemannian manifold.

It is natural to study the Dirichlet energy associated to such measure: for example, Dello Schiavo [44] considered measures $\mathfrak{m}$ that satisfy an integration-by-parts formula; starting from the regular class of cylindrical functions on $(\mathcal{P}_2(X), W_2)$, in [44] the author proves a Rademacher-type result for Lipschitz functions.

Since $(\mathcal{P}_2(X), W_2, \mathfrak{m})$ is a particular example of complete metric measure space, it is natural to compare the approach of Sturm and Dello Schiavo with the one coming from the general metric theory [19, 57, 62, 108]. This aspect is even more interesting because of the crucial role that Optimal Transport played in the development of the theory of CD and RCD spaces [7, 8, 54, 79, 113, 114]. In particular this studies revealed the importance of the notion of infinitesimal Hilbertianity, meaning that the Cheeger energy associated to the metric-Sobolev space is a quadratic form. In the particular case of $(\mathcal{P}_2(X), W_2, \mathfrak{m})$ the point is thus to understand if the class of cylindrical functions is dense in energy in the metric-Sobolev space.

The third part of the thesis is devoted to these kinds of problems; first we provide a general criterium for sub-algebras of Lipschitz and bounded functions to be dense in Sobolev metric measure spaces. Then we apply the general result to the Sobolev space $H^{1,2}(\mathcal{P}_2(X), W_2, \mathfrak{m})$, for a general Borel positive measure $\mathfrak{m}$ on $\mathcal{P}_2(X)$, showing that, in case X is a (possibly infinite dimensional) Hilbert space or a complete Riemannian manifold, the resulting metric Sobolev space is indeed Hilbertian. We will eventually provide an explicit characterization for the corresponding notion of $\mathfrak{m}$-Wasserstein gradient, showing useful calculus rules and its consistency with the tangent bundle and the Γ-calculus inherited from the Dirichlet form.

The remaining part of this introduction is devoted to the detailed discussion of the three parts of the thesis.

1.1 CONTENT OF THE THESIS AND MAIN RESULTS

1.1.1 *Part I: Unbalanced Optimal Transport*

In order to highlight the analogies between the classical Optimal Transport setting and the results we have obtained in the unbalanced case, we first give a brief account of some of the fundamental results of the classical theory.

The classical Optimal Transport case. We fix two complete and separable metric spaces X_1 and X_2, two Borel probability measures $\mu_i \in \mathcal{P}(X_i)$, $i = 1, 2$, and a proper (i.e. not identically $+\infty$) and lower semicontinuous cost function $c : X_1 \times X_2 \to [0, +\infty]$. As outlined in the first part of the introduction, the primal formulation of the optimal transport problem is given by

$$\mathsf{OT}_c(\mu_1, \mu_2) := \inf \left\{ \int_{X_1 \times X_2} c(x_1, x_2) \, d\boldsymbol{\gamma}(x_1, x_2) \mid \boldsymbol{\gamma} \in \Gamma(\mu_1, \mu_2) \right\} \tag{1.1.1}$$

where $\Gamma(\mu_1, \mu_2)$ can be defined as

$$\Gamma(\mu_1, \mu_2) := \left\{ \boldsymbol{\gamma} \in \mathcal{P}(X_1 \times X_2) \mid x^i_\sharp \boldsymbol{\gamma} = \mu_i, \ i = 1, 2 \right\},$$

with $x^i : X_1 \times X_2 \to X_i$ given by $x^i(x_1, x_2) = x_i$, for $i = 1, 2$. This is of course just an alternative way to say that the marginals of $\boldsymbol{\gamma}$ are μ_1 and μ_2. If we consider on $\mathcal{P}(X_1 \times X_2)$ the narrow topology (the topology induced by the duality with bounded and continuous functions in $X_1 \times X_2$), the existence of minimizers in (1.1.1) is a direct consequence of the narrow compactness of $\Gamma(\mu_1, \mu_2)$ and of the narrow lower semicontinuity of the map $\boldsymbol{\gamma} \mapsto \int c \, d\boldsymbol{\gamma}$.

The celebrated Kantorovich duality theorem states that

$$\mathsf{OT}_c(\mu_1,\mu_2) = \sup\{\mathcal{D}(\varphi_1,\varphi_2;\mu_1,\mu_2) \mid (\varphi_1,\varphi_2) \in \Psi_c\} \tag{1.1.2}$$

where $\mathcal{D}$ is simply the duality pairing

$$\mathcal{D}(\varphi_1,\varphi_2;\mu_1,\mu_2) := \int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 \tag{1.1.3}$$

and the set of admissible pairs Ψ_c is defined as

$$\Psi_c := \left\{ (\varphi_1,\varphi_2) \in C_b(X_1) \times C_b(X_2) \;\middle|\; \begin{array}{l} \varphi_1(x_1) + \varphi_2(x_2) \leqslant c(x_1,x_2) \\ \text{for every } (x_1,x_2) \in X_1 \times X_2 \end{array} \right\}.$$

While the proof of the $\geqslant$ inequality in (1.1.2) is immediate, the proof of the converse inequality is generally more involved and, usually, it is first carried out in a simplified setting (discrete measures, compact spaces, etc...) using some convex analysis tool ultimately relying on Hahn-Banach theorem, and then extended to the general setting by exploiting the structure of the problem (see e.g. [68, 72, 118]).

The existence of maximizers for the dual problem (1.1.2) (also called optimal potentials) in the class Ψ_c is not guaranteed in general, unless stronger hypotheses are assumed on the spaces and on the cost. For example if X_1 and X_2 are compact and c is continuous, one can get the existence of a pair $(\varphi_1,\varphi_2) \in \Psi_c$ realizing the equality

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 = \mathsf{OT}_c(\mu_1,\mu_2).$$

The proof (see e.g. [107, Proposition 1.11]) of this result is usually based on the c-transform technique: given an admissible pair (φ_1,φ_2), the pair $(\varphi_1^{cc},\varphi_1^{c})$ defined as

$$\begin{aligned} \varphi_1^c(x_2) &:= \inf_{x_1 \in X_1} \{c(x_1,x_2) - \varphi_1(x_1)\}, \quad x_2 \in X_2, \\ \varphi_1^{cc}(x_1) &:= \inf_{x_2 \in X_2} \{c(x_1,x_2) - \varphi_1^c(x_2)\}, \quad x_1 \in X_1, \end{aligned} \tag{1.1.4}$$

is still admissible, both functions have the same (uniform) modulus of continuity of c and the new pair $(\varphi_1^{cc},\varphi_1^c)$ does better than the previous one, meaning that

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 \leqslant \int_{X_1} \varphi_1^{cc} \, d\mu_1 + \int_{X_2} \varphi_1^c \, d\mu_2.$$

Finally, observing that the "shifted" pair $(\varphi_1 - k, \varphi_2 + k)$, $k \in \mathbb{R}$, is again admissible and realizes the same value in the dual formulation, it is possible to impose that $\min \varphi_1 = 0$, so that one gets also uniform boundedness of the potentials and an application of Ascoli-Arzelà theorem leads to the existence of a maximizing pair in Ψ_c.

The existence of a sufficiently regular maximazing pair (φ_1, φ_2) can be also employed to prove the existence of an Optimal Transport map, if the cost c and the probabilities μ_1, μ_2 satisfy some additional hypotheses. Let us briefly sketch the idea in a simple setting: suppose that $X_1 = X_2 = \Omega$ is a smooth bounded domain in $\mathbb{R}^d$, $\mu_1 \ll \mathcal{L}^d|_{\Omega}$, c and φ_1 are Lipschitz continuous and $\boldsymbol{\gamma}$ is an Optimal Transport plan in $\Gamma(\mu_1, \mu_2)$; we know that $\varphi_1(x_1) + \varphi_2(x_2) \leqslant c(x_1, x_2)$ for every $x_1, x_2 \in \Omega$ and that on the support of $\boldsymbol{\gamma}$ this must be an equality. Under the present assumptions, it is possible to see that the interior of the set

$$\Sigma := \{(x_1, x_2) \in \mathrm{supp}(\boldsymbol{\gamma}) \mid y \mapsto c(y, x_2) - \varphi_1(y) \text{ is differentiable at } y = x_1\}$$

has full $\boldsymbol{\gamma}$-measure. Thus on $\mathrm{int}(\Sigma)$ we have that $\nabla\varphi_1(x_1) = \partial_1 c(x_1, x_2)$; if we ask to the cost c to satisfy the so called “twist condition", i.e. that $z \mapsto \partial_1 c(x_1, z)$ is invertible for every $x_1 \in \Omega$, we have that there exists a unique $x_2 \in \Omega$ such that $\nabla\varphi_1(x_1) = \partial_1 c(x_1, x_2)$ and thus $\boldsymbol{\gamma}$ is concentrated on the graph of a Borel function T i.e. T is an Optimal Transport map.

In general one cannot hope to get the existence of a maximizing pair in the class Ψ_c, but optimality conditions are (almost) always available in the form of the following result, which relies on the crucial notion of c-cyclical monotonicity: a set $\Lambda \subset X_1 \times X_2$ is said to be c-cyclically monotone if for every $n \in \mathbb{N}$, every family of points $\{(x_1^i, x_2^i)\}_{i=1}^n \subset \Lambda$ and every permutation σ of $\{1, \dots, n\}$ it holds

$$\sum_{i=1}^n c(x_1^i, x_2^i) \leqslant \sum_{i=1}^n c(x_1^i, x_2^{\sigma(i)}).$$

It is clear that the c-cyclical monotonicity of the support of a plan $\boldsymbol{\gamma} \in \Gamma(\mu_1, \mu_2)$ is equivalent to optimality for $\boldsymbol{\gamma}$ in the discrete setting (i.e. if the measures μ_i have a finite support), but it is a remarkable result (see e.g. [5, Theorem 6.1.4]) that this is also the case in general, under a few integrability assumptions on the cost function.
The last fundamental fact about Optimal Transport we want to recall is related to metric and topological properties (see e.g. [5, Proposition 7.1.5]); indeed, if we take as cost $c = d^p$ where d is a distance metrizing X and $p \in [1, +\infty)$, then the resulting optimal transport cost OT_{d^p} is the p-th power of a distance (the p-Wasserstein distance W_p) on

$$\mathcal{P}_p(X) := \left\{\mu \in \mathcal{P}(X) \mid \int_X d^p(x, x_0)\, d\mu(x) < +\infty \text{ for some, hence for all, } x_0 \in X\right\}.$$

The topology induced by W_p is compatible with the narrow topology, in the sense that given a sequence $(\mu_n)_n \subset \mathcal{P}_p(X)$ and $\mu \in \mathcal{P}_p(x)$, we have that

$$W_p(\mu_n, \mu) \to 0 \text{ if and only if } \begin{cases} \int_X d^p(x, x_0) d\mu_n(x) \to \int_X d^p(x, x_0) d\mu(x) \\ \quad \text{for some, hence for all, } x_0 \in X, \\ \mu_n \to \mu \text{ in the narrow topology of } \mathcal{P}(X). \end{cases} \tag{1.1.5}$$

Contributions in Part I. In many of the approaches to generalize the classical Optimal Transport problem to unbalanced measures [37, 76, 95, 96] the general idea is to find an equivalent formulation of the standard Optimal Transport problem which is suitable to be extended to general measures. Let us briefly describe the idea of [76] which was in part an inspiration for our work: in [76] the authors consider two Polish spaces X_1 and X_2, two nonnegative Radon measures $\mu_1 \in \mathcal{M}_+(X_1)$ and $\mu_2 \in \mathcal{M}_+(X_2)$, two Borel functions $F_1, F_2 : [0,+\infty) \to [0,+\infty]$, and a Borel cost function $c : X_1 \times X_2 \to [0,+\infty]$. Given a nonnegative Radon measure $\boldsymbol{\gamma} \in \mathcal{M}_+(X_1 \times X_2)$, they define the functionals

$$\mathcal{F}_i(\boldsymbol{\gamma}; F_i) := \int_{X_i} F_i(\sigma_i)\, d\mu_i + (F_i)'_\infty \gamma_i^\perp(X_i), \quad x^i_\sharp \boldsymbol{\gamma} = \sigma_i \mu_i + \gamma_i^\perp, \quad i = 1,2,$$

where $(F_i)'_\infty := \lim_{s \to +\infty} \frac{F_i(s)}{s}$ and $(\sigma_i, \gamma_i^\perp)$ is the Lebesgue decomposition of $x^i_\sharp \boldsymbol{\gamma}$ w.r.t. μ_i. They then define their entropy-transport cost as

$$\mathsf{ET}_{F_1,F_2,c}(\mu_1,\mu_2) := \inf_{\boldsymbol{\gamma} \in \mathcal{P}(X_1 \times X_2)} \sum_i \mathcal{F}_i(\boldsymbol{\gamma}; F_i) + \int_{X_1 \times X_2} c\, d\boldsymbol{\gamma}. \tag{1.1.6}$$

The idea is that $\mathcal{F}_i$ measures the discrepancy between the i-th marginal of $\boldsymbol{\gamma}$ and the measure μ_i, and then one adds the standard Optimal Transport cost induced by c. If $F_i = I_1$ for $i = 1,2$, one gets the classical Optimal Transport cost induced by c, where I_1 is the function equal to 0 at 1 and $+\infty$ elsewhere.

The strategy we propose here is in the same spirit: we first notice that the classical Optimal Transport cost can be expressed as the convex relaxation of a suitable functional on measures and we use this point of view to define a notion of cost for non-negative measures with possibly different masses.
The first key observation is the following: the Optimal Transport cost as in (1.1.1) satisfies

$$\mathsf{OT}_c(\delta_{x_1}, \delta_{x_2}) = c(x_1, x_2) \quad \text{for every } (x_1,x_2) \in X_1 \times X_2 \tag{1.1.7}$$

so that it is natural to define the singular cost $\mathcal{F}_c : \mathcal{M}(X_1) \times \mathcal{M}(X_2) \to [0,+\infty]$ given by

$$\mathcal{F}_c(\mu_1,\mu_2) := \begin{cases} r_1 c(x_1,x_2) & \text{if} \quad \begin{array}{l} \mu_1 = r_1\delta_{x_1},\, \mu_2 = r_1 \delta_{x_2}, \\ x_1 \in X_1, x_2 \in X_2,\, r_1 \geqslant 0, \end{array} \\ +\infty & \text{elsewhere}, \end{cases} \tag{1.1.8}$$

where $\mathcal{M}(X_i)$ denotes the vector space of signed and finite Radon measures on X_i, for $i = 1,2$.
The second remark comes from (1.1.2) which shows that the Optimal Transport cost OT_c, being the supremum of a family of linear and narrowly continuous functionals, is a convex and narrowly lower semicontinuous functional in $\mathcal{P}(X_1) \times \mathcal{P}(X_2)$.
It is therefore natural to consider the extension EOT_c of OT_c to $\mathcal{M}(X_1) \times \mathcal{M}(X_2)$,

obtained by homogeneity if $\mu_1(X_1) = \mu_2(X_2) \geqslant 0$ or set equal to $+\infty$ if one of the measures is negative; in other words, we define $\mathsf{EOT}_c : \mathcal{M}(X_1) \times \mathcal{M}(X_2) \to [0, +\infty]$ as

$$\mathsf{EOT}_c(\mu_1, \mu_2) := \begin{cases} a\mathsf{OT}_c(\mu_1/a, \mu_2/a) & \text{if } \mu_1(X_1) = \mu_2(X_2) = a \geqslant 0, \\ +\infty & \text{elsewhere.} \end{cases} \tag{1.1.9}$$

The above observations strongly suggest that EOT_c can be characterized as the largest narrowly lower semicontinuous and convex functional below the cost $\mathcal{F}_c$, which is precisely the content of [109] where we thus obtained that

$$\mathsf{EOT}_c = \overline{\mathrm{co}}\,(\mathcal{F}_c)\,. \tag{1.1.10}$$

The nice consequence of the proof of this result is twofold: on one hand it provides a natural and simple proof of the Kantorovich duality. On the other hand, it presents a characterization of the Optimal Transport functional that can be generalized to the unbalanced context. To do that we need to interpret further the equality (1.1.10); what it seems to suggest is that, in a way, the space of non-negative measures $\mathcal{M}_+(X)$ is the closed and convex envelope of "weighted points" in X and that cost functions on X lift to nice cost functions on $\mathcal{M}_+(X)$ through a convexification procedure, meaning that the resulting cost is the corresponding (extended) Optimal Transport one. The rigorous counterpart is the fact that actually the convex envelope of the set of weighted Dirac masses is dense in $\mathcal{M}(X)$ (see Proposition 3.1.1). However, there is a representation issue due to the fact that the null measure can be represented in many (actually infinite) ways as a weighted Dirac mass i.e.

$$0_X = 0 \cdot \delta_x \quad \text{for every } x \in X.$$

This suggests that the correct space to represent weighted Dirac masses is not exactly $X \times \mathbb{R}_+$ but rather the quotient of this space w.r.t. the equivalence relation that sends all the points $(x, 0)$ to the same equivalence class. More rigorously, we define on $X \times \mathbb{R}_+$ the equivalence relation

$$(x, r) \sim (y, s) \quad \overset{\text{def}}{\Leftrightarrow} \quad [x = y, r = s \neq 0 \quad \vee \quad r = s = 0]$$

and the corresponding geometric cone $\mathfrak{C}[X] := (X \times \mathbb{R}_+)/\sim$. Points in $\mathfrak{C}[X]$ are denoted by equivalence classes $[x, r]$; in $\mathfrak{C}[X]$ one can consider a suitable topology (weaker than the quotient one) that makes it isomorphic to the set of weighted Dirac masses endowed with the (restriction of the) narrow topology (see Lemma 3.2.1).

Taken into account the cone construction, the correct way to define, this time, a cost on weighted Dirac masses is thus to consider a Borel function

$$\mathsf{H} : \mathfrak{C}(X_1) \times \mathfrak{C}(X_2) \to [0, +\infty],$$

which we will assume to be proper, lower semicontinuous and 1-homogeneous, in the sense that the map

$$(r_1, r_2) \in \mathbb{R}^2_+ \mapsto \mathsf{H}([x_1, r_1], [x_2, r_2])$$

is 1-homogeneous for every fixed $(x_1, x_2) \in X_1 \times X_2$. While the properness and the lower semicontinuity assumptions are natural, the 1-homogeneity assumption deserves a comment: from a modeling point of view we are saying that moving $mr_1\delta_{x_1}$ to $mr_2\delta_{x_2}$ costs exactly m times moving $r_1\delta_{x_1}$ to $r_2\delta_{x_2}$. In analogy with (1.1.8) we can define the unbalanced singular cost $\mathscr{S}_H : \mathcal{M}(X_1) \times \mathcal{M}(X_2) \to [0, +\infty]$ as

$$\mathscr{S}_H(\mu_1, \mu_2) := \begin{cases} H([x_1, r_1]; [x_2, r_2]) & \text{if} \quad \begin{array}{l} \mu_1 = r_1\delta_{x_1},\ \mu_2 = r_2\delta_{x_2}, \\ x_1 \in X_1, x_2 \in X_2,\ r_1, r_2 \geqslant 0, \end{array} \\ +\infty & \text{elsewhere.} \end{cases}$$

As a primal formulation for the unbalanced problem, given $H : \mathfrak{C}[X_1] \times \mathfrak{C}[X_2] \to [0, +\infty]$ as above, let us consider the functional

$$\mathscr{U}_H(\mu_1, \mu_2) := \inf \left\{ \int H \, d\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2) \right\}, \tag{1.1.11}$$

with

$$\mathfrak{H}^1(\mu_1, \mu_2) := \left\{ \boldsymbol{\alpha} \in \mathfrak{M}^1_+(\mathfrak{C}[X_1] \times \mathfrak{C}[X_2]) : \mathfrak{h}^1_i(\boldsymbol{\alpha}) = \mu_i \right\},$$

where $\mathfrak{M}^1_+(\mathfrak{C}[X_1] \times \mathfrak{C}[X_2])$ is the subset of measures in $\mathcal{M}_+(\mathfrak{C}[X_1] \times \mathfrak{C}[X_2])$ such that $\int (r_1 + r_2) \, d\boldsymbol{\alpha}$ is finite and $\mathfrak{h}^1_i$ is the map sending $\boldsymbol{\alpha} \in \mathfrak{M}^1_+(\mathfrak{C}[X_1] \times \mathfrak{C}[X_2])$ to $x^i_\sharp(r_i\boldsymbol{\alpha}) \in \mathcal{M}_+(X_i)$. Notice that this coincides precisely with the "homogeneous perspective marginal costs" considered in [76, Definition 5.1]) (this was indeed the part of [76] that inspired our work): the authors prove that suitably combining F_1, F_2 and c (see in particular [76, Definition 5.1]) one can obtain a function H (thus depending on F_1, F_2 and c) such that the entropy-transport cost as in (1.1.6) coincides with $\mathscr{U}_H$. Notice that, if the cost function H is given by

$$H([x_1, r_1], [x_2, r_2]) := \begin{cases} r_1 c(x_1, x_2) & \text{if } r_1 = r_2 \geqslant 0, \\ +\infty & \text{elsewhere,} \end{cases} \tag{1.1.12}$$

for some proper and lower semicontinuous function $c : X_1 \times X_2 \to [0, +\infty]$, then $\mathscr{U}_H = \mathsf{EOT}_c$ and $\mathscr{S}_H = \mathcal{F}_c$, so that, at least at level of singular costs and primal formulations, these are truly generalizations of the (extended) Optimal Transport problem. Moreover, in analogy with what happens for EOT_c, the primal formulation-cost $\mathscr{U}_H$ enjoys nice properties, we have equality between $\overline{\text{co}}\,(\mathscr{S}_H)$ and $\mathscr{U}_H$, and we can prove a duality formula (see Proposition 4.1.3 and Theorems 4.1.4 and 4.2.4).

Theorem 1.1.1. *Let X_1, X_2 be Polish spaces. Then for every $(\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$, there exists $\boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2)$ such that*

$$\mathscr{U}_H(\mu_1, \mu_2) = \int_{\mathfrak{C}[X_1] \times \mathfrak{C}[X_2]} H \, d\boldsymbol{\alpha}.$$

Moreover $\mathscr{U}_H$ is a lower semicontinuous convex function such that

$$\mathscr{U}_H(r_1\delta_{x_1}, r_2\delta_{x_2}) \leqslant H([x_1, r_1]; [x_2, r_2]) \tag{1.1.13}$$

for every $(x_1, x_2) \in X_1 \times X_2$ *and every* $(r_1, r_2) \in \mathbb{R}^2_+$*. If, in addition,* H *is also convex, meaning that the map sending* $(r_1, r_2) \in \mathbb{R}^2_+$ *to* $H([x_1, r_1], [x_2, r_2])$ *is convex for every fixed* $(x_1, x_2) \in X_1 \times X_2$*, then the above inequality is an equality. Moreover*

$$\mathscr{U}_H(\mu_1, \mu_2) = \overline{\mathrm{co}}\,(\mathscr{S}_H)\,(\mu_1, \mu_2) = \sup\{\mathcal{D}(\varphi_1, \varphi_2; \mu_1, \mu_2) \mid (\varphi_1, \varphi_2) \in \Phi_H\} \quad (1.1.14)$$

for every $(\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$*, where*

$$\Phi_H := \left\{ \begin{array}{l} (\varphi_1, \varphi_2) \in C_b(X_1) \times C_b(X_2) \text{ s.t.} \\ \varphi_1(x_1)r_1 + \varphi_2(x_2)r_2 \leqslant H([x_1, r_1], [x_2, r_2]) \\ \text{for every } (x_1, x_2) \in X_1 \times X_2,\ r_1, r_2 \geqslant 0 \end{array} \right\}.$$

The lower semicontinuity of the cost function $\mathscr{U}_H$ follows by two considerations: the first one, which is strongly based on the 1-homogeneity of H, concerns the possibility of carrying out the minimization procedure among those $\alpha \in \mathfrak{H}^1(\mu_1, \mu_2)$ with support contained in $\{([x_1, r_1], [x_2, r_2]) \mid r_1, r_2 \leqslant R\}$ for some $R > 0$; the second fact is the nice dependence of the set $\alpha \in \mathfrak{H}^1(\mu_1, \mu_2)$ w.r.t. the narrow convergence, as it happens for the canonical set of plans $\Gamma(\mu_1, \mu_2)$ (for the details see Lemma 3.2.6).
The proof of the equality $\mathscr{U}_H(\mu_1, \mu_2) = \overline{\mathrm{co}}\,(\mathscr{S}_H)\,(\mu_1, \mu_2)$ can be carried out in two ways: the first constructive proof (see Theorem 4.1.4) exploits an explicit characterization of $\overline{\mathrm{co}}\,(\mathscr{S}_H)$ and the density of discrete measures to show "by hand" the equality $\overline{\mathrm{co}}\,(\mathscr{S}_H) = \mathscr{U}_H$. The second approach is the same used for the equality (1.1.10) in [109], i.e. just a simple application of the Fenchel-Moreau theorem. Notice that the duality result is completely analogous to the classical Kantorovich duality (1.1.2) and reduces to it in case H has the form in (1.1.12).

Adopting a slightly different point of view (that involves sufficiently rich subalgebras of continuous and bounded functions, see Lemma 3.1.6) in the proof of the above result we can show a reinforcement of (1.1.14) (resp. of (1.1.2)): we can obtain duality formulas with smooth C^∞ functions in finite dimensional Euclidean spaces, Lipschitz functions in metric spaces or smooth cylindrical functions in topological vector spaces. Following the path given by the classical Optimal Transport theory, it is then natural to investigate the existence of potentials in a sufficiently regular setting; to this aim we present two different situations where, assuming that the spaces X_1, X_2 are compact and that H is continuous, 1-homogeneous and convex, it is possible to prove such existence:

1. If H is finite on the whole product cone $\mathfrak{C}[X_1] \times \mathfrak{C}[X_2]$, it is enough to assume that H satisfies a few integrability conditions w.r.t. μ_1 and μ_2 and to have some control on the derivatives of H at the boundary of the product cone. For a detailed discussion see Section 5.1.1.

2. If H is finite only on a smaller cone (depending on the ratio $\mu_1(X_1)/\mu_2(X_2)$), it is sufficient to assume that H diverges to $+\infty$ on the boundary of such smaller cone in a uniform way. For the details see Section 5.1.2.

In both these situations it is possible to define the analogous of the c-transform as in (1.1.4) for a pair $(\varphi_1, \varphi_2) \in \Phi_H$ as

$$\varphi_1^H(x_2) := \inf_{x_1 \in X_1} \inf_{\alpha \geqslant 0} \Big\{ H([x_1, \alpha], [x_2, 1]) - \alpha\varphi_1(x_1) \Big\}, \quad x_2 \in X_2,$$
$$\varphi_1^{HH}(x_1) := \inf_{x_2 \in X_2} \inf_{\alpha \geqslant 0} \Big\{ H([x_1, 1], [x_2, \alpha]) - \alpha\varphi_1^H(x_2) \Big\}, \quad x_1 \in X_1,$$

and prove that the transformed potentials enjoy sufficiently nice properties to use a Ascoli-Arzelà argument. The result is the following (see Theorem 5.1.5).

Theorem 1.1.2. *Assume that X_1, X_2 are compact, that H is continuous, 1-homogeneous, convex and that one of the two settings above is satisfied. Then, there exists a pair $(\varphi_1, \varphi_2) \in \Phi_H$ such that*

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 = \mathscr{U}_H(\mu_1, \mu_2).$$

As for the classical Optimal Transport theory, the existence of regular potentials can be a powerful tool to prove the existence of an Optimal Transport map, under a few additional assumptions on the cost function H, involving its differentiability properties. We thus obtain the following result (see Theorem 5.1.6).

Theorem 1.1.3. *Let $K \subset \mathbb{R}^d$ be a compact and convex set with nonempty interior, let $H : \mathfrak{C}[K] \times \mathfrak{C}[K] \to [0, +\infty)$ be a 1-homogeneous and convex function which is in addition differentiable and Lipschitz continuous on the product cone w.r.t. the cone distance induced by the Euclidean one (see (3.2.4)). Let $\mu_i \in \mathcal{M}_+(K)$ with $\operatorname{supp} \mu_i = K$, $i = 1, 2$; if μ_1 is absolutely continuous w.r.t. $\mathcal{L}^d|_K$ (the d dimensional Lebesgue measure on K) and*

$$\text{for every } x_1 \in \operatorname{int}(K) \text{ the map}$$
$$\mathfrak{C}[K] \ni [y, q] \mapsto \begin{pmatrix} \partial_1 H(x_1, 1; y, q) \\ \partial_2 H(x_1, 1; y, q) \end{pmatrix} \in \mathbb{R}^{d+1} \text{ is invertible,}$$

then there exists a Borel map $T : K \to \mathfrak{C}[K]$ s.t.

$$\mu_2 = (x_\sharp \circ T_\sharp)(\mu_1), \qquad \int_K H([x_1, 1], T(x_1)) \, d\mu_1(x_1) = \mathscr{U}_H(\mu_1, \mu_2).$$

The proof of this result is in the same spirit of to the one of the classical Optimal Transport theory, but several more complicated technical aspects have to be taken into account.

As in the classical Optimal Transport case [5, Theorem 6.1.4], we can investigate general optimality conditions. It is not surprising that also in the unbalanced case the concept of cyclical monotonicity is crucial. In particular, the 1-homogeneity of the cost function H and the cone structure allow to lift the cyclical monotonicity from the support of an admissible plan $\boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2)$ to the convex cone that it generates. More precisely, given $\Gamma \subset \mathfrak{C}[X_1] \times \mathfrak{C}[X_2]$ and $(x_1, x_2) \in X_1 \times X_2$, we define the (x_1, x_2)-*section* of Γ as

$$\Gamma_{x_1,x_2} := \left\{(r_1, r_2) \in \mathbb{R}^2_+ \mid ([x_1, r_1], [x_2, r_2]) \in \Gamma\right\}$$

and the *convex cone generated by* Γ as

$$\hat{\Gamma} := \bigcup_{(x_1,x_2)\in X_1\times X_2} \left\{([x_1, r_1], [x_2, r_2]) \mid (r_1, r_2) \in \hat{\Gamma}_{x_1,x_2}\right\},$$

where $\hat{\Gamma}_{x_1,x_2}$ is the convex cone in $\mathbb{R}^2$ generated by Γ_{x_1,x_2}. Analogously, for $(x_1, x_2) \in X_1 \times X_2$, we denote by H_{x_1,x_2} the map $(r_1, r_2) \mapsto H([x_1, r_1], [x_2, r_2])$. With this notation, we can state the following result (see Proposition 5.2.3 and Theorem 5.2.5).

Theorem 1.1.4. *Let $\mu_i \in \mathcal{M}_+(X_i)$ for $i = 1, 2$, and let $\boldsymbol{\alpha} \in \mathfrak{H}^1_H(\mu_1, \mu_2)$.*
If $\boldsymbol{\alpha}$ is optimal and $\int H \,d\boldsymbol{\alpha} < +\infty$, then $\boldsymbol{\alpha}$ is concentrated on a Borel subset $\Gamma \subset \mathfrak{C}[X_1] \times \mathfrak{C}[X_2]$ s.t. $\hat{\Gamma}$ is H-cyclically monotone.
On the other hand, let Γ be a Borel set on which $\boldsymbol{\alpha}$ is concentrated such that $\hat{\Gamma} \subset D(H)$ and let us suppose that the effective domain of H is independent of $(x_1, x_2) \in X_1 \times X_2$, meaning that $D(H_{x_1,x_2}) = D(H_{y_1,y_2})$ for every $(x_1, x_2), (y_1, y_2) \in X_1 \times X_2$. If we assume moreover that the following conditions are satisfied:

1. *there exists $([\bar{x}_1, \bar{r}_1], [\bar{x}_2, \bar{r}_2]) \in \hat{\Gamma}$ such that $(\bar{r}_1, \bar{r}_2) \in \operatorname{int}(D(\partial H_{\bar{x}_1,\bar{x}_2}))$,*
2. *there exist positive constants a_i, b_i, $i = 1, 2$ s.t.*

$$\begin{aligned} \mu_1\left(\left\{x_1 \in X_1 \mid \int_{X_2} H([x_1, a_1]; [x_2, b_1])\, d\mu_2(x_2) < +\infty\right\}\right) > 0,\\ \mu_2\left(\left\{x_2 \in X_2 \mid \int_{X_1} H([x_1, a_2]; [x_2, b_2])\, d\mu_1(x_1) < +\infty\right\}\right) > 0, \end{aligned} \tag{1.1.15}$$

then, if $\hat{\Gamma}$ is H-cyclically monotone, $\boldsymbol{\alpha}$ is optimal, $\int H \,d\boldsymbol{\alpha} < +\infty$ and there exists a maximizing pair $(\varphi_1, \varphi_2) \in L^1(X_1, \mu_1; \bar{\mathbb{R}}) \times L^1(X_2, \mu_2; \bar{\mathbb{R}})$ for the dual problem i.e.

$$\int_{X_1} \varphi_1 \,d\mu_1 + \int_{X_2} \varphi_2 \,d\mu_2 = \mathscr{U}_H(\mu_1, \mu_2).$$

Finally we treat the case in which H is (the p-th power of) a distance on $\mathfrak{C}[X]$: under this condition, we show that the resulting cost $\mathscr{U}_H$ is itself (the p-th power of) a distance on an appropriate subset of $\mathcal{M}_+(X)$ metrizing the weak convergence of measures, precisely as it is for the standard Optimal Transport problem [5, Proposition 7.1.5]. The result is the following (see Theorems 5.3.7 and 5.3.8).

Theorem 1.1.5. *Let X be a Polish space and let $H : \mathfrak{C}[X] \times \mathfrak{C}[X] \to [0, +\infty)$ be a lower semicontinuous and 1-homogeneous function which is the p-th power of a distance on $\mathfrak{C}[X]$ whose induced topology is stronger than the topology of $\mathfrak{C}[X]$. Then $\mathscr{U}_H$ is the p-th power of a distance on the subset $\mathcal{M}_{H,p}(X)$ of measures with finite p-th moment w.r.t. H defined as*

$$\mathcal{M}_{H,p}(X) := \left\{\mu \in \mathcal{M}_+(X) \mid \int_X H([x, 1], \mathfrak{o})\, d\mu(x) < +\infty\right\}.$$

Moreover, given a sequence $(\mu_n)_n \subset \mathcal{M}_{H,p}(X)$ *and* $\mu \in \mathcal{M}_{H,p}(X)$*, we have that*

$$\lim_{n\to+\infty} \mathscr{U}_H(\mu_n,\mu) = 0 \quad \Longleftrightarrow \quad \begin{cases} \mu_n \rightharpoonup \mu, \\ \int_X H([x,1];\mathfrak{o})\,d\mu_n(x) \to \int_X H([x,1];\mathfrak{o})\,d\mu(x). \end{cases}$$

The first part of the thesis is organized as follows: Chapter 3 is devoted to establish the general setting and a few technical tools that will be used in the sequel; Chapter 4 contains the core of our results: the convexification approach is presented and duality is treated; moreover a last section is devoted to the case of merely Hausdorff spaces, where a suitable definition of narrow topology has to be taken into account; finally Chapter 5 treats the optimality conditions and the dual attainment both in the general case and in the more regular one, together with a few remarks on the metric and topological properties of $\mathscr{U}_H$ in case H is (the p-th power of) a distance.

Part I is the result of a collaboration with Giuseppe Savaré and part of the material presented in Sections 3.1, 4.3 appeared in [109].

1.1.2 *Part II: Dissipative evolutions in Wasserstein spaces*

The study of gradient flow evolutions has always been a very relevant topic in analysis with many applications. An important framework for many PDEs models is the one of a convex[1], proper and lower semicontinuous function $f : H \to (-\infty,+\infty]$ in a Hilbert space H with norm $|\cdot|$. A gradient flow of f starting from $\bar{x}_0 \in D(f)$ is a locally absolutely continuous curve $x : [0,+\infty) \to H$ such that

$$\begin{cases} \dot{x}_t \in -\partial f(x_t) \quad \text{a.e. } t > 0, \\ x_0 = \bar{x}_0, \end{cases} \tag{1.1.16}$$

where $-\partial f(z)$ is the opposite of the subdifferential of f at a point $z \in D(f)$, defined as

$$v \in \partial f(z) \quad \text{if and only if} \quad f(y) - f(z) \geqslant \langle v, y-z\rangle \quad \text{for every } y \in H. \tag{1.1.17}$$

The generalization of gradient flows to general metric spaces has been a very interesting topic that started with the works of De Giorgi and his collaborators [41] and has attracted a growing interest in the Optimal Transport and PDEs communities since the works of Jordan, Kinderlerer and Otto [64, 90] where they noted a Wasserstein-gradient flows structure in some important evolution equations. Let us explain here the main ideas when the metric space under consideration is $(\mathcal{P}_2(\mathbb{H}), W_2)$, where $\mathbb{H}$ is a (possibly infinite dimensional) separable Hilbert space, $\mathcal{P}_2(\mathbb{H})$ is the space of Borel probability measures on $\mathbb{H}$ with finite second

1 A less restrictive notion of convexity, namely λ-convexity, $\lambda \in \mathbb{R}$, can be used in the Hilbertian and Wasserstein settings. While this will be considered in the thesis, we prefer to stick to the simpler convex setting in this introduction

moment and W_2 is the 2-Wassertstein distance induced by the norm $|\cdot|$ of $\mathbb{H}$. In order to formulate (1.1.16) in the Wasserstein setting, we need to find a way to express (1.1.16) without relying on the Hilbertian structure. The first step is to notice that, if we fix a point $y \in H$ and we compute the derivative of the square norm $|x_t - y|^2$ along a solution x of (1.1.16), we get

$$\frac{1}{2}\frac{d}{dt}|x_t - y|^2 = \langle \dot{x}_t, x_t - y\rangle \leqslant f(y) - f(x_t) \quad \text{a.e. } t > 0, \tag{1.1.18}$$

since $\dot{x}_t \in -\partial f(x_t)$. On the other hand, it is clear that, if a locally absolutely continuous curve $x : [0, +\infty) \to H$ satisfies (1.1.18), then it satisfies (1.1.16). Conditions like (1.1.18) are usually called Evolution Variation Inequality (EVI); as it was first noted in [5], the very nice thing is that (1.1.18) is completely independent from the Hilbertian structure and can be set in general: given $\phi : \mathcal{P}_2(\mathbb{H}) \to (-\infty, +\infty]$, we say that a locally absolutely continuous curve $\mu : [0, +\infty) \to X$ is a EVI-gradient flow of ϕ starting from $\bar{\mu}_0 \in D(\phi)$ if

$$\begin{cases} \frac{1}{2}\frac{d}{dt}W_2^2(\mu_t, \nu) \leqslant \phi(\nu) - \phi(\mu_t) & \text{for every } \nu \in \mathcal{P}_2(\mathbb{H}) \text{ and a.e. } t > 0, \\ \mu_0 = \bar{\mu}_0. \end{cases} \tag{1.1.19}$$

We refer to [5] for the general theory of gradient flows in metric spaces and, in particular, to the second part of the book, where the authors treat the Wasserstein case. Both in the metric and Wasserstein settings, the crucial tool is the minimizing movement scheme [41] (which in the present Wasserstein setting is also called JKO-scheme, according to [64]): let us define, for every $\mu \in \mathcal{P}_2(\mathbb{H})$ and $\tau > 0$, the map $\Phi(\tau, \nu; \cdot) : \mathcal{P}_2(\mathbb{H}) \to (-\infty, +\infty]$ as

$$\Phi(\tau, \mu; \nu) := \frac{1}{2\tau}W^2(\mu, \nu) + \phi(\nu), \quad \nu \in \mathcal{P}_2(\mathbb{H}).$$

Starting from the initial point $\bar{\mu}_0$, one iteratively defines the sequence

$$\begin{cases} \mu_0 := \bar{\mu}_0, \\ \mu_n \in \operatorname{arg\,min}_{\nu \in \mathcal{P}_2(\mathbb{H})} \Phi(\tau, \mu_{n-1}; \nu), \quad n \geqslant 1, \end{cases} \tag{1.1.20}$$

and the piece-wise constant interpolating curve $\mu_\tau : [0, +\infty) \to X$. The reason behind this scheme is the following: in the Hilbertian case, minimizers μ_n of $\Phi(\tau, \mu_{n-1}; \cdot)$ are exactly one step of Implcit Euler scheme for ϕ; the corresponding sequence $(x_n)_n \subset H$ defined as in (1.1.20) for $f : H \to (-\infty, +\infty)$ would indeed satisfy

$$\frac{x_n - x_{n-1}}{\tau} \in -\partial f(x_n). \tag{1.1.21}$$

Under suitable hypotheses on ϕ, it is possible to prove (see [5, Chapter 11] that the curves $(\mu_\tau)_{\tau>0}$ have at least an accumulation point μ as $\tau \downarrow 0$ in an appropriate topology and that μ solves (1.1.19). In addition, exploiting the particular structure of this space, in [5], the authors give a notion of subdifferential for a proper, lower semicontinuous and geodesically convex functional

$\phi : \mathcal{P}_2(\mathbb{H}) \to (-\infty, +\infty]$: we denote by $\mathbb{TH} := \mathbb{H} \times \mathbb{H}$ the (flat) tangent bundle to $\mathbb{H}$ and, given $\mu \in D(\phi)$, we say that $\Psi \in \mathcal{P}_2(\mathbb{TH})$ belongs to $\boldsymbol{\partial}\phi(\mu)$, the subdifferential of ϕ at μ, if

$$\text{for every } \nu \in \mathcal{P}_2(\mathbb{H}) \text{ there exists } \boldsymbol{\sigma} \in \Lambda(\Psi, \nu) \text{ such that}$$
$$\phi(\nu) - \phi(\mu) \geqslant \int_{\mathbb{TH}\times\mathbb{H}} \langle v, y - x \rangle \, d\boldsymbol{\sigma}(x, v, y), \tag{1.1.22}$$

where $\Lambda(\Psi, \nu)$ is the set of $\boldsymbol{\sigma} \in \mathcal{P}(\mathbb{TH} \times \mathbb{H})$ such that $(x, v)_\sharp \boldsymbol{\sigma} = \Psi$ and $(x, y)_\sharp \boldsymbol{\sigma} \in \Gamma_o(\mu, \nu)$, the set of Optimal Transport plans connecting μ to ν. It is clear that this definition is an integrated version of the classical (1.1.17). Notice moreover that here an element of $\boldsymbol{\partial}\phi(\mu)$ is defined as a probability on the product space, which would correspond to say that in (1.1.17) $(x, v) \in -\partial f(x)$.

The second building block to better characterize solutions to (1.1.16) is a notion of velocity for an absolutely continuous curve in $\mathcal{P}_2(\mathbb{H})$ which has to be somehow compatible with (1.1.22); this is provided by [5, Theorem 8.3.1, Proposition 8.4.6]: if $\mu : [0, +\infty) \to \mathcal{P}_2(\mathbb{H})$ is a locally absolutely continuous curve, then there exists a (a.e. uniquely determined) Borel vector field $\boldsymbol{v} : [0, +\infty) \times \mathbb{H} \to \mathbb{H}$ such that the continuity equation

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = 0$$

holds in the sense of distributions in $[0, +\infty) \times \mathbb{H}$ and

$$\lim_{h \to 0} \frac{W_2((\boldsymbol{i}_{\mathbb{H}} + h\boldsymbol{v}_t)_\sharp \mu_t, \mu_{t+h})}{|h|} = 0 \quad \text{for a.e. } t > 0, \tag{1.1.23}$$

where $\boldsymbol{i}_{\mathbb{H}}$ is the identity map on $\mathbb{H}$.
With the notions of velocity and subdifferential at our disposal, we can finally characterize a gradient flow evolution in the Wassestein space: we say that a locally absolutely continuous curve $\mu : [0, +\infty) \to \mathcal{P}_2(\mathbb{H})$ is a Wasserstein-gradient flow for ϕ starting from $\bar{\mu}_0 \in D(\phi)$ if

$$\begin{cases} (\boldsymbol{i}_{\mathbb{H}}, -\boldsymbol{v}_t)_\sharp \mu_t \in \boldsymbol{\partial}\phi(\mu_t) & \text{a.e. } t > 0, \\ \mu_0 = \bar{\mu}_0, \end{cases} \tag{1.1.24}$$

where $\boldsymbol{v}_t$ is the vector field associated to the curve μ as above evaluated at time t. The remarkable result [5, Theorem 11.1.3] is that the notion of EVI-gradient flow in (1.1.19) and the Wasserstein one in (1.1.24) coincide, under suitable hypotheses on the functional ϕ.

Since the aim of this second part of the thesis is to study dissipative evolutions in Wasserstein spaces, let us spend a few words on what is well known: a particular example of dissipative operator is the (opposite of the) subdifferential of a convex function $f : H \to (-\infty, +\infty]$, meaning that

$$\langle v - w, x - y \rangle \leqslant 0 \quad \text{for every } (x, v), (y, w) \in -\partial f, \tag{1.1.25}$$

where we are denoting by $-\partial f$ the graph of the multivalued map $x \mapsto -\partial f(x)$ (equivalently $(x, v) \in -\partial f$ if and only if $v \in -\partial f(x)$). It is then natural to consider general dissipative multivalued maps $F : H \to 2^H$ sending a point $x \in H$ to a subset $F(x) \subset H$ satisfying

$$\langle v - w, x - y \rangle \leqslant 0 \quad \text{for every } (x, v), (y, w) \in F, \tag{1.1.26}$$

where with abuse of notation we are denoting by F both the multivalued operator and its graph: $(x, v) \in F$ if $v \in F(x)$. Moreover, it is natural to consider the component v as a velocity, so that we can identify F as a subset of $TH = H \times H$. The corresponding dissipative evolution is thus described by a locally absolutely continuous curve $x : [0, +\infty) \to H$ satisfying

$$\begin{cases} \dot{x}_t \in F(x_t) & \text{a.e. } t > 0, \\ x_0 = \bar{x}_0, \end{cases} \tag{1.1.27}$$

for some $\bar{x}_0 \in D(F)$. These evolutions have been a widely studied topic and we refer to the book of Brézis [26] for a discussion of this theory. The strategy to prove existence of a curve satisfying (1.1.27) in the classical setting relies on the notion of resolvent operator J_τ of F which encodes a step of implicit Euler scheme for F: given $x \in H$ and $\tau > 0$, we say that $y \in J_\tau(x)$ if and only if

$$\frac{y - x}{\tau} \in F(y)$$

which corresponds exactly to (1.1.21) with F in place of $-\partial f$. The crucial point is that, if F is maximal dissipative (meaning that it is maximal in the sense of the graph in the class of dissipative operators), then the resolvent operator is an everywhere defined contraction and it is thus single-valued so that the implicit Euler scheme is well defined. This allows to develop the theory of dissipative evolutions in Hilbert spaces and have existence, uniqueness and stability results for (1.1.27) (see e.g. [26, Theorem 3.1]), thus defining the semigroup operator $S_t : \overline{D(F)} \to H$ associated to F: this is the (unique Lipschitz extension to $\overline{D(F)}$ of the) map sending $\bar{x}_0 \in D(F)$ to the evaluation at time $t > 0$ of the solution x of (1.1.27).

Contributions in Part II. The aim of this second part of the thesis is thus to present a new class of dissipative evolutions in $\mathcal{P}_2(\mathbb{H})$ in the same spirit of the generalization of (1.1.16) to (1.1.24).
In analogy with the metric extension of gradient flow evolutions, the first step is to express a dissipative evolution in a purely metric way (or, at least, in a suitable way for the space $\mathcal{P}_2(\mathbb{H})$). We start again from a natural characterization of (1.1.27) in terms of Evolution Variational Inequality: if we take $(y, w) \in F \subset H \times H$ and we compute the derivative of the squared norm $|x_t - y|^2$ along a solution x of (1.1.27), we get

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} |x_t - y|^2 &= \langle \dot{x}_t, x_t - y \rangle = \langle \dot{x}_t - w, x_t - y \rangle - \langle w, y - x_t \rangle \\ &\leqslant -\langle w, y - x_t \rangle \qquad \text{for a.e. } t > 0, \end{aligned} \tag{1.1.28}$$

since $\dot{x}_t \in F(x_t)$ and F is dissipative. If F is also maximal dissipative, it is easy to see that (1.1.28) is also sufficient to get (1.1.27); thus (1.1.28) is an equivalent way to express that x is the solution of (1.1.28) and we can observe that, even if it is not immediately apparent, this is again a (almost) purely metric formulation: if we consider the exponential map $\exp^s : H \times H \to H$ sending (x, v) to $x + sv$, we can immediately notice that

$$\frac{1}{2}\frac{d}{ds}\Big|_{s=0} |\exp^s(x,v) - y|^2 = \langle v, x - y\rangle$$

so that (1.1.28) can be equivalently expressed as

$$\frac{1}{2}\frac{d}{dt}|x_t - y|^2 \leqslant -\frac{1}{2}\frac{d}{ds}\Big|_{s=0} |\exp^s(y,w) - x_t|^2 \quad \text{for every } (y,w) \in F \text{ and a.e. } t > 0. \tag{1.1.29}$$

Moreover notice that (1.1.26) can be rewritten in terms of derivatives of the squared distance as

$$\frac{1}{2}\frac{d}{ds}\Big|_{s=0} |\exp^s(x,v) - \exp^s(y,w)|^2 \leqslant 0 \quad \text{for every } (x,v),(y,w) \in F. \tag{1.1.30}$$

The problem of extending (1.1.28) and the notion of dissipativity to $\mathcal{P}_2(\mathbb{H})$ thus reduces to find a suitable notion of exp in the space $\mathcal{P}_2(\mathsf{T}\mathbb{H})$ and to study its differentiability properties. The natural candidate is of course the map $\exp^s_\sharp : \mathcal{P}_2(\mathsf{T}\mathbb{H}) \to \mathcal{P}_2(\mathbb{H})$ and the result is the following (see Proposition 7.1.3 and Theorem 7.1.8).

Theorem 1.1.6. *Let $\Phi_0, \Phi_1 \in \mathcal{P}_2(\mathsf{T}\mathbb{H})$ be measures with marginals $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{H})$. Then the maps*

$$s \mapsto \frac{1}{2}W_2^2(\exp^s_\sharp \Phi_0, \mu_1), \quad s \mapsto \frac{1}{2}W_2^2(\exp^s_\sharp \Phi_0, \exp^s_\sharp \Phi_1)$$

are semi-concave and thus right and left differentiable at $s = 0$. If we define

$$[\Phi_0, \mu_1]_r := \lim_{s\downarrow 0} \frac{W_2^2(\exp^s_\sharp \Phi_0, \mu_1) - W_2^2(\mu_0, \mu_1)}{2s}, \tag{1.1.31}$$

$$[\Phi_0, \Phi_1]_r := \lim_{s\downarrow 0} \frac{W_2^2(\exp^s_\sharp \Phi_0, \exp^s_\sharp \Phi_1) - W_2^2(\mu_0, \mu_1)}{2s}, \tag{1.1.32}$$

we have that

$$[\Phi_0, \mu_1]_r = \min\left\{\int_{\mathsf{T}\mathbb{H}\times\mathbb{H}} \langle x_0 - x_1, v_0\rangle \, d\boldsymbol{\sigma} \mid \boldsymbol{\sigma} \in \Lambda(\Phi_0, \mu_1)\right\}, \tag{1.1.33}$$

$$[\Phi_0, \Phi_1]_r = \min\left\{\int_{\mathsf{T}\mathbb{H}\times\mathsf{T}\mathbb{H}} \langle x_0 - x_1, v_0 - v_1\rangle \, d\boldsymbol{\Theta} \mid \boldsymbol{\Theta} \in \Lambda(\Phi_0, \Phi_1)\right\}, \tag{1.1.34}$$

where the sets $\Lambda(\Phi_0, \Phi_1)$ and $\Lambda(\Phi_0, \mu_1)$ are defined as

$$\Lambda(\Phi_0, \mu_1) := \left\{\boldsymbol{\sigma} \in \Gamma(\Phi_0, \mu_1) \mid (x^0, x^1)_\sharp \boldsymbol{\sigma} \in \Gamma_o(\mu_0, \mu_1)\right\},$$

$$\Lambda(\Phi_0, \Phi_1) := \left\{\boldsymbol{\Theta} \in \Gamma(\Phi_0, \Phi_1) \mid (x^0, x^1)_\sharp \boldsymbol{\Theta} \in \Gamma_o(\mu_0, \mu_1)\right\}.$$

Notice that the subscript “r" stands for “right" since in the thesis we will also consider the left derivative, which is however not needed in this introductory part. We will then call a non-empty set $\mathbf{F} \subset \mathcal{P}_2(\mathsf{T}\mathbb{H})$ “Multivalued Probability Vector Field" (MPVF), we say that it is dissipative if

$$[\Phi_0, \Phi_1]_r \leqslant 0 \quad \text{for every } \Phi_0, \Phi_1 \in \mathbf{F} \tag{1.1.35}$$

and that a locally absolutely continuous curve $\mu : [0, +\infty) \to \mathcal{P}_2(\mathbb{H})$ is a EVI-solution for $\mathbf{F}$ starting from $\bar{\mu}_0 \in D(\mathbf{F})$ if

$$\begin{cases} \frac{1}{2}\frac{d}{dt} W_2^2(\mu_t, x_\sharp \Phi) \leqslant -[\Phi, \mu_t]_r & \text{for every } \Phi \in \mathbf{F} \text{ and a.e. } t > 0, \\ \mu_0 = \bar{\mu}_0. \end{cases} \tag{1.1.36}$$

The characterizations in (1.1.33) and (1.1.34) are particularly useful to study semicontinuity properties of the pairings and since, by (1.1.23), every locally absolutely continuous curve $\mu : \mathcal{I} \to \mathcal{P}_2(\mathbb{H})$ behaves locally as $\exp_\sharp^h \Phi_t$ for $\Phi_t = (\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t$, it is not surprising that we can characterize the directional derivatives of the squared Wasserstein distance from a fixed point along absolutely continuous curves (see Theorems 7.2.1 and 7.2.3). This is also important when studying the stability properties of the notion of EVI solutions: the possibility of expressing derivatives as pairings allows also to show the robustness of the notion of solution (see e.g. Proposition 8.1.6).

Notice moreover that, thanks to the result above, we can immediately see that the (opposite of the) subdifferential of a convex and lower semicontinuous functional $\phi : \mathcal{P}_2(\mathbb{H}) \to (-\infty, +\infty]$ can be characterized in the following way: given $\mu \in D(\phi)$, $\Phi \in \mathcal{P}_2(\mathsf{T}\mathbb{H})$ belongs to $-\boldsymbol{\partial}\phi(\mu)$ (meaning that $(x, -v)_\sharp \Phi \in \boldsymbol{\partial}\phi(\mu)$, where $(x, -v)(x, v) = (x, -v)$ for every $(x, v) \in \mathbb{H} \times \mathbb{H}$) if and only if $x_\sharp \Phi = \mu$ and

$$\phi(\nu) - \phi(\mu) \geqslant [\Phi, \nu]_r \text{ for every } \nu \in \mathcal{P}_2(\mathbb{H}).$$

Finally it is not difficult to see that $-\boldsymbol{\partial}\phi$ is a dissipative operator, so that it is natural to compare the notions of gradient flow for ϕ with the one of dissipative evolution for $-\boldsymbol{\partial}\phi$: the result (see Proposition 9.5.2) is that, if e.g. the domain of $\boldsymbol{\partial}\phi$ is geodesically convex, then the two notions coincide.

Of course the framework we are developing here is not the first attempt of defining evolutions of measures under the action of general notions of velocity fields: we mention here two approaches that are in a way connected to ours (the one of Piccoli [94] was partly an inspiration for our work).

The idea of both approaches is to consider maps $\boldsymbol{b} : \mathcal{P}_2(\mathbb{H}) \to C(\mathbb{H}; \mathbb{H})$, taking values in some subset of continuous vector fields in $\mathbb{H}$. The evolution driven by $\boldsymbol{b}$ is thus described by $t \mapsto \mu_t \in \mathcal{P}_2(\mathbb{H})$ such that the continuity equation

$$\partial_t \mu_t + \operatorname{div}(\boldsymbol{v}_t \mu_t) = 0, \qquad \boldsymbol{v}_t = \boldsymbol{b}[\mu_t], \quad \mu_t\text{-a.e. for every } t > 0, \tag{1.1.37}$$

holds in the distributional sense. In particular, in [21] the aim of the authors is to develop a suitable Cauchy-Lipschitz theory in Wasserstein spaces for differential inclusions which generalizes (1.1.37) to multivalued maps $\boldsymbol{b} : \mathcal{P}_b(\mathbb{H}) \rightrightarrows$

$\mathrm{Lip}_{loc}(\mathbb{H};\mathbb{H})$ and requires (1.1.37) to hold for a suitable measurable selection of $\mathbf{b}$. On the other hand, in [94], the map $\mathbf{b}$ is defined in terms of a single valued MPVF, meaning that, given a map $\mathbf{F} : \mathcal{P}_2(\mathbb{H}) \to \mathcal{P}_2(\mathsf{T}\mathbb{H})$, $\mathbf{b}[\mu]$ is defined as the barycenter of $\mathbf{F}[\mu]$:

$$\mathbf{b}[\mu](x) := \int_{\mathbb{H}} v \, d(\mathbf{F}[\mu])_x(v), \quad x \in \mathbb{H},$$

where $(\mathbf{F}[\mu])_x$ is the disintegration of $\mathbf{F}[\mu]$ w.r.t. μ at $x \in \mathbb{H}$ (see Theorem 2.1.1 and Definition 7.1.1). In other words, the notion of evolution given in [94] can be expressed in our framework in the following way: if $\mathbf{F}$ is a MPVF, we say that a locally absolutely continuous curve $\mu : [0, +\infty) \to \mathrm{D}(\mathbf{F})$ satisfies the *barycentric property* if for a.e. $t > 0$ there exists $\Phi_t \in \mathbf{F}[\mu_t]$ such that

$$\frac{d}{dt}\int_{\mathbb{H}} \varphi(x) \, d\mu_t(x) = \int_{\mathsf{T}\mathbb{H}} \langle \nabla\varphi(x), v\rangle \, d\Phi_t(x,v) \quad \text{for every } \varphi \in \mathrm{Cyl}(\mathbb{H}), \tag{1.1.38}$$

where $\mathrm{Cyl}(\mathbb{H})$ is the set of cylindrical function on $\mathbb{H}$, a generalization of smooth functions with compact support to the infinite dimensional space $\mathbb{H}$ (see Definition 2.4.5). The following result, which we state here in a simplified way, compares the two notions of solutions (see Theorems 8.3.4 and 8.3.7 for the more general statements).

Theorem 1.1.7. *Let* $\mathbf{F}$ *be a dissipative MPVF and let* $\mu : [0, +\infty) \to \mathrm{D}(\mathbf{F})$ *be a locally absolutely continuous curve. Then*

1. *If* $\mathrm{D}(\mathbf{F}) = \mathcal{P}_2(\mathbb{H})$, $\mathbf{F}$ *is sequentially closed in* $\mathcal{P}_2(\mathsf{T}\mathbb{H})$, *its sections are convex (meaning that* $\mathbf{F}[\nu]$ *is convex for every* $\nu \in \mathcal{P}_2(\mathbb{H})$*) and* $\mathbf{F}$ *is locally bounded, in the sense that for every* $\mu \in \mathrm{D}(\mathbf{F})$ *there exist constants* $M, \varepsilon > 0$ *such that*

$$\textit{for every } \nu \in \mathcal{P}_2(\mathbb{H}) \textit{ with } W_2(\mu,\nu) < \varepsilon$$
$$\textit{there exists } \Phi \in \mathbf{F}[\nu] \textit{ with } \int_{\mathbb{H}} |v|^2 \, d\Phi(x,v) \leqslant M,$$

   then every EVI solution for $\mathbf{F}$ *has the barycentric property.*

2. *If* μ_t *is a regular probability measure ([5, Definitions 6.2.1, 6.2.2]) for a.e.* $t > 0$ *and the curve* μ *has the barycentric property, then it is a EVI solution for* $\mathbf{F}$.

Let us now come to the question of existence of EVI solutions for a dissipative MPVF. As we have seen, both in the classical setting of Hilbertian dissipative evolutions and of gradient flows in $\mathcal{P}_2(\mathbb{H})$, the main tool is an implicit Euler scheme. In the first case it is possible to construct it thanks to the maximality of the operator (which is equivalent to the global definition of the resolvent), while in the second case it relies on the variational minimizing movements-JKO interpretation of one step of implicit Euler scheme. Both these strategies are not immediately clear in the present setting of dissipative evolutions in $\mathcal{P}_2(\mathbb{H})$ so that we will first carry out the analysis of an explicit Euler scheme, which has the advantage of being easily defined, while we leave for the last part of this introduction the more involved construction of the implicit Euler scheme.

The explicit Euler method. We fix a finite time horizon $T > 0$, a time step $\tau > 0$ and we divide the interval $[0, T]$ in $N(\tau, T) := \lceil T/\tau \rceil$ sub-intervals. Starting from the initial datum $\bar{\mu}_0$ we construct the following sequence

$$M^0_\tau := \bar{\mu}_0, \quad M^{n+1}_\tau := \mathsf{exp}^\tau_\sharp \Phi^n_\tau = (x + \tau v)_\sharp \Phi^n_\tau, \quad \Phi^n_\tau \in \mathbf{F}[M^n_\tau] \tag{1.1.39}$$

and we consider the constant interpolating curve $\bar{M}_\tau : [0, T] \to \mathcal{P}_2(\mathbb{H})$ defined as

$$\bar{M}_\tau(t) := M^{\lfloor t/\tau \rfloor}, \quad t \in [0, T].$$

Obviously, one wants to study the properties of the family of curves $(\bar{M}_\tau)_{\tau>0}$ and hope that, as $\tau \downarrow 0$, we obtain an EVI solution for $\mathbf{F}$. Differently from the implicit Euler method, although $\mathbf{F}$ is dissipative, at each step of the explicit Euler scheme we obtain a perturbation of the distance given by

$$W_2^2(\mathsf{exp}^\tau_\sharp \Phi, \mathsf{exp}^\tau_\sharp \Psi) \leqslant W_2^2(\mu, \nu) + 2\tau [\Phi, \Psi]_r + \tau^2 \left(|\Phi|_2^2 + |\Psi|_2^2 \right),$$

where

$$|\Phi|_2^2 := \int_{\mathsf{T}\mathbb{H}} |v|^2 \, d\Phi(x, v),$$

which thus depends on the second moments of Φ and Ψ, and thus of the magnitude of $\mathbf{F}$ at μ and ν. However, if we impose to the discrete sequence in (1.1.39) to satisfy

$$|\Phi^n_\tau|_2 \leqslant L \quad \text{if } 0 \leqslant n \leqslant N(\tau, T) \tag{1.1.40}$$

for some $L \geqslant 0$, then we can immediately obtain a discrete version of EVI for the family M^n_τ (see Proposition 9.1.3)

$$\frac{1}{2} W_2^2(M^{n+1}_\tau, \nu) - \frac{1}{2} W_2^2(M^n_\tau, \nu) \leqslant \tau [\Phi^n_\tau, \nu]_r + \frac{1}{2} \tau^2 L^2 \tag{1.1.41}$$

for every $0 \leqslant n < N(T, \tau)$ and $\nu \in \mathcal{P}_2(\mathbb{H})$. Heuristically, this tells us that passing to the limit as $\tau \downarrow 0$, we will obtain a solution to EVI. Adapting to the $\mathcal{P}_2(\mathbb{H})$-setting the relaxation approach of [89], based on the doubling variable technique of Kružkov [71] and Crandall-Evans [38] (see also [39]), we prove the main convergence result for the explicit Euler scheme (see Theorems 9.1.5, 9.1.6 and 9.1.8).

Theorem 1.1.8. *Let* $\mathbf{F}$ *be a dissipative MPVF and let* $T > 0$*; let* $(M_\tau)_\tau$ *be a family of constant interpolating curves corresponding to the time steps* $0 < \tau < 1$ *in the interval* $[0, T]$ *and to the initial datum* $\bar{\mu}_0$ *as in* (1.1.39)*. Suppose moreover that there exists some* $L \geqslant 0$ *such that all the sequences* $(M^m_\tau)_{m=0,\dots,N(T,\tau)}$ *satisfy* (1.1.40)*. Then* $(M_\tau)_\tau$ *uniformly converges as* $\tau \downarrow 0$ *to a Lipschitz continuous curve* $\mu : [0, T] \to \mathcal{P}_2(\mathbb{H})$ *which is the unique EVI solution for* $\mathbf{F}$ *starting from* $\bar{\mu}_0$*. Moreover the following error estimate holds:*

$$W_2(\mu_t, \bar{M}_\tau(t)) \leqslant CL\sqrt{\tau(t + \tau)} \quad t \in [0, T], \, n \in \mathbb{N}, \tag{1.1.42}$$

where C *is a universal constant.*

We want to highlight that the estimate in (1.1.42) is sharp [106], does not require any local compactness assumption on the underlying space, and reproduces the celebrated Crandall-Liggett estimate for the generation of dissipative semigroups in Banach spaces [39] in this Wasserstein-metric framework.

Moreover, if μ, ν are two limit solutions starting from μ_0, ν_0 we prove that

$$W_2(\mu_t, \nu_t) \leqslant W_2(\mu_0, \nu_0) \quad \text{for every } t \in [0, T],$$

as it happens in the case of gradient flows of geodesically convex functions.

With the explicit Euler scheme and the notion of EVI solution, it also possible to prove local and global existence results as it is done in the classical ODEs theory, see Section 9.2.

The implicit Euler method. As we have seen, in the context of contraction semigroups generated by dissipative operators in Hilbert spaces, a fundamental role is played by the implicit Euler scheme, which, differently from the explicit one, is unconditionally stable, and thus avoids imposing local boundedness conditions as in (1.1.40). While in the case of gradient flows the construction of the implicit Euler scheme relies on the variational definition as in (1.1.20), if the dissipative operator $\mathbf{F}$ doesn't come from a functional ϕ, it is not possible to use an analogous construction.

The approach for the implicit Euler scheme we present is as follows:

1. We assume that the domain of $\mathbf{F}$ contains a sufficienlty rich set of discrete measures;
2. We "lift" the structure of $\mathcal{P}_2(\mathbb{H})$ to an Hilbert space $\mathcal{H}$ parametrizing measures by random variables defined in an appropriate space;
3. Starting from discrete measures, we define a maximal dissipative operator $F \subset \mathcal{H} \times \mathcal{H}$ which is compatible with $\mathbf{F}$ in a suitable sense;
4. We prove that the dissipative evolution driven by F in $\mathcal{H}$ induces a EVI evolution for $\mathbf{F}$ in $\mathcal{P}_2(\mathbb{H})$.

In particular, point 4. can also be seen in the following way: we build the implicit Euler scheme for F in $\mathcal{H}$ and this induces an implicit Euler scheme for $\mathbf{F}$ in $\mathcal{P}_2(\mathbb{H})$.

We start from a naive guess: we could try to consider

$$F := \{(X, V) \in \mathcal{H} \times \mathcal{H} : (X, V)_\sharp \mathbb{P} \in \mathbf{F}\}, \tag{1.1.43}$$

where $\mathcal{H}$ is any Hilbert space of random variables parametrizing $\mathcal{P}_2(\mathbb{H})$. Unfortunately

$$\langle V - W, X - Y \rangle \geqslant \left[(X, V)_\sharp \mathbb{P}, (Y, W)_\sharp \mathbb{P}\right]_r, \tag{1.1.44}$$

and in general the inequality can be strict, so that this definition does not guarantee the dissipativity of F, even if $\mathbf{F}$ is.

As stated in the first point, we suppose that the domain $D(\mathbf{F})$ contains a nonempty set of discrete measures $C \subset D_f(\mathbf{F})$, where $D_f(\mathbf{F})$ is the subset of $D(\mathbf{F})$ of measures with finite support. Under this assumption, we can look at the properties of $\mathbf{F}$ when computed on discrete measures; consider $\Phi, \Psi \in \mathbf{F}$ and suppose that $\mu := x_\sharp \Phi \in C$ has finite support, meaning that it is concentrated on a set of N distinct points for some $N \in \mathbb{N}$. If also $\nu := x_\sharp \Psi \in C$ is concentrated on the same number N of (non necessarily distinct) points, then there is a unique Optimal Transport plan $\boldsymbol{\gamma} \in \Gamma(\mu, \nu)$ and it is concentrated on a map, meaning that there exists a Borel map $T : \mathbb{H} \to \mathbb{H}$ such that $\boldsymbol{\gamma} = (\boldsymbol{i}_{\mathbb{H}}, T)_\sharp \mu$. For this reason, the set $\Lambda(\Phi, \Psi)$ as in Theorem 1.1.6 is a singleton (see also Remark 7.4.2) and thus

$$\langle V - W, X - Y \rangle = \left[(X, V)_\sharp \mathbb{P}, (Y, W)_\sharp \mathbb{P}\right]_r = [\Phi, \Psi]_r \leqslant 0$$

for every $X, Y, V, W \in \mathcal{H}$, such that $(X, V)_\sharp \mathbb{P} = \Phi$ and $(Y, W)_\sharp \mathbb{P} = \Psi$, where, again, $\mathcal{H}$ is any Hilbert space of random variables parametrizing $\mathcal{P}_2(\mathbb{H})$.

Let us then address the second point; to lift the structure of $\mathcal{P}_2(\mathbb{H})$ to a Hilbert space, it is enough to consider any standard Borel probability space $(\Omega, \mathcal{B}, \mathbb{P})$: this means that $\mathcal{B}$ is a sigma algebra in Ω such that there exists a Polish topology τ on Ω for which $\mathcal{B}$ is the Borel sigma algebra generated by τ and $\mathbb{P}$ is just a diffuse probability measure on $(\Omega, \mathcal{B})$, i.e. $\mathbb{P}(\{\omega\}) = 0$ for every $\omega \in \Omega$. It is a classical result, see also Corollary 6.2.3, that, for any choice of standard Borel probability space $(\Omega, \mathcal{B}, \mathbb{P})$, the L^2 space of random variables in $(\Omega, \mathcal{B}, \mathbb{P})$ with values in $\mathbb{H}$ parametrizes $\mathcal{P}_2(\mathbb{H})$, in the sense that for every $\mu \in \mathcal{P}_2(\mathbb{H})$ there exists $X \in L^2((\Omega, \mathcal{B}, \mathbb{P}); \mathbb{H})$ such that $X_\sharp \mathbb{P} = \mu$. For the sake of clarity, in this introduction we consider only the case in which $\Omega = [0, 1)$, $\mathcal{B}$ is the Borel sigma algebra in $[0, 1)$ and $\mathbb{P}$ is the 1-dimensional Lebesgue measure restricted to $[0, 1)$, while in the thesis we deal with a general standard Borel probability space[2]. We thus define

$$\mathcal{H} := L^2(([0, 1), \mathcal{B}, \mathbb{P}); \mathbb{H}).$$

In particular, in order to parametrize discrete measures, we define the finite sigma algebras $\mathcal{B}_n$ generated by the partitions[3]

$$\mathfrak{P}_n := \{[1, 1/2^n), [1/2^n, 2/2^n), \dots, [(2^n - 1)/2^n, 1)\} \quad \text{for } n \in \mathbb{N}$$

and the sub-Hilbert spaces of piecewise constant random variables

$$\mathcal{H}_n := L^2(([0, 1), \mathcal{B}_n, \mathbb{P}); \mathbb{H}),$$

together with their subsets $\mathcal{O}_n$ of injective maps. In this way, elements of $\mathcal{H}_n$ parametrize discrete measure concentrated on 2^n (non necessarily distinct) points, while elements of $\mathcal{O}_n$ parametrize discrete measure concentrated on exactly 2^n distinct points.

2 The example of standard Borel probability space used here is, in a way, canonical, see Section 6.2 where this and other related matters are addressed in detail.

3 Also in this case, one can consider more general refining families of partitions. These is indeed used in the thesis, but we prefer to consider only this example in this explanatory part.

As outlined in the third point, we can start from the set of discrete measures C and define $F_n \subset \mathcal{H}_n \times \mathcal{H}_n$ as

$$F_n := \left\{ (X, \Pi_n V) \in (\mathcal{O}_n \cap \mathcal{D}_n) \times \mathcal{H}_n : (X, V)_\sharp \mathbb{P} \in \mathbf{F} \right\} \subset \mathcal{H}_n \times \mathcal{H}_n,$$

where $\mathcal{D}_n \subset \mathcal{H}_n$ is the set of random variables $X \in \mathcal{H}_n$ such that $X_\sharp \mathbb{P} \in C$ and $\Pi_n : \mathcal{H} \to \mathcal{H}_n$ is the orthogonal projection. The first compatibility result is the following (see Proposition 9.4.4 and Lemma 9.4.5).

Theorem 1.1.9. *Let* $\mathbf{F}$ *be a dissipative MPVF, let* $n \in \mathbb{N}$ *and suppose that* $\mathbf{F}$ *admits a core* $C \subset D(\mathbf{F})$ *which is in addition convex along couplings (in the sense that, if* $\mu_0, \mu_1 \in C$*, then* $((1-t)x_0 + x_1)_\sharp \boldsymbol{\mu} \in C$ *for every* $t \in [0,1]$ *and every* $\boldsymbol{\mu} \in \Gamma(\mu_0, \mu_1)$*) and open in* $D_f(\mathbf{F})$ *with respect to the* W_∞*-topology. Then the maximal dissipative extension* $\hat{F}_n$ *of* F_n *in* $\mathcal{H}_n \times \mathcal{H}_n$ *is compatible with* $\mathbf{F}$ *in the sense that*

$$\langle V, X - Y \rangle + \left[\Psi, X_\sharp \mathbb{P} \right]_r \leqslant 0 \tag{1.1.45}$$

for every $(X, V) \in \hat{F}_n$, $Y \in D(\hat{F}_n)$ *and every* $\Psi \in \mathbf{F}[Y_\sharp \mathbb{P}]$.

The proof of this result is based on a few technical tools developed in Sections 6.4 and 6.5; in particular, in Theorem 6.5.2, we prove that curves induced by couplings between discrete measures are piece-wise geodesic and, in Proposition 6.4.3, we show that the map $x^t : X^2 \to X$ sending $(x_0, x_1) \mapsto (1-t)x_0 + tx_1$ is *almost injective* on the support of discrete couplings, a crucial property in the computation of the pseudo scalar product $[\cdot, \cdot]_r$ (see Remark 7.4.2).

Finally, the maximal dissipative set $F \subset \mathcal{H} \times \mathcal{H}$ is defined as the maximal dissipative extension in $\mathcal{H} \times \mathcal{H}$ of $F_\infty := \cap_n \hat{F}_n$, which is proven to be non-empty and dissipative. Some of the properties of F are summarized in the following statement (see in particular Theorem 9.3.3 and Proposition 9.4.9).

Theorem 1.1.10. *Let* $\mathbf{F}$ *and* C *be as in Theorem 1.1.9. Then* F *is law invariant in the following sense: if* $X_0 \in D(F)$ *and* $Y_0 \in \mathcal{H}$ *is such that* $(X_0)_\sharp \mathbb{P} = (Y_0)_\sharp \mathbb{P}$*, then* $Y_0 \in D(F)$ *and*

$$(X_0, J_\tau X_0, S_t X_0)_\sharp \mathbb{P} = (Y_0, J_\tau Y_0, S_t Y_0)_\sharp \mathbb{P} \quad \text{for every } \tau > 0,\ t \geqslant 0, \tag{1.1.46}$$

where J_τ *and* S_t *are the the resolvent and the semigroup operators associated to* F*, respectively. Moreover, if* $Y \in D(F_\infty)$ *and* $\Psi \in \mathbf{F}[Y_\sharp \mathbb{P}]$*, we have*

$$\langle V, X - Y \rangle + \left[\Psi, X_\sharp \mathbb{P} \right]_r \leqslant 0. \tag{1.1.47}$$

Finally, to address the fourth point, notice that the result above alredy goes in the correct direction: indeed, the fact that the semigroup generated by F depends only on the law of the starting point and the compatibilty condition (1.1.47) make reasonable to hope that the curve $t \mapsto (S_t X_0)_\sharp \mathbb{P}$ could be an EVI solution for $\mathbf{F}$ starting from $\mu_0 := (X_0)_\sharp \mathbb{P}$. To state the last result we need to assume a few additional hypotheses on the core C and on the geometry of the domain of $\mathbf{F}$. In particular we consider the following approximability condition for the core C

$$\begin{gathered}\text{for every } \mu \in \mathrm{D}(\mathbf{F}) \text{ there exists } \mu_n \in \mathrm{C} \text{ and } \Phi_n \in \mathbf{F}[\mu_n] \text{ such that}\\ W_2(\mu_n, \mu) \to 0, \quad \sup_n |\Phi_n|_2 < +\infty\end{gathered} \tag{1.1.48}$$

and, for every $\mu' \in \overline{\mathrm{D}(\mathbf{F})}$, we define the set

$$\mathcal{S}_{\mathbf{F}}(\mu') := \left\{ \mu'' \in \mathrm{D}_f(\mathbf{F}) \;\middle|\; \begin{array}{l} \mu_t \in \mathrm{D}(\mathbf{F}) \text{ for every } t \in (0,1] \\ \text{and every } (\mu_t)_{t\in[0,1]} \in \mathrm{G}(\mu', \mu'') \end{array} \right\},$$

where $\mathrm{G}(\mu', \mu'')$ is the set of *generalized geodesics* connecting μ' to μ'' (for the precise definition see [5, Definition 9.2.2]). The result is the following (see Theorem 9.4.18).

Theorem 1.1.11. *Let* $\mathbf{F}$ *and* C *be as in Theoerem 1.1.9 and suppose that, in addition,* C *satisfies* (1.1.48) *Assume that for every* $\mu' \in \overline{\mathrm{D}(\mathbf{F})}$ *the set* $\mathcal{S}_{\mathbf{F}}(\mu')$ *is non-empty and open in* $\mathrm{D}_f(\mathbf{F})$ *with respect to the* W_∞*-topology. Let* $\mu_0 \in \mathrm{D}(\mathbf{F})$*, let* $(\mu_t)_{t\geqslant 0}$ *be the Lipschitz curve defined by* $\mu_t := S_t\mu_0$ *for every* $t \in [0, +\infty)$ *and let* $(\nu_t)_{t\geqslant 0}$ *be a locally absolutely continuous EVI solution for* $\mathbf{F}$ *starting from* ν_0*. Then*

$$W_2(\mu_t, \nu_t) \leqslant W_2(\mu_0, \nu_0) \quad \textit{for every } t \in [0, +\infty). \tag{1.1.49}$$

In particular $(\mu_t)_{t\geqslant 0}$ *is the unique locally absolutely continuous* 0*-EVI solution for* $\mathbf{F}$ *starting from* μ_0*.*

The second part is organized as follows: Chapter 6 collects some preliminary material that is used in the subsequent chapters; in particular there is a synthesis of the theory of dissipative evolutions in Hilbert spaces, a fine treatment of Borel partitions and parametrization of measures by random variables, together with a few technical tools related to discrete measures and weak topologies; Chapter 7 contains a throughout study of the notion of dissipativity in Wasserstein spaces, the treatment of pseudo scalar products and of the interaction between dissipative operators in Wasserstein spaces and geodesics; Chapter 8 treats the notion of EVI solutions and its connection with the barycentric property; finally Chapter 9 introduces the explicit Euler scheme, presents its convergence results and treats the construction of the Hilbertian dissipative operator F, besides presenting a few results related to general law invariant dissipative operators.

Part II is the result of a collaboration with Giulia Cavagnari and Giuseppe Savaré, and part of the material presented in Section 6.3 and Chapters 7, 8, 9 appeared in [34].

1.1.3 *Part III: Wasserstein-Sobolev spaces*

Before describing the results presented in this third part of the thesis, we briefly describe the framework of metric Sobolev spaces we deal with. To this aim, let us fix a complete and separable metric space (X, d), a Borel positive measure $\mathfrak{m}$ on X (the triplet $(X, d, \mathfrak{m})$ is called a Polish metric measure space) and an exponent $p \in [1, +\infty)$. We denote by $L^0(X, \mathfrak{m})$ the space of equivalence classes of Borel measurable functions $f : X \to \mathbb{R}$ identified up to $\mathfrak{m}$ measure. This space is naturally endowed with the topology of the convergence in $\mathfrak{m}$ measure.

The first approach to Sobolev functions in metric measures spaces we present, strictly related to the ideas of Cheeger [35](see also [61, 70]), is contained in the work of Ambrosio, Gigli and Savaré [7] where they define the following concept of p-relaxed gradient: if $f \in \mathrm{Lip}_b(X, d)$ the asymptotic Lipschitz constant of f is defined as

$$\mathrm{lip}_d f(x) := \lim_{r \downarrow 0} \mathrm{Lip}(f, B(x, r), d) = \limsup_{y,z \to x,\ y \neq z} \frac{|f(y) - f(z)|}{d(y, z)}, \tag{1.1.50}$$

where $B(x, r)$ denotes the open ball centered at x with radious r and, for $A \subset X$, the quantity $\mathrm{Lip}(f, A, d)$ is defined as

$$\mathrm{Lip}(f, A, d) := \sup_{x,y \in A,\ x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}.$$

We say that $G \in L^p(X, \mathfrak{m})$ is a p-relaxed gradient of $f \in L^0(X, \mathfrak{m})$ if there exist a sequence $(f_n)_n \subset \mathrm{Lip}_b(X, d)$ and $\tilde{G} \in L^p(X, \mathfrak{m})$ such that

1. $f_n \to f$ in $L^0(X, \mathfrak{m})$ and $\mathrm{lip}_d f \rightharpoonup \tilde{G}$ in $L^p(X, \mathfrak{m})$,
2. $\tilde{G} \leqslant G$ $\mathfrak{m}$-a.e. in X.

It is not difficult to check that the set of p-relaxed gradients of $f \in L^0(X, \mathfrak{m})$ is convex and weakly closed, so that it admits an element of minimal $L^p(X, \mathfrak{m})$ norm (which turns out also to be minimal also in the a.e. sense), denoted by $|Df|_\star \in L^p(X, \mathfrak{m})$, and called minimal p-relaxed gradient of f. The p-Cheeger energy of $f \in L^0(X, \mathfrak{m})$ is then defined as

$$\mathsf{CE}_p(f) := \int_X |Df|_\star^p \, d\mathfrak{m}$$

and can be proved to be the relaxation of the so called pre-p-Cheeger energy

$$\mathsf{pCE}_p(f) := \int_X (\mathrm{lip}_d f)^p \, d\mathfrak{m}, \quad f \in \mathrm{Lip}_b(X, d),$$

in the sense that

$$\mathsf{CE}_p(f) = \inf \left\{ \liminf_{n \to +\infty} \mathsf{pCE}_p(f_n) : (f_n)_n \subset \mathrm{Lip}_b(X, d),\ f_n \to f \text{ in } L^0(X, \mathfrak{m}) \right\}. \tag{1.1.51}$$

The Sobolev space *à la Cheeger* $H^{1,p}(X, d, \mathfrak{m})$ is thus the vector space of functions $f \in L^p(X, \mathfrak{m})$ with finite Cheeger energy endowed with the norm

$$|f|^p_{H^{1,p}(X,d,\mathfrak{m})} := \int_X |f|^p \, d\mathfrak{m} + \mathsf{CE}_p(f)$$

which makes it a Banach space. Before moving to the next approach, let us mention the strong approximation property which states that Lipschitz functions are dense in $H^{1,p}(X, d, \mathfrak{m})$ (see also point (2) in Theorem 10.1.2):

$$\text{for every } f \in H^{1,p}(X, d, \mathfrak{m}) \text{ there exists a sequence } (f_n)_n \subset \mathrm{Lip}_b(X, d) \text{ such that}$$
$$f_n \to f, \quad \mathrm{lip}_d f_n \to |Df|_\star \quad \text{in } L^p(X, \mathfrak{m}) \text{ as } n \to +\infty. \tag{1.1.52}$$

Another approach to metric Sobolev spaces is due to Shanmugalingam [111] and it is based on the concept of p-modulus of a family of curves. Let us denote by $\Gamma(X)$ the set of absolutely continuous curves defined in some non-degenerate interval $\mathcal{I} \subset \mathbb{R}$ with values in X. Given $\gamma \in \Gamma(X)$, we denote by $I(\gamma)$ the interval where the curve is defined and by γ_I and γ_F the evaluations of γ at the infimum and supremum of $I(\gamma)$, respectively. If $G : X \to [0, +\infty]$ is a Borel function and $\gamma \in \Gamma(X)$, we set

$$\int_\gamma G := \int_{I(\gamma)} G(\gamma_t))|\dot{\gamma}_t|_d \, dt,$$

where $|\dot{\gamma}_t|_d$ is the metric derivative of γ at $t \in I(\gamma)$ defined as

$$|\dot{\gamma}_t|_d := \limsup_{h \to 0} \frac{d(\gamma_{t+h}, \gamma_t)}{|h|}.$$

If we consider a family of curves $\Gamma \subset \Gamma(X)$, its p-modulus is defined as

$$\mathrm{Mod}_p(\Gamma) := \inf \left\{ \int_X \rho^p \, d\mathfrak{m} \mid \rho : X \to [0, +\infty] \text{ Borel}, \int_\gamma \rho \geqslant 1 \text{ for every } \gamma \in \Gamma \right\}.$$

We say that a property depending on $\gamma \in \Gamma(X)$ holds p-a.e. provided it is satisfied for every γ belonging to some set $\Gamma \subset \Gamma(X)$ with $\mathrm{Mod}_p(\Gamma^c) = 0$. Finally, given Borel functions (not equivalent classes) $f : X \to \mathbb{R}$ and $G : X \to [0, +\infty]$, we say that G is a p-weak upper gradient for f if

$$|f(\gamma_F) - f(\gamma_I)| \leqslant \int_\gamma G \quad \text{for p-a.e. } \gamma.$$

Also in this case, given $f \in L^0(X, \mathfrak{m})$ it is possible to prove that it admits a p-weak upper gradient $|Df|_{Sh} \in L^p(X, \mathfrak{m})$ which is minimal in the following sense: if $\tilde{f} : X \to \mathbb{R}$ is a Borel representative of f and G is a p-weak upper gradient of $\tilde{f}$, then $|Df|_{Sh} \leqslant G$ a.e. in X. Finally the so called Newtonian space $N^{1,p}(X, d, \mathfrak{m})$ is the vector space of functions $f \in L^p(X, \mathfrak{m})$ which admits a p-weak upper gradient endowed with the norm

$$|f|^p_{N^{1,p}(X,d,\mathfrak{m})} := \int_X |f|^p \, d\mathfrak{m} + \int_X |Df|^p_{Sh} \, d\mathfrak{m}$$

which makes it a Banach space.
The remarkable result, which can be found in [7, Theorem 6.2] [6, Theorem 7.4], is that the two approaches are in fact equivalent, meaning that the spaces

$H^{1,p}(X, d, m)$ and $N^{1,p}(X, d, m)$ coincide and the minimal p-gradients are equal, i.e.

$$|Df|_\star = |Df|_{Sh} \quad m\text{-a.e. in } X \text{ and for every } f \in H^{1,p}(X, d, m). \tag{1.1.53}$$

Contributions in Part III. Here we will adopt the Cheeger approach (even if the equivalence of (1.1.53) will be useful, see in particular Section 10.2) with a slight difference [108]: instead on considering the whole space of Lipschitz and bounded functions on X, we fix a unital subalgebra $\mathscr{A} \subset \mathrm{Lip}_b(X, d)$ that separates the points in X and we work with the concept of $(p, \mathscr{A})$-relaxed gradient (and thus with the one of minimal $(p, \mathscr{A})$-relaxed gradient and Cheeger energy $\mathsf{CE}_{p,\mathscr{A}}$) simply substituting $\mathrm{Lip}_b(X, d)$ with $\mathscr{A}$ in the corresponding constructions above. This leads of course to a different Sobolev space $H^{1,p}(X, d, m; \mathscr{A})$ with the norm

$$|f|_{H^{1,p}(X,d,m;\mathscr{A})} := \int_X |f|^p \, dm + \mathsf{CE}_{p,\mathscr{A}}(f) = \int_X |f|^p \, dm + \int_X |Df|^p_{\star,\mathscr{A}} \, dm.$$

Of course every $(p, \mathscr{A})$-relaxed gradient is also a p-relaxed gradient and it holds $|Df|_{\star,\mathscr{A}} \geqslant |Df|_\star$ m-a.e. in X for every $f \in L^0(X, m)$ with a $(p, \mathscr{A})$-relaxed gradient, so that the Sobolev space $H^{1,p}(X, d, m; \mathscr{A})$ is contained in $H^{1,p}(X, d, m)$. The advantage in considering the subalgebra $\mathscr{A}$ consists in the fact that its elements may be more regular and/or computations within the subalgebra may be easier than in the general case.

It is thus relevant to determine sufficient conditions for the Sobolev space $H^{1,2}(X, d, m; \mathscr{A})$ to coincide with the standard $H^{1,2}(X, d, m)$; this would lead also to the strong approximation property (1.1.52) in terms of the algebra $\mathscr{A}$:

$$\begin{gathered}\text{for every } f \in H^{1,p}(X, d, m) \text{ there exists a sequence } (f_n)_n \subset \mathscr{A} \text{ such that}\\ f_n \to f, \quad \mathrm{lip}_d f_n \to |Df|_\star \quad \text{in } L^p(X, m) \text{ as } n \to +\infty.\end{gathered} \tag{1.1.54}$$

It is not difficult to see that (1.1.54) implies that

$$|Dd_y|_{\star,\mathscr{A}} \leqslant 1 \quad m\text{-a.e. in } X \text{ and for every } y \in X, \tag{1.1.55}$$

where $d_y(x) := d(x, y)$ for $x \in X$. The first result (see Theorem 10.2.1) is to prove that actually this condition is also sufficient to get (1.1.54).

Theorem 1.1.12. *Let (X, d, m) be a Polish metric measure space and let $\mathscr{A}$ be a unital separating subalgebra of $\mathrm{Lip}_b(X)$. If (1.1.55) holds true, then for every $f \in L^p(X, m)$ with a p-relaxed gradient there exists a sequence $(f_n)_n \subset \mathscr{A}$ satisfying (1.1.54) i.e. f has a $(p, \mathscr{A})$-relaxed gradient and*

$$|Df|_{\star,\mathscr{A}} = |Df|_\star \quad m\text{-a.e. in } X. \tag{1.1.56}$$

The proof of this results is obtained employing the regularizing properties of the Hopf-Lax semigroup

$$Q_t(f)(x) := \inf_{y \in X} \left\{ \frac{1}{2t} d^2(x, y) + f(y) \right\}, \quad x \in X$$

and a similar argument (although in a different setting) was already contained in [108] (see in particular [108, Theorem 3.2.7]).

In case $p = 2$, the subalgebra viewpoint presents another advantage: if the pre-Cheeger energy pCE_2 satisfies the parallelogram identity at the level of $\mathscr{A}$, then the resulting space $H^{1,2}(X, d, \mathfrak{m}; \mathscr{A})$ (and thus $H^{1,2}(X, d, \mathfrak{m})$ if equality has been proven) is an Hilbert space. More precisely we have the following result (see Theorem 10.2.4).

Theorem 1.1.13 (An Hilbertianity condition). *Let $p = 2$ and let $\mathscr{A}$ be a separating unital subalgebra of $\mathrm{Lip}_b(X)$ satisfying (1.1.55). If for every $f, g \in \mathscr{A}$*

$$\int_X \left(|\mathrm{lip}(f+g)|^2 + |\mathrm{lip}(f-g)|^2 \right) d\mathfrak{m} = 2 \int_X \left(|\mathrm{lip} f|^2 + |\mathrm{lip} g|^2 \right) d\mathfrak{m}, \tag{1.1.57}$$

then $H^{1,2}(X, d, \mathfrak{m})$ is an Hilbert space, CE_2 is a quadratic form, and $\mathscr{A}$ is strongly dense.

The Wasserstein-Sobolev space. As a remarkable application of Theorem 1.1.13 we consider the case of the Sobolev space on the 2-Wasserstein space on $\mathbb{R}^d$. The metric space is thus $\mathcal{P}_2(\mathbb{R}^d)$, the space of probability measures on $\mathbb{R}^d$ with finite second moment, with the Wasserstein distance W_2; we fix then any Borel positive measure $\mathfrak{m}$ on $\mathcal{P}_2(\mathbb{R}^d)$.

The unital subalgebra of functions $\mathscr{A} \subset \mathrm{Lip}_b(\mathcal{P}_2(\mathbb{R}^d), W_2)$ we consider is the one of cylindrical functions (which has nothing to do with the cylindrical functions in $\mathrm{Cyl}(\mathbb{H})$ of Definition 2.4.5, and was already considered by Dello Schiavo [44]): every $\phi \in C^1_b(\mathbb{R}^d)$ induces the function L_ϕ on $\mathcal{P}(\mathbb{R}^d)$

$$L_\phi : \mu \to \int_{\mathbb{R}^d} \phi \, d\mu \tag{1.1.58}$$

which clearly belongs to $\mathrm{Lip}_b(\mathcal{P}_2(\mathbb{R}^d), W_2)$. More generally, if we consider a vector $\boldsymbol{\phi} = (\phi_1, \cdots, \phi_N) \in \left(C^1_b(\mathbb{R}^d)\right)^N$, we denote by $L_{\boldsymbol{\phi}} := (L_{\phi_1}, \cdots, L_{\phi_N})$ the corresponding map from $\mathcal{P}_2(\mathbb{R}^d)$ to $\mathbb{R}^N$. The algebra $\mathscr{A}$ of C^1 cylindrical functions is made of those functions $F : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ of the form

$$F(\mu) = \psi(L_{\boldsymbol{\phi}}(\mu)) \quad \mu \in \mathcal{P}_2(\mathbb{R}^d), \tag{1.1.59}$$

where $\boldsymbol{\phi} \in \left(C^1_b(\mathbb{R}^d)\right)^N$ and $\psi \in C^1_b(\mathbb{R}^N)$ for some $N \in \mathbb{N}$.

Every element $F \in \mathscr{A}$ as in (1.1.59) comes with a natural notion of vector valued gradient $DF : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ defined as

$$DF(\mu, x) := \sum_{i=1}^N \partial_i \psi(L_{\boldsymbol{\phi}}(\mu)) \nabla \phi_i(x), \quad (\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d. \tag{1.1.60}$$

Since the representation of a cylindrical function $F \in \mathscr{A}$ is never unique, in principle the gradient DF of F may depend on the particular choice of ψ and $\boldsymbol{\phi}$ used to represent F; in Proposition 11.1.10 we show that up to integrating DF w.r.t. μ, we obtain a quantity that depends solely on F i.e.

$$\int_{\mathbb{R}^d} |DF(\mu, x)|^2 \, d\mu(x) = (\mathrm{lip} F(\mu))^2 \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{1.1.61}$$

In particular the pre-Cheeger energy satisfies the parallelogram identity (1.1.57) when restricted to the algebra $\mathscr{A}$.

It is then very important to prove that in the present Wasserstein setting the inequality (1.1.55) holds. Let us present a formal motivation for the proof of such result. Let us fix an absolutely continuous probability measure $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ and let us consider the continuous function

$$F_\nu(\mu) := \frac{1}{2} W_2^2(\mu, \nu), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{1.1.62}$$

Let $(\varphi_{1,\mu}, \varphi_{2,\mu})$ be a pair of Kantorovich potentials for ν and μ in the sense that

$$\int_{\mathbb{R}^d} \varphi_{1,\mu} \, d\mu + \int_{\mathbb{R}^d} \varphi_{2,\mu} \, d\nu = \frac{1}{2} \int_{\mathbb{R}^d} |y|^2 \, d\nu(y) + \frac{1}{2} \int_{\mathbb{R}^d} |x|^2 \, d\mu(x) - \frac{1}{2} W_2^2(\nu, \mu)$$

so that

$$F_\nu(\mu) = \int_{\mathbb{R}^d} u_\mu(x) \, d\mu(x), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d),$$

where

$$u_\mu(x) := \frac{1}{2}|x|^2 - \varphi_{1,\mu}(x) + a_\mu, \quad a_\mu := \int_{\mathbb{R}^d} \left(\frac{1}{2}|y|^2 - \varphi_{2,\mu}(y) \right) d\nu(y).$$

If u_μ depends in a continuous way on μ and it can be approximated by elements of $\mathscr{A}$, we can expect that, by (1.1.61), we would have

$$|DF_\nu|^2_{\star,\mathscr{A}}(\mu) \lessapprox (\operatorname{lip} F_\nu(\mu))^2 \approx \int_{\mathbb{R}^d} |\nabla u_\mu(x)|^2 \, d\mu(x) \tag{1.1.63}$$

$$= \int_{\mathbb{R}^d} |x - \nabla \varphi_{1,\mu}(x)|^2 \, d\mu(x) = W_2^2(\mu, \nu), \tag{1.1.64}$$

since $\nabla \varphi_{1,\mu}$ is the Optimal Transport map from μ to ν. Of course the assumption that u_μ depends in a continuous way on μ and can be approximated by elements of $\mathscr{A}$ is not met in general. In the proof of Theorem 11.1.19 and Corollary 11.1.20 we overcome these technical difficulties by a more refined variational argument and we obtain the following result.

Theorem 1.1.14. $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ *is a Hilbert space and the algebra* $\mathscr{A}$ *of cylindrical functions is dense in energy: for every* $F \in H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ *there exists a sequence* $F_n \in \mathscr{A}$, *such that*

$$F_n \to F, \quad \operatorname{lip}(F_n) \to |DF|_\star \textit{ in } L^2(X, \mathfrak{m}). \tag{1.1.65}$$

Another advantage in using the algebra $\mathscr{A}$ of cylindrical functions is a consequence of the fact, as already highlighted, that these functions comes already with a notion of gradient as in (1.1.60), and not only of norm of the gradient as in the approaches to the Sobolev theory in metric spaces mentioned above. Using the density in (1.1.65) we hence obtain a notion of gradient for Sobolev functions as a relaxation of the differential DF for cylindrical functions: if we denote by $\mathfrak{m}$ the measure in $\mathcal{P}(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d)$ defined as

$$\mathfrak{m} := \int_{\mathcal{P}_2(\mathbb{R}^d)} \delta_\mu \otimes \mu \, \mathrm{d}\mathfrak{m}(\mu),$$

there is a linear continuous Wasserstein-gradient operator

$$D_{\mathfrak{m}} : H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m}) \to L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$$

representing the bilinear form associated to the Cheeger energy as

$$\mathsf{CE}_2(F, G) = \int D_{\mathfrak{m}} F(\mu, x) \cdot D_{\mathfrak{m}} G(\mu, x) \, \mathrm{d}\mathfrak{m}(\mu, x) \tag{1.1.66}$$

satisfying useful calculus rules which are typical of Γ-calculus for Dirichlet form and allow for an explicit characterization of the tangent bundle $L^2(T\mathcal{P}_2(\mathbb{R}^d))$ in the sense of Gigli [55, 57].

Finally, we remark that it is not difficult to extend the theory developed so far to $(\mathcal{P}_2(X), W_2, \mathfrak{m})$, where X is a Riemmanian manifold (using the Nash embedding Theorem) or a (possibly infinitely dimensional) separable Hilbert space (using projections on finite dimensional subspaces whose union is dense in the space).

The third part consists of two Chapters: Chapter 10 contains a few preliminaries on the construction of Sobolev spaces with the adaptation of the Cheeger approach to the presence of the subalgebra $\mathscr{A}$ and the proof of the general criterium for the density in energy of such subalgebra; Chapter 11 presents the applications of the results of the previous Chapter to Wasserstein-Sobolev spaces: in particular, Section 10.2 presents the general setting and the proof of Theorem 1.1.14 while the remaining sections are devoted to the calculus rules for the Wasserstein gradient and the extension of the results to separable Hilbert spaces and Riemmanian manifolds.

Part III is the result of a collaboration with Massimo Fornasier and Giuseppe Savaré.

2 PRELIMINARIES

In this chapter we discuss the main notation used in the thesis and we describe the framework in which it is set. Most of the material is basic but we prefer to state the definitions clearly in order to avoid ambiguities.

2.1 TOPOLOGICAL AND MEASURE-THEORETIC FRAMEWORK

Given a set X, we denote by 2^X the set of subset of X. A topological space is a pair (X, τ) where X is a non-empty set and $\tau \subset 2^X$ is a collection of subsets of X containing $\emptyset$ and X and which is closed under finite intersections and arbitrary unions. We say that τ is a topology on X and its elements are called open sets. Complements of open sets are called closed sets. When the topology on a set X is understood, we simply say that X is a topological space. A neighbourhood of a point $x \in X$ in the topological space (X, τ) is an open set containing x.

Given two topological spaces (X, τ), (Y, σ), we denote the space of continuous functions from X to Y as $C((X, \tau); (Y, \sigma))$ (or simply with $C(X; Y)$) when the topologies are clear from the context). If $Y = \mathbb{R}$, we write $C(X)$ ($C_b(X)$ if we are dealing with continuous and bounded functions).

An extended distance on a non-empty set X is a map $d : X \times X \to [0, +\infty]$ such that

- $d(x, y) = 0$ if and only if $x = y$,
- $d(x, y) = d(y, x)$ for every $x, y \in X$,
- $d(x, z) \leqslant d(x, y) + d(y, z)$ for every $x, y, z \in X$.

If, in addition, $d(x, y) < +\infty$ for every $x, y \in X$, we say that d is a distance on X. The pair (X, d) is then called an extended metric space (resp. a metric space). Every extended metric space (X, d) is endowed with a natural topology τ_d whose basis is given by the open balls

$$B(x, r) := \{y \in X \mid d(y, x) < r\}, \quad x \in X, \ r \in (0, +\infty).$$

We will deal with some different kinds of topological spaces which enjoy specific properties. A topological space (X, τ) is said to be

- Hausdorff, if for every pair of distinct points x, y there exist disjoint neighbourhoods of x and y,
- Completely regular, if it is Hausdorff and for every closed set C and point $x \in X \setminus C$ there exists a function $f \in C_b(X)$ such that $f(x) = 0$ and $f(C) = \{1\}$,
- Polish, if it is homeomorphic to a complete and separable metric space,

- Lusin, if it Hausdorff and it is the image of a Polish space through a continuous and injective function,
- Suslin, if it is Hausdorff and it is the image of a Polish space through a continuous function.

Every (extended) metric space is completely regular, Polish spaces are Lusin and Lusin spaces are Suslin. These inclusions can be proven to be strict. For a complete account of the theory of Lusin/Suslin spaces we refer to [20, Chapter 6].

A set $\mathbb{L}$ is said to be directed w.r.t. a partial order relation $\prec$ if for every pair of elements $i, j \in \mathbb{L}$ there exists $h \in \mathbb{L}$ such that $i \prec h$ and $j \prec h$. A net in a topological space (X, τ) is a map from a directed set $\mathbb{A}$ to X. A subnet of a net $f : \mathbb{L} \to X$ is the composition of f with a final monotone function $\mathbb{A} \to \mathbb{L}$, where $\mathbb{A}$ is some directed set. We denote nets miming the notation for sequences, e.g. with $(x_i)_{i \in \mathbb{L}}$ and we say that a net $(x_i)_{i \in \mathbb{L}}$ converges to $x \in X$ if for every neighbourhood U of x there exists an index $j \in \mathbb{L}$ such that $x_i \in U$ for every $i \in \mathbb{L}$, $j \prec i$. We refer e.g. to [50] for a discussion about nets.

A measurable space is a pair $(X, \mathcal{F})$ where X is a non-empty set and $\mathcal{F} \subset 2^X$ is a sigma-algebra on X, meaning that $\mathcal{F}$ contains $\emptyset$ and X, it is closed under countable unions and complementation. Given $\mathcal{E} \subset 2^X$, we denote by $\sigma(\mathcal{E})$ the smallest sigma-algebra on X containing $\mathcal{E}$. In particular, when (X, τ) is a topological space and $\mathcal{E} = \tau$, the sigma algebra $\sigma(\tau)$ is denoted by $\mathcal{B}((X, \tau))$ (or $\mathcal{B}(X)$, if the topology is understood) and called Borel sigma algebra on (X, τ). Given two topological spaces (X, τ), (Y, σ) we say that a function $f : X \to Y$ is Borel (measurable) if it is measurable w.r.t. the Borel sigma algebras of X and Y. When the target space is $\mathbb{R}$, we denote by $B(X)$ (resp. $B_b(X)$) the space of Borel (resp. Borel and bounded) functions from X to $\mathbb{R}$.

Given a measurable space $(X, \mathcal{F})$, a (finite) measure on $(X, \mathcal{F})$ is a sigma-additive map $\mu : \mathcal{F} \to \mathbb{R}$. We say that μ is non-negative if $\mu(A) \geqslant 0$ for every $A \in \mathcal{F}$. The triplet $(X, \mathcal{F}, \mu)$ is called a measure space. We say that the measure μ is carried by (or it is concentrated on) $X' \subset X$ if there exists a set $X'' \in \mathcal{F}$ such that $\mu(X'') = 0$ and $X' \subset X''$. If $(X, \mathcal{F})$ and $(Y, \mathcal{G})$ are measurable spaces, μ is a measure on $(X, \mathcal{F})$ and $f : X \to Y$ is a measurable map, the push forward of the measure μ through f, denoted by $f_\sharp \mu$, is a measure on $(Y, \mathcal{G})$ defined by

$$f_\sharp \mu(A) := \mu\left(f^{-1}(A)\right), \quad A \in \mathcal{G}.$$

Given a topological space (X, τ), we say that μ is a

- Borel measure, if it is a measure on $(X, \mathcal{B}((X, \tau))$,
- Radon measure, if it is a Borel measure and for every $A \in \mathcal{B}((X, \tau))$ and every $\varepsilon > 0$ there exists a compact set $K \subset A$ such that $|\mu|(A \setminus K) < \varepsilon$, where $|\mu|$ is the total variation measure associated to μ.

Every Borel measure in a Suslin space is a Radon measure concentrated on a sigma-compact set (i.e. the countable union of compact sets), see e.g. [20, Theorem 7.4.3]. The support of a measure μ on a topological space (X, τ), denoted by

$\mathrm{supp}(\mu)$, is the smallest closed set on which μ is concentrated. We introduce the following notation for measures on a topological space (X, τ):

- $\mathcal{M}(X)$ is the set of Radon measures on X,
- $\mathcal{M}_+(X)$ is the set of non-negative Radon measures on X,
- $\mathcal{P}(X)$ is the set of probabilities on X, i.e. the elements $\mu \in \mathcal{M}_+(X)$ such that $\mu(X) = 1$,
- $\mathrm{Discr}(X)$ is the set of discrete measures on X i.e. the elements $\mu \in \mathcal{M}(X)$ such that $\mathrm{supp}(\mu)$ is a finite (possibly empty) set,
- $\mathrm{Discr}_+(X) := \mathrm{Discr}(X) \cap \mathcal{M}_+(X)$,
- $\Delta(X)$ is the set of measures $\mu \in \mathcal{M}_+(X)$ of the form $\mu = r\delta_x$ for some $x \in X$, $r \geqslant 0$,
- 0_X is the null measure on X.

Notice that $\mathrm{Discr}_+(X) \subset \mathrm{Discr}(X) \subset \mathcal{M}(X)$ and $\Delta(X) \subset \mathcal{M}(X)$. Unless otherwise stated, we will deal with Radon measures on topological spaces and we will just say "μ is a measure on X", meaning that μ is a Radon measure on the topological space X, where the topology is understood. Usually the topological space will be Polish (or Suslin) so that all Borel measures will be Radon measures. Only in few cases we will distinguish between Radon measures and Borel measures.

When dealing with a product space $X \times Y$ we will often use the notation π^X (resp. π^Y) or π^x (resp. π^y) to denote the projection on X (resp. on Y) i.e. the map sending (x, y) to x (resp. to y). In this case, if X and Y are topological spaces, $\mu \in \mathcal{M}(X)$ and $\nu \in \mathcal{M}(Y)$, we define

$$\Gamma(\mu, \nu) := \left\{ \boldsymbol{\gamma} \in \mathcal{M}(X \times Y) \mid \pi^X_\sharp \boldsymbol{\gamma} = \mu,\ \pi^Y_\sharp \boldsymbol{\gamma} = \nu \right\}. \tag{2.1.1}$$

If $\boldsymbol{\gamma} \in \Gamma(\mu, \nu)$, we say that μ and ν are the marginals of $\boldsymbol{\gamma}$ and that $\boldsymbol{\gamma}$ is a plan between μ and ν.
The following is the well known disintegration theorem.

Theorem 2.1.1. *Let $\mathbb{X}, X$ be Lusin completely regular topological spaces, let $\boldsymbol{\mu} \in \mathcal{P}(\mathbb{X})$ and let $r : \mathbb{X} \to X$ be a Borel map. Denote with $\mu = r_\sharp \boldsymbol{\mu} \in \mathcal{P}(X)$. Then there exists a μ-a.e. uniquely determined Borel family of probability measures $\{\boldsymbol{\mu}_x\}_{x \in X} \subset \mathcal{P}(\mathbb{X})$ such that $\boldsymbol{\mu}_x(\mathbb{X} \setminus r^{-1}(x)) = 0$ for μ-a.e. $x \in X$, and*

$$\int_{\mathbb{X}} \varphi(\boldsymbol{x})\, d\boldsymbol{\mu}(\boldsymbol{x}) = \int_X \left(\int_{r^{-1}(x)} \varphi(\boldsymbol{x})\, d\boldsymbol{\mu}_x(\boldsymbol{x}) \right) d\mu(x)$$

for every bounded Borel map $\varphi : \mathbb{X} \to \mathbb{R}$.

Remark 2.1.2. When $\mathbb{X} = X \times Y$ and $r = \pi^X$, we can canonically identify the disintegration $\{\boldsymbol{\mu}_x\}_{x \in X} \subset \mathcal{P}(\mathbb{X})$ of $\boldsymbol{\mu} \in \mathcal{P}(X \times Y)$ w.r.t. $\mu = \pi^X_\sharp \boldsymbol{\mu}$ with a family of probability measures $\{\mu_x\}_{x \in X} \subset \mathcal{P}(Y)$. We write $\boldsymbol{\mu} = \int_X \mu_x \, d\mu(x)$.

If X is a completely regular topological space, there is a canonical duality map $\langle\cdot,\cdot\rangle$ between $\mathcal{M}(X)$ and $C_b(X)$ given by

$$\langle\mu,\varphi\rangle := \int_X \varphi\, d\mu \quad \text{for every } \mu \in \mathcal{M}(X),\ \varphi \in C_b(X). \tag{2.1.2}$$

The map in (2.1.2) defines a real nondegenerate bilinear form in $\mathcal{M}(X) \times C_b(X)$, for if a Radon measure $\mu \in \mathcal{M}(X)$ satisfies $\int_X \varphi\, d\mu = 0$ for every $\varphi \in C_b(X)$, then $|\mu|(B) = 0$ for every $B \in \mathcal{B}(X)$ (e.g. by the approximation result [20, Lemma 7.2.8]) so that μ is the null measure. Hence we can endow $\mathcal{M}(X)$ with the narrow (sometimes also called weak) topology $\sigma(\mathcal{M}(X), C_b(X))$: the coarsest topology on $\mathcal{M}(X)$ for which the maps $\mu \mapsto \int_X \varphi\, d\mu$ are continuous for every $\varphi \in C_b(X)$.
Notice that, in general, $\mathcal{M}(X)$ with the narrow topology is not first-countable while $\mathcal{P}(X)$ and $\mathcal{M}_+(X)$ are metrizable (resp. Polish) if and only if X is metrizable (resp. Polish) (see e.g. [20, Theorem 8.9.4]). Unless otherwise stated, we will endow $\mathcal{M}(X)$ (and its subsets $\mathcal{M}_+(X)$ and $\mathcal{P}(X)$) with the narrow topology. For a net $(\mu_\lambda)_{\lambda\in\mathbb{L}} \subset \mathcal{M}(X)$ (resp. $\mathcal{P}(X)$) and a point $\mu \in \mathcal{M}(X)$ (resp. $\mathcal{P}(X)$), we write $\mu_\lambda \to \mu$ in $\mathcal{M}(X)$ (resp. $\mathcal{P}(X)$) or $\lim_{\lambda\in\mathbb{L}} \mu_\lambda = \mu$ in $\mathcal{M}(X)$ (resp. $\mathcal{P}(X)$) to mean that μ_λ converges to μ in the narrow topology of $\mathcal{M}(X)$ (resp. $\mathcal{P}(X)$).
We list here some useful properties related to the narrow topology. see [43, 54, 58, 59 Chap. III] for the proofs of the last three claims.

Lemma 2.1.3. *Let X, Y be completely regular spaces.*

1. *If $f : X \to Y$ is continuous then the map $f_\sharp : \mathcal{M}(X) \to \mathcal{M}(Y)$ is continuous.*
2. *If $\varphi : X \to (-\infty, +\infty]$ is lower semicontinuous and bounded from below and $(\mu_\lambda)_{\lambda\in\mathbb{L}}$ is a net weakly converging to μ in $\mathcal{M}_+(X)$ then*
$$\liminf_{\lambda\in\mathbb{L}} \int_X \varphi\, d\mu_\lambda \geqslant \int_X \varphi\, d\mu.$$
3. *If $\iota : X \to Y$ is a topological embedding (i.e. a continuous map providing a homeomorphism between X and $\iota(X)$ with the topology induced by the inclusion in Y), then $\iota_\sharp : \mathcal{M}(X) \to \mathcal{M}(Y)$ is a topological embedding as well, with*
$$\iota_\sharp(\mathcal{M}(X)) = \mathcal{M}(\iota(X), Y) := \Big\{\mu \in \mathcal{M}(Y) : \mu \text{ is concentrated on } \iota(X)\Big\}.$$
4. *If X is compact then for every $M \geqslant 0$ the set $\Big\{\mu \in \mathcal{M}(X) : |\mu|(X) \leqslant M\Big\}$ is compact.*

We recall the celebrated Prokhorov theorem ([110, Appendix]) which characterizes compactness in the narrow topology.

Theorem 2.1.4 (Prokhorov)**.** *Let X be a completely regular topological space and let $\mathcal{F} \subset \mathcal{M}(X)$ be a equibounded (i.e. $\sup|\mu|(X) < +\infty$) and tight subset i.e.*

$$\textit{for all } \varepsilon > 0 \textit{ there exists } K_\varepsilon \subset X \textit{ compact s.t. } \sup_{\mu\in\mathcal{F}} |\mu|(X \setminus K_\varepsilon) < \varepsilon.$$

Then $\mathcal{F}$ is relatively compact in $\mathcal{M}(X)$ w.r.t. the narrow topology. If X is Polish, also the converse implication holds true: if a set $\mathcal{F} \subset \mathcal{M}(X)$ is relatively compact in the narrow topology, then it is equibounded and tight.

2.2 FUNCTIONS, CONVEXIFICATION, RELAXATION

We briefly fix some notation related to relaxation and convexification of functions. If X is a set and $f : \mathbb{R} \to (-\infty, +\infty]$ is a function we denote the effective domain of f as $D(f)$ which is defined as

$$D(f) := \{x \in X \mid f(x) < +\infty\}.$$

If X is also a topological space, we denote by Γf the lower semicontinuous relaxation of f i.e.

$$\Gamma f(x) := \inf\left\{\liminf_{\lambda} f(x_\lambda) \mid (x_\lambda)_{\lambda\in\mathbb{L}} \subset X,\ x_\lambda \to x\right\}, \quad x \in X; \tag{2.2.1}$$

We have that Γf is the largest lower semicontinuous function below f. If X is also a topological vector space and $A \subset X$, we denote by $\operatorname{co}(A)$ (resp. by $\overline{\operatorname{co}}(A)$) the convex (resp. closed and convex) envelope of A. If $g : A \to (-\infty, +\infty]$ is a function, we denote its convex envelope and its closed convex envelope by $\operatorname{co}(g)$ and $\overline{\operatorname{co}}(g)$, defined by

$$\operatorname{co}(g)(x) := \inf\left\{\sum_{i=1}^{n} \alpha_i g(x_i) \mid \{(x_i, \alpha_i)\}_{i=1}^{n} \subset S_n(x),\, n \in \mathbb{N}\right\}, \quad x \in \operatorname{co}(A),$$
$$\overline{\operatorname{co}}(g) := \Gamma \operatorname{co}(g), \quad x \in \overline{\operatorname{co}}(A),$$

where

$$S_n(x) := \left\{\{(x_i, \alpha_i)\}_{i=1}^{n} \mid (x_i, \alpha_i) \in X \times [0,1],\, \sum_{i=1}^{n} \alpha_i = 1,\, \sum_{i=1}^{n} \alpha_i x_i = x\right\}.$$

Clearly $\operatorname{co}(g)$ is the largest convex function below g and $\overline{\operatorname{co}}(g)$ is the largest lower semicontinuous and convex function below g.

2.3 THE OPTIMAL TRANSPORT PROBLEM

We set a few notation related to the Optimal Transport Problem, as presented in the introduction; given two Polish spaces X, Y, probabilities $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$ and a Borel cost function $c : X \times Y \to [0, +\infty]$, the Kantorovich version of the Optimal Transport problem asks to find

$$\inf\left\{\int_{X\times Y} c(x,y)\, d\boldsymbol{\gamma}(x,y) \mid \boldsymbol{\gamma} \in \Gamma(\mu,\nu)\right\}. \tag{2.3.1}$$

If the cost is proper and lower semicontinuous the infumum above is attained in a non-empty, compact and convex subset of $\mathcal{P}(X \times Y)$ denoted by $\Gamma_o^c(\mu, \nu)$, the set of c-optimal plans between μ and ν. On the other hand, the Monge version of the Optimal Transport problem asks to find

$$\inf\left\{\int_{X} c(x, T(x))\, d\mu(x) \mid T \in \mathrm{Tr}(\mu,\nu)\right\}, \tag{2.3.2}$$

where $\mathrm{Tr}(\mu,\nu)$ is the set of admissible transport maps from μ to ν defined as

$$\mathrm{Tr}(\mu,\nu) := \left\{ T \in B(X;Y) \mid T_\sharp \mu = \nu \right\}.$$

It is clear that to every map $T \in \mathrm{Tr}(\mu,\nu)$ it can be associated a plan $\boldsymbol{\gamma}_T \in \Gamma(\mu,\nu)$ defined as $\boldsymbol{\gamma}_T := (\boldsymbol{i}_X \times T)_\sharp \mu$ satisfying

$$\int_{X\times Y} c \, d\boldsymbol{\gamma}_T = \int_X c(x,T)(x,y)\, d\mu(x)$$

so that the infimum in (2.3.1) is less than the one in (2.3.2). In case the cost function is continuous and μ is non-atomic (i.e. $\mu(\{x\}) = 0$ for every $x \in X$) it can be proven that the two infima coincide ([97]) but, even in this case, the infimum in (2.3.2) may not be attained.

2.4 WASSERSTEIN DISTANCES

We focus now on the case when the cost function c in the Optimal Transport problem is a distance d, following the approach of [5].
If (X,d) is a complete and separable metric space and $p \in [1,+\infty)$, we define

$$\mathcal{P}_p(X) := \left\{ \mu \in \mathcal{P}(X) \mid \int_X d^p(x,x_0)\, d\mu(x) < +\infty \text{ for some } x_0 \in X \right\}.$$

The p-Wasserstein distance between $\mu,\nu \in \mathcal{P}_p(X)$, denoted by $W_p(\mu,\nu)$, is defined as

$$W_p(\mu,\nu) := \left(\inf \left\{ \int_{X\times X} d^p \, d\boldsymbol{\gamma} \mid \boldsymbol{\gamma} \in \Gamma(\mu,\nu) \right\} \right)^{1/p}.$$

The set of optimal transport plans is denoted in this case by $\Gamma_o^p(\mu,\nu) \subset \mathcal{P}(X \times X)$. In case $p = 2$, we will just talk about Wasserstein distance and optimal plans, also omitting the index in the symbol denoting the collection of optimal plans. The metric space $(\mathcal{P}_p(X), W_p)$ is complete and separable and its topology is stronger than the narrow topology (in particular W_p convergent sequences are tight). More precisely, for a sequence $(\mu_n)_n \subset \mathcal{P}_p(X)$ and a point $\mu \in \mathcal{P}_p(X)$, we have

$$W_p(\mu_n,\mu) \to 0 \text{ if and only if } \begin{cases} \int_X d^p(x,x_0) d\mu_n(x) \to \int_X d^p(x,x_0) d\mu(x) \\ \qquad \text{for some } x_0 \in X, \\ \mu_n \to \mu \text{ in } \mathcal{P}(X). \end{cases} \tag{2.4.1}$$

Moreover, the p-Wasserstein distance is narrowly lower semicontinuous, meaning that, if $(\mu_n)_n, (\mu'_n)_n \subset \mathcal{P}_p(X)$, $\mu,, \mu' \in \mathcal{P}_p(X)$ and $\mu_n \to \mu$, $\mu'_n \to \mu'$ both narrowly, then we have

$$\liminf_n W_p(\mu_n,\mu'_n) \geqslant W_p(\mu,\mu').$$

The following is a criterion for compactness in $(\mathcal{P}_2(X), W_p)$).

Lemma 2.4.1. *A subset $\mathcal{K} \subset \mathcal{P}_p(X)$ is relatively compact w.r.t. the W_p-topology if and only if*

1. *$\mathcal{K}$ is tight,*
2. *$\mathcal{K}$ is uniformly p-integrable, i.e.*

$$\lim_{k\to\infty} \sup_{\mu\in\mathcal{K}} \int_{X\setminus B(x_0,k)} d^p(x_0,x)\,d\mu(x) = 0 \text{ for some } x_0 \in X. \tag{2.4.2}$$

Proof. The necessity of (1) is clear; regarding (2.4.2) notice that the functions

$$F_k : \mathcal{P}_p(X) \to [0,\infty), \quad F_k(\mu) := \int_{X\setminus B(x_0,k)} d^p(x_0,x)\,d\mu(x)$$

are upper semicontinuous, are decreasing w.r.t. k, and converge to 0 for every $\mu \in \mathcal{P}_p(X)$. Then, if $\mathcal{K}$ is relatively compact, they converge uniformly to 0 thanks to Dini's Theorem. This proves the necessity of (2).

To see that (1) and (2) are also sufficient to obtain the relative compactness, it is enough to check that every sequence $(\mu_n)_{n\in\mathbb{N}}$ in $\mathcal{K}$ has a convergent subsequence. Applying Theorem 2.1.4, we can find $\mu \in \mathcal{P}(X)$ and a subsequence $k \mapsto \mu_{n_k}$ such that $\mu_{n_k} \to \mu$ in $\mathcal{P}(X)$. Since $\int_X d^p(x,x_0)\,d\mu(x)$ is uniformly bounded, then $\mu \in \mathcal{P}_p(X)$. Applying [5, Lemma 5.1.7], we also obtain that

$$\lim_{k\to\infty} \int_X d^p(x,x_0)\,d\mu_n(x) = \int_X d^p(x,x_0)\,d\mu(x)$$

so that, by (2.4.1), we conclude that

$$\lim_{k\to\infty} W_p(\mu_{n_k},\mu) = 0.$$

□

We recall that, given $p, q \in [1,+\infty)$ with $q \leqslant p$ we have $\mathcal{P}_p(X) \subset \mathcal{P}_q(X)$ and

$$W_q(\mu,\nu) \leqslant W_p(\mu,\nu) \quad \text{for every } \mu,\nu \in \mathcal{P}_p(X). \tag{2.4.3}$$

The Kantorovich duality for the p-Wasserstein distance states that, for every $\mu,\nu \in \mathcal{P}_p(X)$, we have

$$W_p^p(\mu,\nu) = \sup\left\{\int_X \varphi d\mu + \int_X \psi d\nu \mid (\varphi,\psi) \in \mathrm{Adm}_p\right\}, \tag{2.4.4}$$

where Adm_p is the set of pairs $(\varphi,\psi) \in C_b(X) \times C_b(X)$ such that

$$\varphi(x) + \psi(y) \leqslant d^p(x,y) \quad \text{for every } x,y \in X.$$

When $p = 1$, (2.4.4) can be strengthened in

$$W_1(\mu,\nu) = \sup\left\{\int_X \varphi d(\mu-\nu) \mid \varphi \in \mathrm{Lip}_{b,1}(X)\right\},$$

where $\mathrm{Lip}_{b,1}(X)$ is the set of 1-Lipschitz continuous and bounded functions on X. It is easy to check that for every $f \in \mathrm{Lip}(X)$

$$\int_X f \, d(\mu - \nu) \leqslant \mathrm{Lip}(f, X) W_2(\mu, \nu), \tag{2.4.5}$$

since choosing $\boldsymbol{\mu} \in \Gamma_o(\mu, \nu)$ and setting $L := \mathrm{Lip}(f, X)$,

$$\begin{aligned}\int_X f \, d(\mu - \nu) &= \int (f(x) - f(y)) \, d\boldsymbol{\mu}(x, y) \\ &\leqslant L \int d \, d\boldsymbol{\mu} \\ &\leqslant L \Big(\int d^2 \, d\boldsymbol{\mu} \Big)^{1/2} = L W_2(\mu, \nu).\end{aligned}$$

As outlined in the introduction, the spaces $(\mathcal{P}_p(X), W_p)$ enjoys many interesting properties in case X has more structure i.e. when X is an Hilbert space or a Riemmanain manifold. For this reason, in the last part of this section, we will describe the properties of $(\mathcal{P}_2(\mathbb{H}), W_2)$ where $\mathbb{H}$ is a (possibily infinite dimensional) separable Hilbert space: many of the results that follow are still valid for a generic $p \in [1, +\infty)$ and also in case X is a Riemannian manifold, but since we will use them only in case $p = 2$ and X is a Hilbert space, we prefer to state them in this simpler context.

We use the notation

$$x^t : \mathbb{H}^2 \to \mathbb{H}, \quad x^t(x_0, x_1) := (1 - t)x_0 + t x_1, \quad (x_0, x_1) \in \mathbb{H}^2, \ t \in [0, 1]. \tag{2.4.6}$$

Definition 2.4.2 (Geodesics)**.** A map $\mu : [0, 1] \to \mathcal{P}_2(\mathbb{H})$ is said to be a *(constant speed) geodesic* if for every $0 \leqslant s \leqslant t \leqslant 1$ we have

$$W_2(\mu_s, \mu_t) = (t - s) W_2(\mu_0, \mu_1),$$

where we denoted by μ_t the evaluation of μ at time $t \in [0, 1]$. We also say that μ is a *geodesic from* μ_0 *to* μ_1.

Definition 2.4.3 (Convexity)**.** We say that $A \subset \mathcal{P}_2(\mathbb{H})$ is a *geodesically convex* set if for every pair $\mu_0, \mu_1 \in A$ there exists a geodesic μ from μ_0 to μ_1 such that $\mu_t \in A$ for every $t \in [0, 1]$.
We say that $A \subset \mathcal{P}_2(\mathbb{H})$ is *convex along couplings* if for any pair $\mu_0, \mu_1 \in A$ and any coupling $\boldsymbol{\gamma} \in \Gamma(\mu_0, \mu_1)$, we have that $(x^t)_\sharp \boldsymbol{\gamma} \in A$ for any $t \in [0, 1]$.

The following result gives a useful characterization of geodesics (see [5, Theorem 7.2.1, Theorem 7.2.2] for a proof).

Theorem 2.4.4 (Properties of geodesics)**.** *Let* $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{H})$ *and* $\boldsymbol{\mu} \in \Gamma_o(\mu_0, \mu_1)$. *Then* $\mu : [0, 1] \to \mathcal{P}_2(\mathbb{H})$ *defined by*

$$\mu_t := (x^t)_\sharp \boldsymbol{\mu}, \quad t \in [0, 1], \tag{2.4.7}$$

is a (constant speed) geodesic from μ_0 *to* μ_1. *Conversely, any (constant speed) geodesic* μ *from* μ_0 *to* μ_1 *admits the representation* (2.4.7) *for a suitable plan* $\boldsymbol{\mu} \in \Gamma_o(\mu_0, \mu_1)$.

Finally, if $\mu : [0,1] \to \mathcal{P}_2(\mathbb{H})$ *is a geodesic connecting* μ_0 *to* μ_1*, then for every* $t \in (0,1)$ *there exists a unique optimal plan* $\boldsymbol{\mu}_{t0}$ *between* μ_t *and* μ_0 *(resp.* $\boldsymbol{\mu}_{t1}$ *between* μ_t *and* μ_1*) and it is concentrated on a map w.r.t.* μ_t*, meaning that there exist Borel maps* $\mathbf{r}_t, \mathbf{r}'_t : \mathbb{H} \to \mathbb{H}$ *such that*

$$\boldsymbol{\mu}_{t0} = (\boldsymbol{i}_X, \mathbf{r}_t)_\sharp \mu_t, \quad \boldsymbol{\mu}_{t1} = (\boldsymbol{i}_X, \mathbf{r}'_t)_\sharp \mu_t.$$

The following definition generalizes the notion of smooth function in $\mathbb{R}^d$.

Definition 2.4.5 (Cyl($\mathbb{H}$))**.** We denote by $\Pi_d(\mathbb{H})$ the space of linear maps $\pi : \mathbb{H} \to \mathbb{R}^d$ of the form $\pi(x) = (\langle x, e_1 \rangle, \cdots, \langle x, e_d \rangle)$ for an orthonormal set $\{e_1, \cdots, e_d\}$ of $\mathbb{H}$. A function $\varphi : \mathbb{H} \to \mathbb{R}$ belongs to the space of cylindrical functions on $\mathbb{H}$, Cyl($\mathbb{H}$), if it is of the form

$$\varphi = \psi \circ \pi$$

where $\pi \in \Pi_d(\mathbb{H})$ and $\psi \in C_c^\infty(\mathbb{R}^d)$ for some $d \in \mathbb{N}$.

The following result (see [5, Theorem 8.3.1, Proposition 8.4.5 and Proposition 8.4.6]) characterizes locally absolutely continuous curves in $\mathcal{P}_2(\mathbb{H})$ defined on an open interval $\mathcal{I} \subset \mathbb{R}$. We use again the notation μ_t for the evaluation at time $t \in \mathcal{I}$ of a map $\mu : \mathcal{I} \to \mathcal{P}_2(\mathbb{H})$.

Theorem 2.4.6 (Wasserstein velocity field)**.** *Let* $\mu : \mathcal{I} \to \mathcal{P}_2(\mathbb{H})$ *be a locally absolutely continuous curve defined in an open interval* $\mathcal{I} \subset \mathbb{R}$*. There exist a Borel vector field* $\boldsymbol{v} : \mathcal{I} \times \mathbb{H} \to \mathbb{H}$ *and a set* $A(\mu) \subset \mathcal{I}$ *with* $\mathcal{L}(\mathcal{I} \setminus A(\mu)) = 0$ *such that the following hold:*

1. $\boldsymbol{v}_t \in \mathrm{Tan}_{\mu_t} \mathcal{P}_2(\mathbb{H}) := \overline{\{\nabla\varphi \mid \varphi \in \mathrm{Cyl}(\mathbb{H})\}}^{L^2_{\mu_t}(\mathbb{H};\mathbb{H})}$*, for every* $t \in A(\mu)$*;*
2. $\displaystyle\int_{\mathbb{H}} |\boldsymbol{v}_t|^2 \, d\mu_t = |\dot{\mu}_t|^2 := \lim_{h\to 0} \frac{W_2^2(\mu_{t+h}, \mu_t)}{h^2}$*, for every* $t \in A(\mu)$*;*
3. *the continuity equation*

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = 0$$

holds in the sense of distributions in $\mathcal{I} \times \mathbb{H}$*.*

Moreover, $\boldsymbol{v}_t$ *is uniquely determined in* $L^2_{\mu_t}(\mathbb{H};\mathbb{H})$ *for* $t \in A(\mu)$ *and*

$$\lim_{h\to 0} \frac{W_2((\boldsymbol{i}_X + h\boldsymbol{v}_t)_\sharp \mu_t, \mu_{t+h})}{|h|} = 0 \quad \textit{for every } t \in A(\mu). \tag{2.4.8}$$

We conclude this section recalling (see e.g. [5, Proposition 6.1.4]) that, given $\mu, \nu \in \mathcal{P}_2(\mathbb{H})$, $\gamma \in \Gamma(\mu,\nu)$ is optimal if and only if its support is cyclically monotone i.e.

$$\text{for every } N \in \mathbb{N} \text{ and } \{(x_n, y_n)\}_{n=1}^N \subset \operatorname{supp} \boldsymbol{\gamma} \text{ with } x_0 := x_N \text{ we have}$$
$$\sum_{n=1}^N \langle y_n, x_n - x_{n-1} \rangle \geqslant 0. \tag{2.4.9}$$

Part I

UNBALANCED OPTIMAL TRANSPORT

We present a new class of Optimal Transport costs for non-negative measures with possibly different masses. These are obtained by a convex relaxation procedure of a cost for non-negative Dirac masses. As a byproduct of our analysis, we show that the classical Optimal Transport cost can be obtained by the same procedure. A primal-dual formulation of the cost, optimality conditions and metric-topological properties are also presented.

3 TECHNICAL TOOLS AND EXAMPLES OF COST FUNCTIONS

The aim of this chapter is to introduce some of the preliminary material used in the sequel of Part I. In particular, in Section 3.1 we present some approximation results related to discrete measures and we prove that sufficiently rich subalgebras of continuous functions generate the narrow topology; Section 3.2 is completely devoted to the construction and to the properties of the geometric cone $\mathfrak{C}[X]$ on a completely regular space X, besides also listing some definition related to functions defined on the cone; finally Section 3.3 presents a few examples of admissible cost functions H for the unbalanced Optimal Transport problem.

This Chapter is the result of a collaboration with Giuseppe Savaré and part of Section 3.1 appeared in [109].

3.1 APPROXIMATION THROUGH DISCRETE MEASURES AND FUNCTIONS

Proposition 3.1.1. *Let X be a completely regular space. Then*

$$\overline{\mathrm{Discr}_+(X)} = \mathcal{M}_+(X).$$

Proof. The null measure is already a discrete measure. Notice moreover that $\mathcal{M}_+(X) \setminus \{0_X\}$ is open. Hence the thesis is equivalent to prove that $\mathrm{Discr}_+(X)$ intersects every open set contained in $\mathcal{M}_+(X) \setminus \{0_X\}$. Take $\mu_0 \in \mathcal{M}_+(X) \setminus \{0_X\}$ and any U open neighbourhood of μ_0 in $\mathcal{M}_+(X) \setminus \{0_X\}$. There exist $n \geqslant 1$, $\varepsilon > 0$, $\{f_i\}_{i=1}^n \subset C_b(X)$ s.t.

$$\mu_0 \in V := \{\mu \in \mathcal{M}_+(X) \mid |\mu(f_i) - \mu_0(f_i)| < \varepsilon \ i = 1, \ldots, n\} \subset U$$

hence it is enough to show that $\mathrm{Discr}_+(X)$ intersects V i.e. that there exists $\nu \in \mathrm{Discr}_+(X)$ s.t.

$$|\nu(f_i) - \mu_0(f_i)| < \varepsilon \quad \text{for every } i = 1, \ldots, n. \tag{3.1.1}$$

Take simple functions $\{g_i\}_{i=1}^n$ s.t. $\sup_{x \in X} |f_i(x) - g_i(x)| < \frac{\varepsilon}{4\mu_0(X)}$. Then, if we can find $\nu \in \mathrm{Discr}_+(X)$, s.t. $\nu(g_i) = \mu_0(g_i)$ for every $i = 1, \ldots, n$ with $\nu(X) = \mu_0(X)$, (3.1.1) holds immediately because

$$|\nu(f_i) - \mu_0(f_i)| \leqslant |\nu(f_i) - \nu(g_i)| + |\nu(g_i) - \mu_0(g_i)| + |\mu_0(g_i) - \mu_0(f_i)| \leqslant \varepsilon/2.$$

Now, every g_i is of the form

$$g_i := \sum_{k=1}^{m_i} \alpha_k^i \chi_{E_i^k}$$

with $\{\alpha_k^i\}_{k=1}^{m_i} \subset \mathbb{R}$ and $\{E_i^k\}_{k=1}^{m_i} \subset \mathscr{B}(X)$ disjoint for every $i = 1, \dots, n$. Take the partition $\{P_k\}_{k=1}^m$ obtained as the intersection of all the other partitions, so that

$$g_i = \sum_{k=1}^{m} \beta_k^i \chi_{P_k}$$

for some $\{\beta_k^i\}_{k=1}^m \subset \mathbb{R}$ for every $i = 1, \dots, n$. Hence we have $\nu(g_i) = \mu_0(g_i)$ for every $i = 1, \dots, n$ if and only if

$$\sum_{k=1}^{m} \beta_k^i \nu(P_k) = \sum_{k=1}^{m} \beta_k^i \mu_0(P_k) \quad \text{for every } i = 1, \dots, n$$

and then it is enough to have $\nu(P_k) = \mu_0(P_k)$ for every $k = 1, \dots, m$. Take $\{x_k\}_{k=1}^m \subset X$ s.t. $x_k \in P_k$ for every $k = 1, \dots, m$. Then we can define

$$\nu := \sum_{k=1}^{m} \mu_0(P_k)\delta_{x_k}.$$

This is the sought ν. □

Remark 3.1.2. Observe that in the above proof we have shown that for every $\mu_0 \in \mathcal{M}_+(X)$ and every $U \subset \mathcal{M}(X)$ open neighbourhood of μ_0 there exists $\nu \in \mathrm{Discr}_+(X)$ with the same mass of μ_0 s.t. $\nu \in U$. Hence

$$\overline{\mathrm{Discr}_+(X) \cap \{\mu \in \mathcal{M}_+X \mid \mu(X) = c\}} = \{\mu \in \mathcal{M}_+X \mid \mu(X) = c\},$$

where c is a nonnegative real number.

Remark 3.1.3. It holds

$$\overline{\mathrm{co}}\,(\Delta_+(X)) = \mathcal{M}_+(X).$$

This is an immediate consequence of the fact that $\mathrm{co}\,(\Delta_+(X)) = \mathrm{Discr}_+(X)$ and Proposition 3.1.1.

The following Lemma is a refinement of Proposition 3.1.1 showing that, given a Borel function f, we can construct an approximating sequence of discrete measures for which we have convergence also of the integral of f.

Lemma 3.1.4. *Let X be a completely regular space and let $\alpha \in \mathcal{M}_+(X)$. Let $f : X \to [0, +\infty]$ be a Borel function. Then there exists a net $(\gamma_\lambda)_{\lambda \in \mathbb{L}} \subset \mathrm{Discr}_+(X) \cap \{\mu \in \mathcal{M}_+(X) \mid \mu(X) = \alpha(X)\}$ s.t.*

$$\lim_{\lambda \in \mathbb{L}} \gamma_\lambda = \alpha, \quad \lim_{\lambda \in \mathbb{L}} \int_X f \, d\gamma_\lambda = \int_X f \, d\alpha.$$

Proof. By Lusin's theorem, we can find an increasing sequence of closed sets $\tilde{X}_k \subset \tilde{X}_{k+1} \subset X$ s.t.

$$\alpha(X \setminus \tilde{X}_k) \leqslant \frac{1}{2k}, \quad f|_{\tilde{X}_k} \text{ is continuous} \qquad \text{for every } k \geqslant 1.$$

Moreover observe that, since α is a Radon measure, we can find an increasing sequence of compact sets $B_k \subset B_{k+1} \subset X$ s.t.

$$\alpha(X \setminus B_k) \leqslant \frac{1}{2k} \quad \text{for every } k \geqslant 1.$$

Hence we can define

$$X_k := \tilde{X}_k \cap B_k \quad \text{for every } k \geqslant 1.$$

Then, this new sequence of compact sets is such that

$$X_k \subset X_{k+1} \quad \text{for every } k \geqslant 1, \quad \alpha(X \setminus X_k) \leqslant \frac{1}{k}, \quad f|_{X_k} \text{ is bounded and continuous.}$$

Consider now the family of measures $\{\alpha_k\}_{k\geqslant 1} \subset \{\mu \in \mathcal{M}_+(X) \mid \mu(X) = \alpha(X)\}$ defined as

$$\alpha_k := \frac{\alpha(X)}{\alpha(X_k)} \alpha|_{X_k} \quad \text{for every } k \geqslant 1.$$

We can easily observe that

$$\lim_k \alpha_k = \alpha$$

indeed, if $\varphi \in C_b(X)$, we have

$$\lim_k \int_X \varphi \, d\alpha_k = \lim_k \frac{\alpha(X)}{\alpha(X_k)} \int_X \varphi \chi_{X_k} \, d\alpha = \int_X \varphi \, d\alpha$$

by monotone convergence. The same argument shows that we also have

$$\lim_k \int_X f \, d\alpha_k = \int_X f \, d\alpha. \tag{3.1.2}$$

By Proposition 3.1.1 and Remark 3.1.2, for every $k \geqslant 1$, we can find a net $\{\gamma^k_{\lambda_k}\}_{\lambda_k \in \mathbb{L}_k} \subset \mathrm{Discr}_+(X_k) \cap \{\mu \in \mathcal{M}_+(X) \mid \mu(X) = \alpha(X)\}$, such that

$$\lim_{\lambda_k \in \mathbb{L}_k} \gamma^k_{\lambda_k} = \alpha_k.$$

Moreover, since $f|_{X_k}$ is bounded and continuous, it holds

$$\lim_{\lambda_k \in \mathbb{L}_k} \int_X f \, d\gamma^k_{\lambda_k} = \lim_{\lambda_k \in \mathbb{L}_k} \int_{X_k} f \, d\gamma^k_{\lambda_k} = \int_{X_k} f \, d\alpha_k = \int_X f \, d\alpha_k.$$

This allows us to find, for every $k \geqslant 1$, some $\bar{m}(k) \in \mathbb{L}_k$ s.t.

$$\left| \int_X f \, d\gamma^k_{\lambda_k} - \int_X f \, d\alpha_k \right| \leqslant \frac{1}{k} \quad \text{for every } \lambda_k \geqslant \bar{m}(k).$$

Hence we can consider, for every $k \geqslant 1$, the directed sets $\mathbb{E}_k := \{\lambda_k \in \mathbb{L}_k \mid \lambda_k \geqslant \bar{m}(k)\}$ and the corresponding new sequence of nets $\{\gamma^k_{\lambda_k}\}_{\lambda_k \in \mathbb{E}_k}$ on varying of $k \geqslant 1$. Obviously it holds

$$\lim_{\lambda_k \in \mathbb{E}_k} \gamma^k_{\lambda_k} = \alpha_k, \quad \left| \int_X f \, d\gamma^k_{\lambda_k} - \int_X f \, d\alpha_k \right| \leqslant \frac{1}{k} \quad \text{for every } \lambda_k \in \mathbb{E}_k. \tag{3.1.3}$$

Define now the directed set

$$\mathbb{N} \otimes \mathbb{E}_k := \{(k,\lambda) \mid \lambda \in \mathbb{E}_k\} \text{ with } (k,\lambda) \leqslant (k',\lambda') \iff k < k' \text{ or } (k = k' \wedge \lambda \leqslant \lambda').$$

By the diagonal principle for nets, we can find a directed set $\mathbb{B}$ and a monotone final function

$$h : \mathbb{B} \to \mathbb{N} \otimes \mathbb{E}_k, \quad h(\beta) = (h_1(\beta), h_2(\beta)) \text{ with } h_2(\beta) \in \mathbb{E}_{h_1(\beta)} \quad \text{for every } \beta \in \mathbb{B}$$

such that the diagonal net $\{\gamma_\beta\}_{\beta \in \mathbb{B}} := \{\gamma^{h_1(\beta)}_{h_2(\beta)}\}_{\beta \in \mathbb{B}} \subset \mathrm{Discr}_+(X) \cap \{\mu \in \mathcal{M}_+(X) \mid \mu(X) = \alpha(X)\}$ converges to α. We only need to prove that also the integral of f converges:

$$\begin{aligned}\left|\int_X f\,d\gamma_\beta - \int_X f\,d\alpha\right| &\leqslant \left|\int_X f\,d\gamma_\beta - \int_X f\,d\alpha_{h_1(\beta)}\right| + \left|\int_X f\,d\alpha_{h_1(\beta)} - \int_X f\,d\alpha\right| \\ &= \left|\int_X f\,d\gamma^{h_1(\beta)}_{h_2(\beta)} - \int_X f\,d\alpha_{h_1(\beta)}\right| + \left|\int_X f\,d\alpha_{h_1(\beta)} - \int_X f\,d\alpha\right| \\ &\leqslant \frac{1}{h_1(\beta)} + \left|\int_X f\,d\alpha_{h_1(\beta)} - \int_X f\,d\alpha\right|,\end{aligned}$$

where we have used (3.1.3). Now it is enough to observe that $h_1 : \mathbb{B} \to \mathbb{N}$ is a final monotone function i.e. it is an increasing monotone sequence converging to $+\infty$. Passing to $\lim_{\beta \in \mathbb{B}}$ and using (3.1.2), we conclude. □

The following statement concerns sub-algebras $A \subset C_b(X)$ which are rich enough to characterize weak convergence. We first state the relevant definition.

Definition 3.1.5. (Adapted algebra of continuous functions)
Let X be a completely regular space. We say that a unital subalgebra $A \subset C_b(X)$ is *adapted* if the topology of X coincides with the initial topology induced by A. Equivalently, for every net $(x_\lambda)_{\lambda \in \mathbb{L}}$ in X

$$\lim_{\lambda \in \mathbb{L}} x_\lambda = x \quad \Leftrightarrow \quad \lim_{\lambda \in \mathbb{L}} f(x_\lambda) = f(x) \quad \text{for every } f \in A.$$

Since X is Hausdorff, it is immediate to check that an adapted algebra A separates the points of X. It is interesting that the above condition is also sufficient to recover the weak topology of $\mathcal{M}(X)$.

Lemma 3.1.6. *Let X be a completely regular space and let $A \subset C_b(X)$ be an adapted algebra. Then a net $(\mu_\lambda)_{\lambda \in \mathbb{L}}$ in $\mathcal{M}(X)$ weakly converges to μ if and only if*

$$\lim_{\lambda \in \mathbb{L}} \int_X f\,d\mu_\lambda = \int_X f\,d\mu \quad \textit{for every } f \in A. \tag{3.1.4}$$

Equivalently, the weak topology of $\mathcal{M}(X)$ coincides with $\sigma(\mathcal{M}(X), A)$.

Proof. We consider only the nontrivial implication and we will show that a net $(\mu_\lambda)_{\lambda \in \mathbb{L}}$ satisfying (3.1.4) weakly converges in $\mathcal{M}(X)$.

Let us set $I_f := [\inf_X f, \sup_X f] \subset \mathbb{R}$ and let us consider the product space $Y = \prod_{f \in A} I_f$ endowed with the product topology; the component of a point in

$y \in Y$ will be denoted as y_f with $f \in A$. Y is compact by Tychonoff's Theorem. Since A is adapted, the map

$$\iota : X \to Y \quad \text{defined by} \quad \iota(x)_f := f(x) \quad \text{for every } x \in X$$

is a topological embedding. By Lemma 2.1.3(3) it is then sufficient to show that the net $\hat{\mu}_\lambda := \iota_\sharp \mu_\lambda$ weakly converges to $\hat{\mu} := \iota_\sharp \mu$ in $\mathcal{M}(Y)$. Let $\mathcal{B}$ be the unital algebra obtained by functions of the form

$$\varphi_{F,P}(y) = P(y_{f_1}, y_{f_2}, \cdots, y_{f_k}), \quad y \in Y,\ F = \{f_1, f_2, \cdots, f_k\} \subset A,\ P \text{ polynomial.}$$

Since $\mathcal{B}$ contains the unit and separates the points of Y, by Stone-Weierstrass theorem $\mathcal{B}$ is uniformly dense in $C_b(Y)$, so that in order to check the convergence of $\hat{\mu}_\lambda$ is sufficient to test them against functions of $\mathcal{B}$. We have

$$\begin{aligned}
\lim_{\lambda \in \mathbb{L}} \int_Y \varphi_{F,P}(y)\, d\hat{\mu}_\lambda(y) &= \lim_{\lambda \in \mathbb{L}} \int_X P(\iota(x)_{f_1}, \iota(x)_{f_2}, \cdots, \iota(x)_{f_k})\, d\mu_\lambda(x) \\
&= \lim_{\lambda \in \mathbb{L}} \int_X P(f_1(x), f_2(x), \cdots, f_k(x))\, d\mu_\lambda(x) \\
&= \int_X P(f_1(x), f_2(x), \cdots f_k(x))\, d\mu(x) \\
&= \int_Y \varphi_{F,P}(y)\, d\hat{\mu}(y),
\end{aligned}$$

where we used (3.1.4) and the fact that the function $x \mapsto P(f_1(x), f_2(x), \cdots f_k(x))$ belongs to the algebra A as well. □

3.2 THE CONE CONSTRUCTION

It will be natural to state some definitions and results in the context of the so called geometric cone: we introduce on $X \times \mathbb{R}_+$ the equivalence relation

$$(x, r) \sim (y, s) \quad \stackrel{\text{def}}{\Leftrightarrow} \quad [x = y, r = s \neq 0 \quad \vee \quad r = s = 0]$$

and the corresponding geometric cone $\mathfrak{C}[X] = (X \times \mathbb{R}_+)/\sim$, whose points are denoted by gothic letters like $\mathfrak{y}$. We denote by $\mathfrak{p}$ the quotient map $\mathfrak{p} : X \times \mathbb{R}_+ \to \mathfrak{C}[X]$ sending a point (x, r) to its equivalence class $[x, r]$. Notice that $\mathfrak{p}$ is just the identity map except for those points with $r = 0$, which are all sent to the same equivalence class, the so called vertex of the cone that we denote with $\mathfrak{o}$.

On the cone we introduce the projections on $\mathbb{R}_+$ and X simply defined as $r([x, r]) = r$ and $x([x, r]) = x$ if $r > 0$ and $x([x, r]) = \bar{x}$ if $r = 0$, where $\bar{x} \in X$ is some fixed point. We omit the dependence of x on $\bar{x}$ since in the constructions where x is involved this will be irrelevant.

We can define a right inverse of $\mathfrak{p}$ as

$$\mathfrak{q}([x, r]) = (x([x, r]), r([x, r])). \tag{3.2.1}$$

Notice that also $\mathfrak{q}$ depends on the choice of $\bar{x}$ in the definition of x but, again, this will be irrelevant.

On $\mathfrak{C}[X]$ we consider the following topology, weaker than the quotient one: a local system of neighbourhoods of a point $[x, r]$ is just the image trough $\mathfrak{p}$ of the local system of neighbourhoods given by the product topology at $(x, r) \in X \times \mathbb{R}_+$, if $r > 0$. A local system of neighbourhoods at $\mathfrak{o}$ is given by

$$\{\{[x, r] \in \mathfrak{C}[X] \mid 0 \leqslant r < \varepsilon\} \mid \varepsilon > 0\}.$$

If the topology of X is induced by a distance d, then the topology of $\mathfrak{C}[X]$ is induced by the distance $d_{\mathfrak{C}} : \mathfrak{C}[X] \times \mathfrak{C}[X] \to [0, +\infty)$ defined as

$$d_{\mathfrak{C}}([x, r], [y, s]) := \left(r^2 + s^2 - 2rs\cos(d(x, y) \wedge \pi)\right)^{\frac{1}{2}}, \quad [x, r], [y, s] \in \mathfrak{C}[X]. \tag{3.2.2}$$

With the above topology, $\mathfrak{C}[X]$ is completely regular and it is the right object to consider when one wants to represent elements in $\Delta_+(X)$; in particular we have the following result.

Lemma 3.2.1. *Let X be a completely regular space. Then $\Delta_+(X)$ is homeomorphic to $\mathfrak{C}[X]$.*

Proof. The map $\varphi : \mathfrak{C}[X] \to \Delta_+(X)$ given by

$$\varphi([x, r]) := \begin{cases} r\delta_x & \text{if } r > 0, \\ 0_X & \text{if } r = 0 \end{cases}$$

can be checked to be the sought homeomorphism. □

If $R > 0$, we define

$$\mathfrak{C}_R[X] := \{[x, r] \in \mathfrak{C}[X] \mid 0 \leqslant r \leqslant R\} \tag{3.2.3}$$

and we will often identify measures on $\mathfrak{C}[X]$ with support contained in $\mathfrak{C}_R[X]$ with elements of $\mathcal{M}(\mathfrak{C}_R[X])$. There is a natural product operation on the cone given by

$$\lambda[x, r] := [x, \lambda r] \quad \text{for every } \lambda, r \geqslant 0,\ x \in X.$$

For every $p \geqslant 1$, we introduce moreover the set

$$\mathfrak{M}^p_+(\mathfrak{C}[X]) := \left\{\alpha \in \mathcal{M}_+(\mathfrak{C}[X]) \mid \int_{\mathfrak{C}[X]} r^p \, d\alpha < +\infty\right\},$$

and the map

$$\mathfrak{h}^p : \mathfrak{M}^p_+(\mathfrak{C}[X]) \to \mathcal{M}_+(X), \quad \mathfrak{h}^p(\alpha) = (x)_\sharp(r^p\alpha).$$

Notice that the map $\mathfrak{h}^p$ does not depend on the point $\bar{x}$ in the definition of x.
We introduce now the product cone: given X_1 and X_2 completely regular spaces, we define $\mathfrak{C}[X_1, X_2] := \mathfrak{C}[X_1] \times \mathfrak{C}[X_2]$ endowed with the product topology. Points in the product cone are denoted by bold gothich letters like $\boldsymbol{\mathfrak{y}} = (\mathfrak{y}_1, \mathfrak{y}_2) = ([x_1, r_1]; [x_2, r_2])$. On the product cone we can consider the projections on the two components $\pi^{\mathfrak{C}_i} : \mathfrak{C}[X_1, X_2] \to \mathfrak{C}[X_i]$ sending $([x_1, r_1]; [x_2, r_2])$ to $[x_i, r_i]$ and

the projections on $\mathbb{R}_+$ and X_i simply defined as $\mathsf{r}_i := \mathsf{r} \circ \pi^{\mathfrak{C}_i}$ and $\mathsf{x}_i := \mathsf{x} \circ \pi^{\mathfrak{C}_i}$ (x_i depends on the choice of a point $\bar{x}_i \in X_i$, but this will be irrelevant). In analogy with (3.2.2), if the topologies of X_1 and X_2 are induced by distances d_1 and d_2 respectively, the topology of the product cone is induced by the distance

$$(d_1 \otimes_{\mathfrak{C}} d_2)((\mathfrak{y}_1, \mathfrak{y}_2), (\mathfrak{w}_1, \mathfrak{w}_2)) := \left(d^2_{1,\mathfrak{C}}(\mathfrak{y}_1, \mathfrak{w}_1) + d^2_{2,\mathfrak{C}}(\mathfrak{y}_2, \mathfrak{w}_2)\right)^{\frac{1}{2}}, \tag{3.2.4}$$

for every $(\mathfrak{y}_1, \mathfrak{y}_2), (\mathfrak{w}_1, \mathfrak{w}_2) \in \mathfrak{C}[X_1, X_2]$. As in (3.2.3), given $R > 0$, we define

$$\mathfrak{C}_R[X_1, X_2] := \mathfrak{C}_R[X_1] \times \mathfrak{C}_R[X_2] = \{\mathfrak{y} \in \mathfrak{C}[X_1, X_2] \mid 0 \leqslant \mathsf{r}_i(\mathfrak{y}) \leqslant R,\ i = 1, 2\} \tag{3.2.5}$$

and as in the previous case, we identify measures on $\mathfrak{C}[X_1, X_2]$ with support contained in $\mathfrak{C}_R[X_1, X_2]$ with elements of $\mathcal{M}(\mathfrak{C}_R[X_1, X_2])$. For every $p \geqslant 1$, we introduce the set

$$\mathfrak{M}^p_+(\mathfrak{C}[X_1, X_2]) := \left\{\boldsymbol{\alpha} \in \mathcal{M}_+(\mathfrak{C}[X_1, X_2]) \mid \int_{\mathfrak{C}[X_1, X_2]} (\mathsf{r}_1^p + \mathsf{r}_2^p)\, d\boldsymbol{\alpha} < +\infty\right\},$$

and the maps

$$\mathfrak{h}^p_i : \mathfrak{M}^p_+(\mathfrak{C}[X_1, X_2]) \to \mathcal{M}_+(X_i), \quad \mathfrak{h}^p_i(\boldsymbol{\alpha}) = (\mathsf{x}_i)_\sharp(\mathsf{r}_i^p \boldsymbol{\alpha}).$$

Notice that the map $\mathfrak{h}^p_i$ does not depend on the point $\bar{x}_i \in X_i$ in the definition of x_i.

Finally we define, for every $(\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$ and every $p \geqslant 1$, the set

$$\mathfrak{H}^p(\mu_1, \mu_2) := \left\{\boldsymbol{\alpha} \in \mathfrak{M}^p_+(\mathfrak{C}[X_1, X_2]) \mid \mathfrak{h}^p_i(\boldsymbol{\alpha}) = \mu_i,\ i = 1, 2\right\}. \tag{3.2.6}$$

If $\boldsymbol{\alpha} \in \mathfrak{H}^p(\mu_1, \mu_2)$, we say that μ_1 and μ_2 are the p-*homogeneous marginals of* $\boldsymbol{\alpha}$. The following result comes from [76].

Lemma 3.2.2. *Let X_i for $i = 1, 2$ be completely regular spaces and let $p \geqslant 1$. Given $\boldsymbol{\alpha} \in \mathcal{M}_+(\mathfrak{C}[X_1, X_2])$ and $\vartheta : \mathfrak{C}[X_1, X_2] \to (0, +\infty)$ Borel measurable in $L^p(\mathfrak{C}[X_1, X_2], \boldsymbol{\alpha})$ we can define*

$$\mathrm{prd}_\vartheta(\mathfrak{y}) := \left(\vartheta(\mathfrak{y})^{-1}\mathfrak{y}_1, \vartheta(\mathfrak{y})^{-1}\mathfrak{y}_2\right), \quad \mathfrak{y} \in \mathfrak{C}[X_1, X_2],$$
$$\mathrm{dil}_{\vartheta,p}(\boldsymbol{\alpha}) := (\mathrm{prd}_\vartheta)_\sharp(\vartheta^p \boldsymbol{\alpha}).$$

Then we have

$$\mathfrak{h}_i(\mathrm{dil}_{\vartheta,p}(\boldsymbol{\alpha})) = \mathfrak{h}^p_i(\boldsymbol{\alpha}), \quad i = 1, 2.$$

In particular, if we define

$$\vartheta_{\boldsymbol{\alpha},p}(\mathfrak{y}) := \frac{1}{r^*(\boldsymbol{\alpha})} \begin{cases} \mathsf{r}_1^p(\mathfrak{y}) + \mathsf{r}_2^p(\mathfrak{y}) & \text{if } \mathfrak{y} \neq (o, o) \\ 1 & \text{if } \mathfrak{y} = (o, o) \end{cases},$$

where $r^*(\boldsymbol{\alpha})$ is a normalization constant s.t. $\int_{\mathfrak{C}[X_1, X_2]} \vartheta^p_{\boldsymbol{\alpha},p}\, d\boldsymbol{\alpha} = 1$ given by

$$r^*(\boldsymbol{\alpha}) := \int_{\mathfrak{C}[X_1, X_2]} (\mathsf{r}_1^p(\mathfrak{y}) + \mathsf{r}_2^p(\mathfrak{y}))\, d\boldsymbol{\alpha} + \boldsymbol{\alpha}((o, o)),$$

we have that $\mathrm{dil}_{\vartheta_{\boldsymbol{\alpha},p},p}(\boldsymbol{\alpha}) \in \mathcal{P}(\mathfrak{C}[X_1, X_2])$, has the same p-homogeneous marginals of $\boldsymbol{\alpha}$ and its support is contained in $\mathfrak{C}_{r^*(\boldsymbol{\alpha})}[X_1, X_2]$.

Theorem 3.2.3. *(Compactness from converging marginals)*
Let X_i for $i = 1, 2$ be completely regular spaces. Let $(\boldsymbol{\gamma}_\lambda)_{\lambda\in\mathbb{L}}$ be a net in $\mathcal{M}_+(X_1 \times X_2)$ with $\mu_{i,\lambda} := \pi^i_\sharp \boldsymbol{\gamma}_\lambda \in \mathcal{M}_+(X_i)$, $i = 1, 2$, $\lambda \in \mathbb{L}$. If $(\mu_{i,\lambda})_{\lambda\in\mathbb{L}}$ weakly converge to some μ_i in $\mathcal{M}(X_i)$, then there exists a subnet $(\boldsymbol{\gamma}'_\alpha)_{\alpha\in\mathbb{A}}$ weakly convergent to some $\boldsymbol{\gamma} \in \Gamma(\mu_1, \mu_2)$ in $\mathcal{M}(X_1 \times X_2)$.

Proof. We recall that every completely regular space can be topologically embedded in a compact Hausdorff space (e.g. by the construction we used in the proof of Lemma 3.1.6: this property, in fact, characterizes completely regular spaces). Up to an identification of X_i with its homeomorphic image, we can thus assume that X_i is a subset of a compact Hausdorff spaces $\hat{X}_i$; thanks to Lemma 2.1.3(3), we can also identify the measures $\mu_{i,\lambda}$, μ_i in $\mathcal{M}_+(X_i)$ with corresponding measures $\hat{\mu}_{i,\lambda}$, $\hat{\mu}_i$ in $\mathcal{M}_+(X_i, \hat{X}_i)$ concentrated on X_i s.t. $\hat{\mu}_{i,\lambda} \to \hat{\mu}_i$ weakly in $\mathcal{M}(\hat{X}_i)$. Similarly, we can identify each $\boldsymbol{\gamma}_\lambda$ with a measure $\hat{\boldsymbol{\gamma}}_\lambda$ in $\mathcal{M}_+(\hat{X}_1 \times \hat{X}_2)$ concentrated on $X_1 \times X_2$. Since $\hat{X}_1 \times \hat{X}_2$ is compact and the total mass $\hat{\boldsymbol{\gamma}}_\lambda(\hat{X}_1 \times X_2)$ is converging and thus it is eventually bounded, by Lemma 2.1.3(4) there exist some $\tilde{\boldsymbol{\gamma}} \in \mathcal{M}_+(\hat{X}_1 \times \hat{X}_2)$ and a subnet $\boldsymbol{\gamma}'_\alpha = \boldsymbol{\gamma}_{\lambda(\alpha)}$ (with corresponding subnet $\hat{\boldsymbol{\gamma}}'_\alpha = \hat{\boldsymbol{\gamma}}_{\lambda(\alpha)}$) induced by a map $\alpha \mapsto \lambda(\alpha)$, $\alpha \in \mathbb{A}$ (see [50, Theorem 4.29]), such that

$$\hat{\boldsymbol{\gamma}}'_\alpha \to \tilde{\boldsymbol{\gamma}} \quad \text{weakly in } \mathcal{M}(\hat{X}_1 \times \hat{X}_2).$$

On the other hand, since the marginals of $\hat{\boldsymbol{\gamma}}_\lambda$ are $\hat{\mu}_{i,\lambda}$ and $\hat{\mu}_{i,\lambda} \to \hat{\mu}_i$ weakly in $\mathcal{M}(\hat{X}_i)$, we deduce that the marginals of $\tilde{\boldsymbol{\gamma}}$ on $\hat{X}_i$ are $\hat{\mu}_i$. Since $\hat{\mu}_i$ are Radon measures concentrated on two sigma compact subsets $D_i \subset X_i$, we have

$$\tilde{\boldsymbol{\gamma}}((\hat{X}_1 \times \hat{X}_2) \setminus (D_1 \times D_2)) \leqslant \mu_1(X_1 \setminus D_1) + \mu_2(X_2 \setminus D_2) = 0.$$

It follows that $\tilde{\boldsymbol{\gamma}}$ is concentrated on $X_1 \times X_2$, and therefore can be written as $\tilde{\boldsymbol{\gamma}} = \iota_\sharp \boldsymbol{\gamma}$ for a measure $\boldsymbol{\gamma} \in \mathcal{M}(X_1 \times X_2)$. A further application of Lemma 2.1.3(3) yields that $\boldsymbol{\gamma}'_\alpha$ weakly converges to $\boldsymbol{\gamma}$ in $\mathcal{M}(X_1 \times X_2)$ and Lemma 2.1.3(1) shows that $\boldsymbol{\gamma} \in \Gamma(\mu_1, \mu_2)$. □

Let us state an immediate consequence of the previous result.

Corollary 3.2.4. *(Compactness from compact marginals)*
Let X_i for $i = 1, 2$ be completely regular spaces. Let $\mathcal{K}_i \subset \mathcal{M}_+(X_i)$ be compact in the weak topology, $i = 1, 2$. Then the set $\mathcal{K} := \left\{ \boldsymbol{\gamma} \in \mathcal{M}_+(X_1 \times X_2) \mid \pi^i_\# \boldsymbol{\gamma} \in \mathcal{K}_i \right\}$ is compact in the weak topology of $\mathcal{M}(X_1 \times X_2)$.

Proof. Since $\mathcal{K}$ is closed in $\mathcal{M}(X_1 \times X_2)$ thanks to Claim 1 of Lemma 2.1.3, it is sufficient to prove that every net $(\boldsymbol{\gamma}_\lambda)_{\lambda\in\mathbb{L}}$ in $\mathcal{K}$ has a converging subnet.

Setting $\mu_{i,\lambda} := \pi^i_\sharp \boldsymbol{\gamma}_\lambda$, thanks to the compactness of $\mathcal{K}_1 \times \mathcal{K}_2$ we can find a subnet $(\mu'_{1,\alpha}, \mu'_{2,\alpha})_{\alpha\in\mathbb{A}}$, $\mu'_{i,\alpha} = \mu_{i,\alpha(\lambda)}$, converging to $(\mu_1, \mu_2) \in \mathcal{K}_1 \times \mathcal{K}_2$ in $\mathcal{M}(X_1) \times \mathcal{M}(X_2)$.

Applying Theorem 3.2.3 we can find a further subnet $(\boldsymbol{\gamma}''_\beta)_{\beta\in\mathbb{B}}$ of $(\boldsymbol{\gamma}'_\alpha)_{\alpha\in\mathbb{A}}$ converging to a measure $\boldsymbol{\gamma}$ in $\mathcal{M}(X_1 \times X_2)$. □

Remark 3.2.5. In locally compact or Polish spaces Corollary 3.2.4 could also be proven by using Prokhorov's characterization of compact subsets of $\mathcal{M}_+(X_i)$ in terms of uniform tightness (see Theorem 2.1.4). The argument we are presenting here is more direct (once Radon measures are involved) and works in completely regular spaces as well. In the case of arbitrary topological spaces, one has to deal with a more refined definition of the weak topology and Corollary 3.2.4 can also be extended to this general setting. Since we think that this result is of independent interest, we added its proof in the Section 4.3.

Lemma 3.2.6. *Let X_i, $i = 1, 2$ be completely regular spaces, let $p \geqslant 1$ and let $(\boldsymbol{\alpha}_\lambda)_{\lambda\in\mathbb{L}}$ be a net in $\mathcal{P}(\mathfrak{C}_R[X_1, X_2])$ for some $R > 0$ with $\mu_{i,\lambda} := (\mathfrak{h}^p_i)_\sharp \boldsymbol{\alpha}_\lambda \in \mathcal{M}_+(X_i)$, $i = 1, 2$, $\lambda \in \mathbb{L}$. If $(\mu_{i,\lambda})_{\lambda\in\mathbb{L}}$ converge to some μ_i in $\mathcal{M}(X_i)$, then there exists a subnet $(\boldsymbol{\alpha}'_\beta)_{\beta\in\mathbb{B}}$ of $(\boldsymbol{\alpha}_\lambda)_{\lambda\in\mathbb{L}}$ convergent to some $\boldsymbol{\alpha} \in \mathfrak{H}^p(\mu_1, \mu_2)$.*

Proof. Thanks to Theorem 3.2.3, it is enough to prove that $\{\pi^{\mathfrak{C}_i}_\sharp \boldsymbol{\alpha}_\lambda\}_{\lambda\in\mathbb{L}}$ converges, up to a subnet, for every $i = 1, 2$. Then, let X be a completely regular space, let $(\mu_\lambda)_{\lambda\in\mathbb{L}} \subset \mathcal{M}_+(X)$ be convergent and let $(\boldsymbol{\alpha}_\lambda)_{\lambda\in\mathbb{L}} \subset \mathcal{P}(\mathfrak{C}_R[X])$ with $\mathfrak{h}^p(\boldsymbol{\alpha}_\lambda) = \mu_\lambda$ for every $\lambda \in \mathbb{L}$. Define, for every $\lambda \in \mathbb{L}$,

$$\boldsymbol{\beta}_\lambda := \mathfrak{q}_\sharp(r^p \boldsymbol{\alpha}_\lambda) \in \mathcal{M}_+([0, R] \times X). \tag{3.2.7}$$

Notice that this definition does not depend on the point $\bar{x}$ w.r.t. which $\mathfrak{q}$ is defined. Observe that $\pi^{[0,R]}_\sharp \boldsymbol{\beta}_\lambda \in \mathcal{M}_+([0, R])$ with mass bounded by R^p and $\pi^X_\sharp \boldsymbol{\beta}_\lambda = \mu_\lambda$. Then we can apply Theorem 3.2.3 to $(\boldsymbol{\beta}_\lambda)_{\lambda\in\mathbb{L}}$ and obtain that, up to passing to a subnet, there exists $\boldsymbol{\beta} \in \mathcal{M}_+([0, R] \times X)$ s.t. $\lim_{\lambda\in\mathbb{L}} \boldsymbol{\beta}_\lambda = \boldsymbol{\beta}$. Now we define

$$O_n := \left\{ [x, r] \in \mathfrak{C}[X] \mid 0 \leqslant r \leqslant \frac{1}{n} \right\}, \quad n \geqslant 1$$

and, for every $n \geqslant 1$, the nets of real numbers

$$m_{\lambda,n} := \boldsymbol{\alpha}_\lambda(O_n).$$

Observe that $0 \leqslant m_{\lambda,n} \leqslant 1$ for every $n \geqslant 1$ and $\lambda \in \mathbb{L}$ then, up to passing to a subnet, they converge in $\lambda \in \mathbb{L}$ to some $m_n \in [0, 1]$. Define then $m := \inf_{n\geqslant 1} m_n$. We claim then that

$$\lim_{\lambda\in\mathbb{L}} \boldsymbol{\alpha}_\lambda = \frac{1}{r^p} \mathfrak{p}_\sharp \boldsymbol{\beta} + m\delta_{\mathfrak{o}} =: \boldsymbol{\alpha}.$$

Take any $\Omega \subset \mathfrak{C}[X]$ open; if $\mathfrak{o} \notin \Omega$, we have

$$\begin{aligned}\liminf_{\lambda\in\mathbb{L}} \alpha_\lambda(\Omega) &= \liminf_{\lambda\in\mathbb{L}} \int_{[0,R]\times X} (\chi_\Omega \circ \mathfrak{q})(x, r) \frac{1}{(r^p \circ \mathfrak{q})(x, r)} \, d\boldsymbol{\beta}_\lambda(x, r) \\ &\geqslant \int_{[0,R]\times X} (\chi_\Omega \circ \mathfrak{q})(x, r) \frac{1}{(r^p \circ \mathfrak{q})(x, r)} \, d\boldsymbol{\beta}(x, r) \\ &= \boldsymbol{\alpha}(\Omega),\end{aligned}$$

since everything is bounded, staying away from $\mathfrak{o}$. If, on the other hand, $\mathfrak{o} \in \Omega$, we have that $O_N \subset \Omega$ for some $N \geqslant 1$; calling $\Omega_n := \Omega \setminus O_n$ (which is an open set), we have, for every $n \geqslant N$, that

$$\begin{aligned}\liminf_{\lambda\in\mathbb{L}} \boldsymbol{\alpha}_\lambda(\Omega) &\geqslant \liminf_{\lambda\in\mathbb{L}} \boldsymbol{\alpha}_\lambda(O_n) + \liminf_{\lambda\in\mathbb{L}} \boldsymbol{\alpha}_\lambda(\Omega_n) \\ &\geqslant \boldsymbol{\alpha}(\Omega_n) + m_n.\end{aligned}$$

Now we pass to the limit as $n \to +\infty$ and, using the monotone convergence theorem and the fact that $\Omega_n \uparrow \Omega \setminus \{o\}$, we obtain

$$\liminf_{\lambda \in \mathbb{L}} \alpha_\lambda(\Omega) \geqslant \alpha(\Omega \setminus \{o\}) + m = \alpha(\Omega),$$

and this concludes the proof thanks to Portmanteau theorem (see e.g. [20, Corollary 8.2.10]. □

3.2.1 *Functions on the product cone*

In this subsection X_1 and X_2 are completely regular spaces.

Definition 3.2.7. Let $H : \mathfrak{C}[X_1, X_2] \to [0, +\infty]$ be a function. For every $(x_1, x_2) \in X_1 \times X_2$ we define

$$\begin{aligned} H_{x_1,x_2} : \mathbb{R}^2_+ &\to [0, +\infty] \\ (r_1, r_2) &\mapsto H([x_1, r_1], [x_2, r_2]). \end{aligned}$$

We say that

- H is 1-*homogeneous* if H_{x_1,x_2} is 1-homogeneous for every $(x_1, x_2) \in X_1 \times X_2$ i.e.

$$H_{x_1,x_2}(\lambda r_1, \lambda r_2) = \lambda H_{x_1,x_2}(r_1, r_2) \quad \text{for every } \lambda \geqslant 0,\ (r_1, r_2) \in \mathbb{R}^2_+;$$

- H is *convex* if H_{x_1,x_2} is convex for every $(x_1, x_2) \in X_1 \times X_2$.

Finally we define the functions $\operatorname{co}(H), \operatorname{cof}(H), \overline{\operatorname{co}}(H) : \mathfrak{C}[X_1, X_2] \to [0, +\infty]$ as

$$\begin{aligned} \operatorname{co}(H)([x_1, r_1], [x_2, r_2]) &:= \operatorname{co}(H_{x_1,x_2})(r_1, r_2), \\ \operatorname{cof}(H)([x_1, r_1], [x_2, r_2]) &:= \overline{\operatorname{co}}(H_{x_1,x_2})(r_1, r_2), \\ \overline{\operatorname{co}}(H) &:= \Gamma \operatorname{cof}(H) = \Gamma \operatorname{co}(H). \end{aligned}$$

Remark 3.2.8. Notice that, in particular, if H is 1-homogeneous, then $H(o, o) = 0$, since we adopt the convention that $0 \cdot \infty = 0$.

Remark 3.2.9. We will show in Corollary 4.2.6 that, if H is lower semicontinuous, then $\operatorname{cof}(H) = \overline{\operatorname{co}}(H)$ and a useful representation formula will be presented.

Lemma 3.2.10. *Let X_i, $i = 1, 2$ be completely regular spaces, let $H : \mathfrak{C}[X_1, X_2] \to [0, +\infty]$ be a 1-homogeneous Borel function and let $(\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$. Then*

$$\begin{aligned} &\inf\left\{\int_{\mathfrak{C}[X_1,X_2]} H \, d\alpha \mid \alpha \in \mathfrak{H}^1(\mu_1, \mu_2)\right\} \\ &= \inf\left\{\int_{\mathfrak{C}[X_1,X_2]} H \, d\alpha \mid \alpha \in \mathfrak{H}^1(\mu_1, \mu_2) \cap \mathcal{P}(\mathfrak{C}_{R(\mu_1,\mu_2)}[X_1, X_2])\right\}, \end{aligned}$$

where

$$R(\mu_1, \mu_2) := \mu_1(X_1) + \mu_2(X_2). \tag{3.2.8}$$

Proof. It is of course enough to prove the $\geqslant$ inequality. If $\boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2)$ we can assume that $\boldsymbol{\alpha}(\{\mathfrak{o}, \mathfrak{o}\}) = 0$, since $H(\mathfrak{o}, \mathfrak{o}) = 0$. By Lemma 3.2.2, we have that

$$\tilde{\boldsymbol{\alpha}} := \operatorname{dil}_{\vartheta_{\alpha,1},1}(\boldsymbol{\alpha}) \in \mathfrak{H}^1(\mu_1, \mu_2) \cap \mathcal{P}(\mathfrak{C}_{R(\mu_1,\mu_2)}[X_1, X_2])$$

and

$$\int_{\mathfrak{C}[X_1,X_2]} H\, d\boldsymbol{\alpha} = \int_{\mathfrak{C}[X_1,X_2]} H\, d\tilde{\boldsymbol{\alpha}}.$$

This concludes the proof. □

3.3 EXAMPLES OF COST FUNCTIONS

The aim of this section is to present some examples of cost functions $H : \mathfrak{C}[X_1, X_2] \to [0, +\infty]$, where X_1 and X_2 are completely regular spaces, satisfying (some of) the hypotheses we will assume throughout the rest of Part I.

3.3.1 *Mass-space product costs*

We consider cost functions of the form

$$H([x_1, r_1], [x_2, r_2]) := H_+(r_1, r_2) + H_-(r_1, r_2)c(x_1, x_2),$$

where $H_+, H_- : \mathbb{R}^2_+ \to \mathbb{R}$ are convex, 1-homogeneous and continuous and $c : X_1 \times X_2 \to [0, +\infty)$ is a continuous function satisfying

$$H_+(r_1, r_2) \geqslant -H_-(r_1, r_2) \sup_{(x_1,x_2) \in X_1 \times X_2} c(x_1, x_2) \quad \text{for every } r_1, r_2 \geqslant 0.$$

Possible choices of H_+ and H_- are given by e.g.

1. $m_p(r_1, r_2) := \left(\frac{1}{2}(r_1^p + r_2^p)\right)^{\frac{1}{p}}, \quad p \in [1, +\infty),$
2. $m_p(r_1, r_2) := -m_{-p}(r_1, r_2), \quad p \in (-\infty, 0) \cup (0, 1),$
3. $m_\infty(r_1, r_2) = r_1 \vee r_2, \quad m_{-\infty}(r_1, r_2) = r_1 \wedge r_2, \quad m_0 = \sqrt{r_1 r_2},$
4. $|r_1^\alpha - r_2^\alpha|^{1/\alpha}, \quad |r_1^\alpha + r_2^\alpha|^{1/\alpha}, \quad 0 < \alpha \leqslant 1,$

3.3.2 *Homogeneous marginal perspective functional*

Following [76, Section 5] we can build H starting from two *entropy functions* $F_i : X_i \to [0, +\infty]$, $i = 1, 2$ and a proper and lower semicontinuous cost function $c : X_1 \times X_2 \to [0, +\infty]$. Assuming that each F_i, $i = 1, 2$ is convex lower semicontinuous and finite in at least one positive point, we can define, for every number $c \in [0, +\infty]$, the function $H_c : \mathbb{R}^2_+ \to [0, +\infty]$, as the lower semicontinuous envelope of

$$\tilde{H}_c(r_1, r_2) := \begin{cases} \inf_{\theta > 0} \{r_1 F_1(\theta/r_1) + r_2 F_2(\theta/r_2) + \theta c\}, & \text{if } c \in [0, +\infty), \\ F_1(0) r_1 + F_2(0) r_2 & \text{if } c = +\infty, \end{cases}$$

for $r_1, r_2 \in \mathbb{R}^2_+$. The function $H : \mathfrak{C}[X_1, X_2] \to [0, +\infty]$ is then defined as

$$H([x_1, r_1], [x_2, r_2]) := H_{c(x_1,x_2)}(r_1, r_2), \quad ([x_1, r_1], [x_2, r_2]) \in \mathfrak{C}[X_1, X_2].$$

Such function H is convex and 1-homogeneous (see [76, Lemma 5.3]. Possible choices (see e.g. [42, 75]) for F_i are given by

1. Power like entropies: for $p \in \mathbb{R}$ we define

$$U_p(s) := \begin{cases} \frac{1}{p(p-1)}\left(s^p - p(s-1) - 1\right) & \text{if } p \neq 0, 1, \\ s \log s - s + 1, & \text{if } p = 1, \\ s - 1 - \log s, & \text{if } p = 0, \end{cases}$$

   for $s > 0$, with $U_p(0) = 1/p$ if $p > 0$ and $U_p(0) = +\infty$ if $p \leqslant 0$.

2. Indicator functions: for numbers $0 \leqslant a \leqslant 1 \leqslant b \leqslant +\infty$ we define

$$I_{[a,b]}(s) := \begin{cases} 0 & \text{if } s \in [a, b], \\ +\infty & \text{if } s \notin [a, b]. \end{cases}$$

3. χ^α divergences: for a parameter $\alpha \geqslant 1$ we define

$$\chi^\alpha(s) := |s - 1|^\alpha, \quad s \in \mathbb{R}.$$

4

CONVEXIFICATION AND DUALITY

In this chapter we present the convexification/relaxation procedure that allows to connect the cost on Dirac masses to the one defined on general non-negative measures. Section 4.1 presents the general setting and shows the equality between the relaxation $\overline{\text{co}}\,(\mathscr{S}_H)$ and the primal formulation-cost $\mathscr{U}_H$; Section 4.2 is devoted to the duality theorem and the application of the concept of adapted algebra of continuous functions; Section 4.2 treats the case of merely Hausdorff spaces, where the notion of narrow topology has to be refined. If not stated otherwise, in this section X_1 and X_2 are completely regular spaces.

This Chapter is the result of a collaboration with Giuseppe Savaré and Section 4.3 appeared in [109].

4.1 LOWER SEMICONTINUITY OF THE TRANSPORT COST AND CONVEXIFICATION

Definition 4.1.1. Let $H : \mathfrak{C}[X_1, X_2] \to [0, +\infty]$ be a proper Borel function. We define the *singular cost* $\mathscr{S}_H : \mathcal{M}(X_1) \times \mathcal{M}(X_2) \to [0, +\infty]$ as

$$\mathscr{S}_H(\mu_1, \mu_2) := \begin{cases} H([x_1, r_1]; [x_2, r_2]) & \text{if} \quad \begin{array}{l} \mu_1 = r_1\delta_{x_1},\ \mu_2 = r_2\delta_{x_2}, \\ x_1 \in X_1, x_2 \in X_2,\ r_1, r_2 \geqslant 0, \end{array} \\ +\infty & \text{elsewhere.} \end{cases}$$

and the *unbalanced Optimal Transport cost* $\mathscr{U}_H : \mathcal{M}(X_1) \times \mathcal{M}(X_2) \to [0, +\infty]$ as

$$\mathscr{U}_H(\mu_1, \mu_2) := \inf \left\{ \int_{\mathfrak{C}[X_1, X_2]} H \, d\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2) \right\}$$

if $(\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$ and equal to $+\infty$ elsewhere.

The aim of this section is to study the relations between $\mathscr{S}_H$ and $\mathscr{U}_H$; in particular we are interested in studying the lower semicontinuous and convex relaxation of $\mathscr{S}_H$. With this in mind, we have the following remark.

Remark 4.1.2. If $H : \mathfrak{C}[X_1, X_2] \to [0, +\infty]$ is a function, then

$$\Gamma\mathscr{S}_H = \mathscr{S}_{\Gamma H}. \tag{4.1.1}$$

Indeed, both are equal to $+\infty$ outside the closed set $\Delta_+(X_1) \times \Delta_+(X_2)$ and the equality on $\Delta_+(X_1) \times \Delta_+(X_2)$ follows by Lemma 3.2.1.

For this reason and to exploit Lemma 3.2.10, we will assume for the rest of this section that

$$H : \mathfrak{C}[X_1, X_2] \to [0, +\infty] \text{ is a proper, 1-homogeneous and lower semicontinuous function.} \tag{4.1.2}$$

In the following result we prove that $\mathscr{U}_H$ is a lower semicontinuous convex function.

Proposition 4.1.3. *Let* H *be as in* (4.1.2) *and let* $\mathscr{U}_H$ *be as in Definition 4.1.1. Then, for every* $(\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$, *there exists a probability measure* $\boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2) \cap \mathcal{P}(\mathfrak{C}_{R(\mu_1,\mu_2)}[X_1, X_2])$, *where* $R(\mu_1, \mu_2)$ *is as in* (3.2.8), *such that*

$$\mathscr{U}_H(\mu_1, \mu_2) = \int_{\mathfrak{C}[X_1,X_2]} H \, d\boldsymbol{\alpha}.$$

Moreover $\mathscr{U}_H$ *is a lower semicontinuous convex function such that*

$$\mathscr{U}_H(r_1\delta_{x_1}, r_2\delta_{x_2}) \leqslant H([x_1, r_1]; [x_2, r_2]) \tag{4.1.3}$$

for every $(x_1, x_2) \in X_1 \times X_2$ *and every* $(r_1, r_2) \in \mathbb{R}^2_+$. *If, in addition,* H *is also convex, then* (4.1.3) *is an equality.*

Proof. Let $(\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$; by Lemma 3.2.10, it holds

$$\mathscr{U}_H(\mu_1, \mu_2) = \inf\left\{\int_{\mathfrak{C}[X_1,X_2]} H \, d\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2) \cap \mathcal{P}(\mathfrak{C}_{R(\mu_1,\mu_2)}[X_1, X_2])\right\}. \tag{4.1.4}$$

Thanks to Lemma 3.2.6, we have that $\mathfrak{H}^1(\mu_1, \mu_2) \cap \mathcal{P}(\mathfrak{C}_{R(\mu_1,\mu_2)}[X_1, X_2])$ is compact and the lower semicontinuity of H gives that the functional

$$\boldsymbol{\alpha} \mapsto \int_{\mathfrak{C}[X_1,X_2]} H \, d\boldsymbol{\alpha} \tag{4.1.5}$$

is lower semicontinuous. We can thus conclude that a minimizer exists by the direct method in Calculus of Variations.

The convexity of $\mathscr{U}_H$ follows by the convexity of the constraints: if $(\mu_1^1, \mu_2^1), (\mu_1^2, \mu_2^2)$ are measures in $\mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$ and $\gamma, \beta \in [0, 1]$ s.t $\gamma + \beta = 1$, we can take $\boldsymbol{\alpha}_1 \in \mathfrak{H}^1(\mu_1^1, \mu_2^1)$ and $\boldsymbol{\alpha}_2 \in \mathfrak{H}^1(\mu_1^2, \mu_2^2)$ such that

$$H(\mu_1^1, \mu_2^1) = \int_{\mathfrak{C}[X_1,X_2]} H \, d\boldsymbol{\alpha}_1, \quad H(\mu_1^2, \mu_2^2) = \int_{\mathfrak{C}[X_1,X_2]} H \, d\boldsymbol{\alpha}_2.$$

It is then enough to observe that $\boldsymbol{\alpha} := \gamma\boldsymbol{\alpha}_1 + \beta\boldsymbol{\alpha}_2 \in \mathfrak{H}^1(\gamma\mu_1^1 + \beta\mu_1^2, \gamma\mu_2^1 + \beta\mu_2^2)$.

The lower semicontinuity of $\mathscr{U}_H$ is a consequence of Lemma 3.2.10: if $\{(\mu_1^\lambda, \mu_2^\lambda)\}_{\lambda\in\mathbb{L}}$ is a net in $\mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$ converging to $(\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$, we can consider, for every $\lambda \in \mathbb{L}$, some $\boldsymbol{\alpha}_\lambda \in \mathfrak{H}^1(\mu_1^\lambda, \mu_2^\lambda) \cap \mathcal{P}(\mathfrak{C}_{R(\mu_1^\lambda,\mu_2^\lambda)}[X_1, X_2])$ such that

$$\mathscr{U}_H(\mu_1^\lambda, \mu_2^\lambda) = \int_{\mathfrak{C}[X_1,X_2]} H \, d\boldsymbol{\alpha}_\lambda.$$

Since μ_i^λ are converging for $i = 1, 2$ and $R(\mu_1^\lambda, \mu_2^\lambda)$ is bounded from above, we can use Lemma 3.2.6 to extract a convergent subnet of $(\boldsymbol{\alpha}_\lambda)_{\lambda\in\mathbb{L}}$ with limit

$\alpha \in \mathfrak{H}^1(\mu_1, \mu_2)$. Using again the lower semicontinuity of the functional in (4.1.5), we can conclude that $\mathscr{U}_H$ is lower semicontinuous.

(4.1.3) follows by the fact that

$$\alpha = \delta_{[x_1,r_1]} \otimes \delta_{[x_2,r_2]}$$

is an element of $\mathfrak{H}^1(r_1\delta_{x_1}, r_2\delta_{x_2})$.

If, in addition, H is convex, we can take $\alpha \in \mathfrak{H}^1(r_1\delta_{x_1}, r_2\delta_{x_2})$ such that

$$\mathscr{U}_H(r_1\delta_{x_1}, r_2\delta_{x_2}) = \int_{\mathfrak{C}[X_1,X_2]} H\, d\alpha$$

and we observe that α is concentrated on

$$\{\lambda_1[x_1, 1], \lambda_2[x_2, 1] \mid \lambda_1, \lambda_2 \geqslant 0\}$$

with

$$\int_{\mathfrak{C}[X_1,X_2]} \mathsf{r}_1\, d\alpha = r_1, \quad \int_{\mathfrak{C}[X_1,X_2]} \mathsf{r}_2\, d\alpha = r_2.$$

Hence, using Jensen's inequality and the convexity of H_{x_1,x_2}, we have

$$\begin{aligned}
\mathscr{U}_H(r_1\delta_{x_1}, r_2\delta_{x_2}) &= \int_{\mathfrak{C}[X_1,X_2]} H\, d\alpha \\
&= \int_{\mathbb{R}^2_+} H_{x_1,x_2}\, d(\mathsf{r}_1, \mathsf{r}_2)_\sharp \alpha \\
&\geqslant H_{x_1,x_2}\left(\int_{\mathbb{R}_+} r_1\, d(\mathsf{r}_1)_\sharp \alpha(r_1), \int_{\mathbb{R}_+} r_2\, d(\mathsf{r}_2)_\sharp \alpha(r_2)\right) \\
&= H_{x_1,x_2}(r_1, r_2) \\
&= H([x_1, r_1], [x_2, r_2]).
\end{aligned}$$

□

Thanks to Proposition 4.1.3, given $\mu_1 \in \mathcal{M}_+(X_1)$ and $\mu_2 \in \mathcal{M}_+(X_2)$, the set

$$\mathfrak{H}^1_H(\mu_1, \mu_2) := \left\{\alpha \in \mathfrak{H}^1(\mu_1, \mu_2) \mid \int_{\mathfrak{C}[X_1,X_2]} H\, d\alpha = \mathscr{U}_H(\mu_1, \mu_2)\right\} \tag{4.1.6}$$

is not empty. Elements of $\mathfrak{H}^1_H(\mu_1, \mu_2)$ are called *Unbalanced Optimal Transport plans*.

In the following result we prove that $\mathscr{U}_H$ is the lower semicontinuous convex envelope of $\mathscr{S}_H$ in a direct way. Another proof is given in Theorem 4.2.4.

Theorem 4.1.4. *Let* H *be as in* (4.1.2) *and let* $\mathscr{S}_H$ *and* $\mathscr{U}_H$ *be as in Definition 4.1.1. Then*

$$\overline{\text{co}}\,(\mathscr{S}_H) = \mathscr{U}_H. \tag{4.1.7}$$

Proof. We only need to prove equality on $\mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$ being both functions equal to $+\infty$ outside it. First of all let us compute $\mathrm{co}\,(\mathscr{S}_H)$ on the set $\mathrm{co}\,(\Delta_+(X_1) \times \Delta_+(X_2)) = \mathrm{co}\,(\Delta_+(X_1)) \times \mathrm{co}\,(\Delta_+(X_2))$. If we take an element (μ_1, μ_2) of $\mathrm{co}\,(\Delta_+(X_1)) \times \mathrm{co}\,(\Delta_+(X_2))$, then we can write

$$\mathrm{co}\,(\mathscr{S}_H)\,(\mu_1, \mu_2) = \inf\Bigg\{ \sum_{ij} \alpha_{ij} \mathscr{S}_H(\mu_1^i, \mu_2^j) \mid (\mu_1, \mu_2) = \sum_{ij} \alpha_{ij}(\mu_1^i, \mu_2^j), \\ \alpha_{ij} \geqslant 0, \sum_{ij} \alpha_{ij} = 1,\ \mu_k^i \in \Delta_+(X_k) \Bigg\}.$$

Since $(\mu_1, \mu_2) \in \mathrm{co}\,(\Delta_+(X_1)) \times \mathrm{co}\,(\Delta_+(X_2))$ it holds

$$\mu_1 = \sum_h m_h \delta_{x_1^h} \qquad \mu_2 = \sum_k n_k \delta_{x_2^k}$$

for some $m_h, n_k \in \mathbb{R}_+$ and some $x_1^h \in X_1$, $x_2^k \in X_2$ distinct points. Then it holds that

$$\mu_1 = \sum_h m_h \delta_{x_1^h} = \sum_{ij} \alpha_{ij} \mu_1^i = \sum_{ij} \alpha_{ij} r_1^i \delta_{\tilde{x}_1^i},$$
$$\mu_2 = \sum_k n_k \delta_{x_2^k} = \sum_{ij} \alpha_{ij} \mu_2^j = \sum_{ij} \alpha_{ij} r_2^j \delta_{\tilde{x}_2^j}$$

for some $r_1^i, r_2^j \in \mathbb{R}_+$ and some $\tilde{x}_1^i \in X_1$, $\tilde{x}_2^j \in X_2$ not necessarily distinct points. Clearly it must be

$$\cup_h \{x_1^h\}_h = \cup_i \{\tilde{x}_1^i\}_i, \quad \cup_k \{x_1^k\}_k = \cup_j \{\tilde{x}_2^j\}_j.$$

We can then group the r_1^i and the α_{ij} as follows

$$\{r_1^{ph}\}_p := \{r_1^i \mid \tilde{x}_1^i = x_1^h\}, \quad \{\alpha_{pjh}\}_p := \{\alpha_{ij} \mid \tilde{x}_1^i = x_1^h\}$$

obtaining

$$\mu_1 = \sum_h m_h \delta_{x_1^h} = \sum_{pjh} \alpha_{pjh} r_i^{ph} \delta_{x_1^h},$$
$$\mu_2 = \sum_k n_k \delta_{x_2^k} = \sum_{pjh} \alpha_{pjh} r_2^j \delta_{\tilde{x}_2^j}.$$

Analogously we can group the r_2^j and again the α_{pjh} as follows

$$\{r_2^{qk}\}_q := \{r_2^j \mid \tilde{x}_2^j = x_2^k\}, \quad \{\alpha_{pqhk}\}_q := \{\alpha_{pjh} \mid \tilde{x}_2^j = x_2^k\}$$

obtaining

$$\mu_1 = \sum_h m_h \delta_{x_1^h} = \sum_{pqhk} \alpha_{pqhk} r_i^{pqhk} \delta_{x_1^h},$$
$$\mu_2 = \sum_k n_k \delta_{x_2^k} = \sum_{pqhk} \alpha_{pqhk} r_2^{qk} \delta_{x_2^k}.$$

We are then left with the compatibility conditions

$$\begin{cases} m_h = \sum_{pqk} \alpha_{pqhk} r_1^{ph} & \text{for every} h, \\ n_k = \sum_{pqh} \alpha_{pqhk} r_2^{qk} & \text{for every } k, \\ \sum_{pqhk} \alpha_{pqhk} = 1 \end{cases} \tag{4.1.8}$$

The convex envelope of $\mathscr{S}_H$ can be then written in two arbitrary points

$$(\mu_1, \mu_2) = \left(\sum_h m_h \delta_{x_1^h}, \sum_k n_k \delta_{x_2^k} \right) \in \operatorname{co}(\Delta_+(X_1)) \times \operatorname{co}(\Delta_+(X_2))$$

as

$$\begin{aligned} \operatorname{co}(\mathscr{S}_H)(\mu_1, \mu_2) = \inf \Bigg\{ & \sum_{pqhk} \alpha_{pqhk} H([x_1^h, r_1^{ph}], [x_2^k, r_2^{pk}]) \text{ s.t.} \\ & \{\alpha_{pqhk}\}_{pqhk}, \{r_1^{ph}\}_{ph}, \{r_2^{qk}\}_{qk} \subset \mathbb{R}_+ \\ & \text{and (4.1.8) holds} \Bigg\}. \end{aligned}$$

This formula tells us that $\operatorname{co}(\mathscr{S}_H)$ can be written as

$$\operatorname{co}(\mathscr{S}_H)(\mu_1, \mu_2) = \inf \left\{ \int_{\mathfrak{C}[X_1, X_2]} H \, d\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2) \cap \text{P-Discr}_+(\mathfrak{C}[X_1, X_2]) \right\},$$

where $\text{P-Discr}_+(\mathfrak{C}[X_1, X_2]) = \text{Discr}_+(\mathfrak{C}[X_1, X_2]) \cap \mathcal{P}(\mathfrak{C}[X_1, X_2])$. Reasoning as in the proof of Lemma 3.2.10, we have that

$$\operatorname{co}(\mathscr{S}_H)(\mu_1, \mu_2) = \inf \left\{ \int_{\mathfrak{C}[X_1, X_2]} H \, d\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2) \cap \text{P-Discr}_+(\mathfrak{C}_*) \right\},$$

where $\mathfrak{C}_* := \mathfrak{C}_{R(\mu_1, \mu_2)}[X_1, X_2]$. Thus

$$\mathscr{U}_H(\mu_1, \mu_2) \leqslant \operatorname{co}(\mathscr{S}_H)(\mu_1, \mu_2) \quad \text{for every } (\mu_1, \mu_2) \in \operatorname{co}(\Delta_+(X_1)) \times \operatorname{co}(\Delta_+(X_2)).$$

Moreover $\mathscr{U}_H$ is lower semicontinuous and convex hence, by definition of $\overline{\operatorname{co}}(\mathscr{S}_H)$, it must hold

$$\overline{\operatorname{co}}(\mathscr{S}_H)(\mu_1, \mu_2) \geqslant \mathscr{U}_H(\mu_1, \mu_2) \quad \text{for every } (\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_1).$$

Then, in order to prove equality, we only need to prove the other inequality. To do so, fixed $(\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$, we prove that there exists a net $\{(\mu_1^\eta, \mu_2^\eta)\}_{\eta \in \mathbb{E}} \subset \operatorname{co}(\Delta_+(X_1)) \times \operatorname{co}(\Delta_+(X_2))$ s.t. $\lim_\eta (\mu_1^\eta, \mu_2^\eta) = (\mu_1, \mu_2)$ and a net

$$\{\boldsymbol{\gamma}_\eta\}_{\eta \in \mathbb{E}} \subset \text{Discr}_+(\mathfrak{C}_*) \cap \mathcal{P}(\mathfrak{C}_*)$$

s.t. $\boldsymbol{\gamma}_\eta \in \mathfrak{H}^1(\mu_1^\eta, \mu_2^\eta)$ for every $\eta \in \mathbb{E}$ satisfying

$$\lim_{\eta \in \mathbb{E}} \int_{\mathfrak{C}[X_1, X_2]} H \, d\boldsymbol{\gamma}_\eta = \int_{\mathfrak{C}[X_1, X_2]} H \, d\boldsymbol{\alpha}^*$$

where $\boldsymbol{\alpha}^* \in \mathfrak{H}^1_H(\mu_1, \mu_2) \cap \mathcal{P}(\mathfrak{C}_*)$. If we are able to do so, we conclude, indeed

$$\begin{aligned}\overline{\mathrm{co}}\,(\mathscr{S}_H)\,(\mu_1, \mu_2) = \inf\Big\{ &\liminf_\lambda \mathrm{co}\,(\mathscr{S}_H)\,(\mu_1^\lambda, \mu_2^\lambda) \text{ with} \\ &\{(\mu_1^\lambda, \mu_2^\lambda)\}_{\lambda\in\mathbb{L}} \subset \mathrm{co}\,(\Delta_+(X_1) \times \Delta_+(X_2)) \text{ and} \\ &(\mu_1, \mu_2) = \lim_\lambda (\mu_1^\lambda, \mu_2^\lambda) \Big\} \\ &\leqslant \liminf_\eta \mathrm{co}\,(\mathscr{S}_H)\,(\mu_1^\eta, \mu_2^\eta) \\ &\leqslant \liminf_\eta \int_{\mathfrak{C}[X_1,X_2]} H \, d\boldsymbol{\gamma}_\eta \\ &= \liminf_\eta \int_{\mathfrak{C}[X_1,X_2]} H \, d\boldsymbol{\gamma}_\eta \\ &= \int_{\mathfrak{C}[X_1,X_2]} H \, d\boldsymbol{\alpha}^* \\ &= \mathscr{U}_H(\mu_1, \mu_2).\end{aligned}$$

To do so, we use Lemma 3.1.4 with $X := \mathfrak{C}_*$, $f = H$, $\alpha := \boldsymbol{\alpha}^*$ and we find $\{\boldsymbol{\gamma}_\eta\}_{\eta\in\mathbb{E}} \subset \mathrm{Discr}_+(\mathfrak{C}_*) \cap \mathcal{P}(\mathfrak{C}_*)$ s.t.

$$\lim_{\eta\in\mathbb{E}} \boldsymbol{\gamma}_\eta = \boldsymbol{\alpha}^*, \quad \lim_{\eta\in\mathbb{E}} \int_{\mathfrak{C}_*} H \, d\boldsymbol{\gamma}_\eta = \int_{\mathfrak{C}_*} H \, d\boldsymbol{\alpha}^*.$$

Finally we can define

$$\mu_1^\eta := \mathfrak{h}_1^1(\boldsymbol{\gamma}_\eta), \quad \mu_2^\eta := \mathfrak{h}_2^1(\boldsymbol{\gamma}_\eta) \qquad \text{for every } \eta \in \mathbb{E}.$$

Obviously $\boldsymbol{\gamma}_\eta \in \mathfrak{H}^1(\mu_1^\eta, \mu_2^\eta)$ and $\mu_i = \lim_{\eta\in\mathbb{E}} \mu_i^\eta$, indeed if $\varphi_i \in C_b(X_i)$, then

$$\begin{aligned}\lim_{\eta\in\mathbb{E}} \int_{X_i} \varphi_i \, d\mu_i^\eta &= \lim_{\eta\in\mathbb{E}} \int_{X_i} \varphi_i \, d\mathfrak{h}_i^1(\boldsymbol{\gamma}_\eta) = \lim_{\eta\in\mathbb{E}} \int_{\mathfrak{C}[X_1,X_2]} (\varphi \circ \mathbf{x}_i)\mathbf{r}_i \, d\boldsymbol{\gamma}_\eta \\ &= \lim_{\eta\in\mathbb{E}} \int_{\mathfrak{C}_*} (\varphi_i \circ \mathbf{x}_i)\mathbf{r}_i \, d\boldsymbol{\gamma}_\eta = \int_{\mathfrak{C}_*} (\varphi_i \circ \mathbf{x}_i)\mathbf{r}_i \, d\boldsymbol{\alpha}^* \\ &= \int_{\mathfrak{C}[X_1,X_2]} (\varphi_i \circ \mathbf{x}_i)\mathbf{r}_i \, d\boldsymbol{\alpha}^* = \int_{X_i} \varphi_i \, d\mathfrak{h}_i^1(\boldsymbol{\alpha}^*) \\ &= \int_{X_i} \varphi_i \, d\mu_i,\end{aligned}$$

where we have used that $(\varphi \circ \mathbf{x}_i)\mathbf{r}_i \in C_b(\mathfrak{C}_*)$ and the convergence of $\boldsymbol{\gamma}_\eta$ to $\boldsymbol{\alpha}^*$ in $\mathcal{P}(\mathfrak{C}_*)$. Notice that, in general, it is not true that $(\varphi \circ \mathbf{x}_i)\mathbf{r}_i \in C_b(\mathfrak{C}[X_1, X_2])$. □

4.2 DUALITY

In this section, we still assume that H is as in (4.1.2) and we study the dual formulation of the definition of $\mathscr{U}_H$.

Definition 4.2.1. We define the following set of continuous functions

$$\Phi_H := \left\{ \begin{array}{l} (\varphi_1, \varphi_2) \in C_b(X_1) \times C_b(X_2) \text{ s.t.} \\ \varphi_1(x_1)r_1 + \varphi_2(x_2)r_2 \leqslant H([x_1, r_1], [x_2, r_2]) \\ \text{for every } (x_1, x_2) \in X_1 \times X_2,\ r_1, r_2 \geqslant 0 \end{array} \right\} \tag{4.2.1}$$

and, for every $(\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$, the functional $\mathcal{D}(\cdot; \mu_1, \mu_2) : C_b(X_1) \times C_b(X_2) \to \mathbb{R}$ given by

$$\mathcal{D}(\varphi_1, \varphi_2; \mu_1, \mu_2) := \int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2, \quad (\varphi_1, \varphi_2) \in C_b(X_1) \times C_b(X_2). \tag{4.2.2}$$

Before stating the main duality result, let us briefly recall the Fenchel-Moreau Theorem in the framework of a pair of vector spaces E, F placed in duality by a nondegenerate bilinear map $\langle \cdot, \cdot, \rangle$, see e.g. [46]. We endow E with the weak topology $\sigma(E, F)$, the coarsest topology for which all the functions $e \mapsto \langle e, f \rangle$, $f \in F$, are continuous.

Definition 4.2.2. Let $\mathscr{F} : E \to (-\infty, +\infty]$ be not identically $+\infty$ and satisfying

$$\mathscr{F}(e) \geqslant \langle e, f \rangle - c \quad \text{for some } f \in F,\, c \in \mathbb{R} \text{ and every } e \in E. \tag{4.2.3}$$

The *polar* (or conjugate) function of $\mathscr{F}$ is the function $\mathscr{F}^* : F \to (-\infty, +\infty]$ defined by

$$\mathscr{F}^*(f) := \sup_{e \in E} \langle e, f \rangle - \mathscr{F}(e) \quad \text{for every } f \in F.$$

Theorem 4.2.3 (Fenchel-Moreau)**.** *Let E and F be vector spaces placed in duality and let $\mathscr{F} : E \to (-\infty, +\infty]$ be satisfying (4.2.3) and not identically $+\infty$. Then the lower semicontinuous convex envelope of $\mathscr{F}$ is given by the dual formula*

$$\overline{\text{co}}\,(\mathscr{F}) = \mathscr{F}^{**}(e) := \sup_{f \in F} \langle e, f \rangle - \mathscr{F}^*(f) \quad \textit{for every } e \in E.$$

In particular,

$$\textit{if } \mathscr{F} \textit{ is convex and lower semicontinuous then} \quad \mathscr{F} = \mathscr{F}^{**}.$$

Theorem 4.2.4. *Let H be as in (4.1.2), let $\mathscr{S}_H$ and $\mathscr{U}_H$ be as in Definition 4.1.1 and let $A_i \subset C_b(X_i)$ be adapted algebras of continuous functions as in Definition 3.1.5. Then*

$$\mathscr{U}_H(\mu_1, \mu_2) = \sup \{ \mathcal{D}(\varphi_1, \varphi_2; \mu_1, \mu_2) \mid (\varphi_1, \varphi_2) \in \Phi_H \cap (A_1 \times A_2) \} \tag{4.2.4}$$

for every $(\mu_1, \mu_2) \in \mathcal{M}_+(X_1) \times \mathcal{M}_+(X_2)$.

Proof. Set $E := \mathcal{M}(X_1) \times \mathcal{M}(X_2)$ and $F := A_1 \times A_2$ with the bilinear form

$$\langle \cdot, \cdot \rangle : E \times F \to \mathbb{R}, \qquad ((\mu_1, \mu_2), (\varphi_1, \varphi_2)) \mapsto \mathcal{D}(\varphi_1, \varphi_2; \mu_1, \mu_2).$$

This is a well defined nondegenerate bilinear form. We endow then E with the topology $\sigma(E, F)$ which coincides exactly with the product weak topology by Lemma 3.1.6.
Consider then the function $\mathscr{S}_H : E \mapsto (-\infty, +\infty]$ defined as in Definition 4.1.1. Using Theorem 4.1.4 and Theorem 4.2.3 we have that

$$\mathscr{U}_H = \overline{\text{co}}\,(\mathscr{S}_H) = \mathscr{S}_H^{**}.$$

Moreover

$$\begin{aligned}
\mathscr{S}_H^*(\varphi_1,\varphi_2) &= \sup_{(\mu_1,\mu_2)\in E} \{\langle(\mu_1,\mu_2),(\varphi_1,\varphi_2)\rangle - \mathscr{S}_H(\mu_1,\mu_2)\} \\
&= \sup_{(\mu_1,\mu_2)\in\Delta_+(X_1)\times\Delta_+(X_2)} \{\langle(\mu_1,\mu_2),(\varphi_1,\varphi_2)\rangle - \mathscr{S}_H(\mu_1,\mu_2)\} \\
&= \sup_{x_1,r_1,x_2,r_2} \{\varphi_1(x_1)r_1 + \varphi_2(x_2)r_2 - H(x_1,r_1;x_2,r_2)\} \\
&= \begin{cases} 0 & \text{if } (\varphi_1,\varphi_2)\in\Phi_H\cap(A_1\times A_2), \\ +\infty & \text{elsewhere.} \end{cases}
\end{aligned}$$

Hence

$$\mathscr{S}_H^{**}(\mu_1,\mu_2) = \sup\{\mathcal{D}(\varphi_1,\varphi_2;\mu_1,\mu_2) \mid (\varphi_1,\varphi_2)\in\Phi_H\cap(A_1\times A_2)\}.$$

□

Remark 4.2.5. We remark that equality (4.2.4) does not require Theorem 4.1.4. Indeed, from the proof of Theorem 4.2.4, we immediately have that

$$\overline{\mathrm{co}}\,(\mathscr{S}_H)\,(\mu_1,\mu_2) = \sup\left\{\int_{X_1}\varphi_1\,d\mu_1 + \int_{X_2}\varphi_2\,d\mu_2 \mid (\varphi_1,\varphi_2)\in\Phi_H\right\}. \quad (4.2.5)$$

In Proposition 4.1.3 we have proven that $\mathscr{U}_H$ is convex and lower semicontinuous and stays below $\mathscr{S}_H$, then $\mathscr{U}_H \leqslant \overline{\mathrm{co}}\,(\mathscr{S}_H)$. The other inequality is immediate: take any $\boldsymbol{\alpha}\in\mathfrak{H}^1(\mu_1,\mu_2)$ and any $(\varphi_1,\varphi_2)\in\Phi_H$; then

$$\begin{aligned}
\int_{X_1}\varphi_1\,d\mu_1 + \int_{X_2}\varphi_2\,d\mu_2 &= \int_{\mathfrak{C}[X_1,X_2]} ((\varphi_1\circ\mathsf{x}_1)\mathsf{r}_1 + (\varphi_2\circ\mathsf{x}_2)\mathsf{r}_2)\,d\boldsymbol{\alpha} \\
&\leqslant \int_{\mathfrak{C}[X_1,X_2]} H\,d\boldsymbol{\alpha}.
\end{aligned}$$

Passing to the supremum in Φ_H and to the infimum in $\mathfrak{H}^1(\mu_1,\mu_2)$ and using (4.2.5), we conclude that $\mathscr{U}_H \geqslant \overline{\mathrm{co}}\,(\mathscr{S}_H)$.

We conclude this section showing the equality $\overline{\mathrm{co}}\,(H) = \mathrm{cof}\,(H)$ as a consequence of the above result.

Corollary 4.2.6. *Let* H *be as in* (4.1.2) *and let* $\overline{\mathrm{co}}\,(H)$ *and* $\mathrm{cof}\,(H)$ *be as in Definition 3.2.7. Then*

$$\begin{aligned}
\overline{\mathrm{co}}\,(H)\,([x_1,r_1],[x_2,r_2]) &= \mathrm{cof}\,(H)\,([x_1,r_1],[x_2,r_2]) \\
&= \sup\{\varphi(x_1)r_1 + \varphi_2(x_2)r_2 \mid (\varphi_1,\varphi_2)\in\Phi_H\}
\end{aligned}$$

for every $(x_1,x_2)\in X_1\times X_2$ *and every* $(r_1,r_2)\in\mathbb{R}_+^2$*. Moreover*

$$\overline{\mathrm{co}}\,(\mathscr{S}_H) = \overline{\mathrm{co}}\left(\mathscr{S}_{\overline{\mathrm{co}}(H)}\right) = \mathscr{U}_H = \mathscr{U}_{\overline{\mathrm{co}}(H)}$$

Proof. We denote by $U_H : \mathfrak{C}[X_1,X_2] \to [0,+\infty]$ the restriction of $\mathscr{U}_H$ to $\Delta_+(X_1)\times\Delta_+(X_2)\cong\mathfrak{C}[X_1,X_2]$ (see Lemma 3.2.1).
It is clear that $\overline{\mathrm{co}}\,(H)\leqslant\mathrm{cof}\,(H)\leqslant H$ so that

$$\mathscr{U}_{\overline{\mathrm{co}}(H)} \leqslant \mathscr{U}_{\mathrm{cof}(H)} \leqslant \mathscr{U}_H$$

and

$$\overline{\text{co}}\,(H) = U_{\overline{\text{co}}(H)} \leqslant \text{cof}\,(H) = U_{\text{cof}(H)} \leqslant U_H \leqslant H,$$

where we used Proposition 4.1.3 and the convexity of $\overline{\text{co}}\,(H)$ and $\text{cof}\,(H)$. Moreover, since U_H is, lower semicontinuous, convex and stays below H, we have that $U_H \leqslant \overline{\text{co}}\,(H)$. This gives that

$$\overline{\text{co}}\,(H) = U_{\overline{\text{co}}(H)} = \text{cof}\,(H) = U_{\text{cof}(H)} = U_H$$

and in particular, using Theorem 4.2.4, that

$$\begin{aligned}\overline{\text{co}}\,(H)\,([x_1,r_1],[x_2,r_2]) &= \text{cof}\,(H)\,([x_1,r_1];[x_2,r_2]) \\ &= \sup\{\varphi(x_1)r_1 + \varphi_2(x_2)r_2 \mid (\varphi_1,\varphi_2) \in \Phi_H\}.\end{aligned}$$

The fact that $\overline{\text{co}}\,(H) = U_H$ gives that $\mathscr{U}_H \leqslant \mathscr{S}_{\overline{\text{co}}(H)}$ so that $\mathscr{U}_H \leqslant \overline{\text{co}}\left(\mathscr{S}_{\overline{\text{co}}(H)}\right)$. However, by Proposition 4.1.3, we know that $\mathscr{U}_H = \overline{\text{co}}\,(\mathscr{S}_H)$ so that $\overline{\text{co}}\,(\mathscr{S}_H) \leqslant \overline{\text{co}}\left(\mathscr{S}_{\overline{\text{co}}(H)}\right)$. Since, obviously, the other inequality holds, we have $\overline{\text{co}}\,(\mathscr{S}_H) = \overline{\text{co}}\left(\mathscr{S}_{\overline{\text{co}}(H)}\right)$. Applying again Proposition 4.1.3 to $\overline{\text{co}}\,(H)$ we conclude that

$$\overline{\text{co}}\,(\mathscr{S}_H) = \overline{\text{co}}\left(\mathscr{S}_{\overline{\text{co}}(H)}\right) = \mathscr{U}_H = \mathscr{U}_{\overline{\text{co}}(H)}.$$

□

4.3 THE CASE OF MERELY HAUSDORFF SPACES

In this last section we show how to generalize Theorem 3.2.3 and Corollary 3.2.4 to arbitrary Hausdorff topological spaces.

Since duality with continuous and bounded functions cannot be used to define a Hausdorff topology in $\mathcal{M}_+(X)$, a natural topology (called *narrow topology*) can be introduced following Topsoe [110, Appendix].

Definition 4.3.1. (Narrow topology)
Let X be a Hausdorff topological space. The narrow topology on $\mathcal{M}_+(X)$ is the coarsest topology which makes all the maps $\mu \mapsto \int_X \varphi \, d\mu$ lower semicontinuous for every bounded and lower semicontinuous function $\varphi : X \to \mathbb{R}$.

In order to state a useful criterium for compactness in $\mathcal{M}_+(X)$ we give the following definition.

Definition 4.3.2. (Domination of compact sets)
Let X be a Hausdorff topological space and let $\mathcal{K}(X)$ (respectively $\mathcal{G}(X)$) be the collection of the compact (resp. open) subsets of X. We say that a collection $\mathcal{G} \subset \mathcal{G}(X)$ *dominates* the compact subsets of X, and we write $\mathcal{G} \succ \mathcal{K}(X)$, if

$$\text{for every } K \in \mathcal{K}(X) \quad \text{there exists } G \in \mathcal{G} : \quad K \subset G.$$

Theorem 4.3.3. *(Topsoe [116])*
Let X be a Hausdorff topological space.

1. *A net* $(\mu_\lambda)_{\lambda\in\mathbb{L}} \subset \mathcal{M}_+(X)$ *is compact (i.e. from every subnet it is possible to extract a narrowly convergent sub-subnet) if and only if* $\limsup_{\lambda\in\mathbb{L}} \mu_\lambda(X) < +\infty$ *and for every* $\mathcal{G} \succ \mathcal{K}(X)$ *and for every* $\varepsilon > 0$ *there exists a finite subset* $\mathcal{G}' \subset \mathcal{G}$ *such that*

$$\limsup_{\lambda\in\mathbb{L}} \min_{G\in\mathcal{G}'} \mu_\lambda(X \setminus G) \leqslant \varepsilon. \tag{4.3.1}$$

2. *A narrowly closed set* $\mathcal{F} \subset \mathcal{M}_+(X)$ *is narrowly compact if and only if it is bounded and for every* $\mathcal{G} \succ \mathcal{K}(X)$ *and for every* $\varepsilon > 0$ *there exists a finite subset* $\mathcal{G}' \subset \mathcal{G}$ *such that*

$$\sup_{\mu\in\mathcal{F}} \min_{G\in\mathcal{G}'} \mu(X \setminus G) \leqslant \varepsilon. \tag{4.3.2}$$

Remark 4.3.4. Condition (4.3.2) is really a relaxation of the usual uniform tightness condition: in fact, the latter guarantees the existence of a singleton $\mathcal{G}'$ satisfying (4.3.2).

We are now able to state and prove the analogous of Theorem 3.2.3 and Corollary 3.2.4.

Theorem 4.3.5. *(Compactness from converging marginals)*
Let X_i*,* $i = 1, 2$ *be Hausdorff topological spaces and let* $(\boldsymbol{\gamma}_\lambda)_{\lambda\in\mathbb{L}}$ *be a net in* $\mathcal{M}_+(X_1 \times X_2)$ *with* $\mu_{i,\lambda} := \pi^i_\sharp \boldsymbol{\gamma}_\lambda \in \mathcal{M}_+(X_i)$*,* $i = 1, 2$*,* $\lambda \in \mathbb{L}$*. If* $(\mu_{i,\lambda})_{\lambda\in\mathbb{L}}$ *narrowly converge to some* μ_i *in* $\mathcal{M}(X_i)$*, then there exists a subnet* $(\boldsymbol{\gamma}'_\alpha)_{\alpha\in\mathbb{A}}$ *narrowly convergent to some* $\boldsymbol{\gamma} \in \Gamma(\mu_1, \mu_2)$ *in* $\mathcal{M}(X_1 \times X_2)$*.*

Proof. Let us first recall (see e.g. [84, §26, Exercise 9]) that whenever $G \subset X_1 \times X_2$ is an open set containing the product $K_1 \times K_2$ of two compact subsets $K_i \subset X_i$, $i = 1, 2$, then there exist open sets $G_i \subset X_i$ such that

$$K_1 \times K_2 \subset G_1 \times G_2 \subset G. \tag{4.3.3}$$

Let $\mathcal{G} \succ \mathcal{K}(X_1 \times X_2)$ and let $\varepsilon > 0$ be fixed. Thanks to (4.3.3), in order to check (4.3.1) it is not restrictive to replace $\mathcal{G}$ with the collection of cartesian open sets

$$\mathcal{G}_c := \Big\{ G_1 \times G_2 \mid \exists\, G \in \mathcal{G} \text{ such that } G_1 \times G_2 \subset G \Big\}.$$

Let us now introduce the disjoint union $X := X_1 \sqcup X_2$ endowed with the finest topology for which the canonical injections $\iota_i : X_i \to X$ are continuous; we can thus identify X_i with $\iota_i(X_i)$ as (open and closed) subsets of X. Since a set $A \subset X$ is open (resp. compact) in X if and only if $A \cap X_i$ is open (resp. compact) in X_i for $i = 1, 2$, it is not difficult to check that the family of open sets in X

$$\hat{\mathcal{G}}_c := \Big\{ G_1 \sqcup G_2 \mid G_1 \times G_2 \in \mathcal{G}_c \Big\}$$

dominates $\mathcal{K}(X)$.

We now consider the net $\mu_\lambda := (\iota_1)_\sharp \mu_{1,\lambda} + (\iota_2)_\sharp \mu_{2,\lambda}$ in $\mathcal{M}_+(X)$; equivalently, $\mu_\lambda(B) := \mu_{1,\lambda}(B \cap X_1) + \mu_{2,\lambda}(B \cap X_2)$ for every Borel set B of X. It is immediate to check that μ_λ narrowly converges to $\mu := (\iota_1)_\sharp\mu_1 + (\iota_2)_\sharp\mu_2$. By Theorem 4.3.3(1) we can find a finite subset $\hat{\mathcal{G}}' = \{G_{1,j} \sqcup G_{2,j}\}_{j=1}^J$ of $\hat{\mathcal{G}}_c$ such that

$$\limsup_{\lambda\in\mathbb{L}} \min_{G\in\hat{\mathcal{G}}'} \mu_\lambda(X \setminus G) \leqslant \varepsilon. \tag{4.3.4}$$

On the other hand we observe that, for every $\lambda \in \mathbb{L}$ and $j \in \{1, \cdots, J\}$, it holds

$$\begin{aligned}\boldsymbol{\gamma}_\lambda(X_1 \times X_2 \setminus G_{1,j} \times G_{2,j})) &\leqslant \boldsymbol{\gamma}_\lambda((X_1 \setminus G_{1,j}) \times X_2) + \boldsymbol{\gamma}_\lambda(X_1 \times (X_2 \setminus G_{2,j}))\\ &= \mu_{1,\lambda}(X_1 \setminus G_{1,j}) + \mu_{2,\lambda}(X_2 \setminus G_{2,j})\\ &= \mu_\lambda(X \setminus G_{1,j} \sqcup G_{2,j}),\end{aligned}$$

so that, setting $\mathcal{G}' := \{G_{1,j} \times G_{2,j}\}_{j=1}^J$, (4.3.4) yields

$$\limsup_{\lambda\in\mathbb{L}} \min_{G\in\mathcal{G}'} \boldsymbol{\gamma}_\lambda(X_1 \times X_2 \setminus G) \leqslant \limsup_{\lambda\in\mathbb{L}} \min_{G\in\hat{\mathcal{G}}'} \mu_\lambda(X \setminus G) \leqslant \varepsilon.$$

□

Arguing as in the proof of Corollary 3.2.4 we eventually obtain the corresponding characterization of compactness in $\mathcal{M}_+(X_1 \times X_2)$.

Corollary 4.3.6. *(Compactness from compact marginals)*
Let X_i, $i = 1, 2$ be Hausdorff topological spaces and let $\mathcal{K}_i \subset \mathcal{M}_+(X_i)$ be compact in the narrow topology, $i = 1, 2$. Then the set $\mathcal{K} := \left\{\boldsymbol{\gamma} \in \mathcal{M}_+(X_1 \times X_2) \mid \pi^i_\#\boldsymbol{\gamma} \in \mathcal{K}_i\right\}$ is compact in the narrow topology of $\mathcal{M}(X_1 \times X_2)$.

5 DUAL ATTAINMENT, OPTIMALITY CONDITIONS AND METRIC PROPERTIES

This chapter is devoted to the study of the dual attainment in (4.2.4) i.e. to understand under which hypotheses one can find a pair of functions (continuous or not) realizing the maximum in (4.2.4). The first Section 5.1 treats the case in which the spaces are compact and H is somehow more regular; Section 5.2 deals with a more general settings and present sufficient and necessary optimality conditions for the primal problem in the definition of $\mathscr{U}_H$; finally in Section 5.3 the case in which H is a distance on the geometric cone is treated.

This Chapter is the result of a collaboration with Giuseppe Savaré.

5.1 THE REGULAR CASE

In this section we provide sufficent conditions for the existence of a maximizing pair $(\varphi_1, \varphi_2) \in \Phi_H$ such that

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 = \mathscr{U}_H(\mu_1, \mu_2).$$

5.1.1 *The case of* H *finite everywhere*

In the following we assume that (X_1, d_1), (X_2, d_2) are compact metric spaces, that $H : \mathfrak{C}[X_1, X_2] \to [0, +\infty)$ is a 1-homogeneous, convex and continuous function and that $\mu_i \in \mathcal{M}_+(X_i)$ are such that $\operatorname{supp} \mu_i = X_i$ for $i = 1, 2$ (notice that this implies $\mu_i \neq 0_{X_i}$ for $i = 1, 2$). We define $H_i : X_i \to [0, +\infty)$ for $i = 1, 2$ as

$$H_1(x_1) := H([x_1, 1], \mathfrak{o}), \quad H_2(x_2) := H(\mathfrak{o}, [x_2, 1]) \quad x_1 \in X_1, \, x_2 \in X_2 \qquad (5.1.1)$$

and we assume that

$$\kappa_1 := \int_{X_1} H_1 \, d\mu_1 < +\infty, \quad \kappa_2 := \int_{X_2} H_2 \, d\mu_2 < +\infty. \qquad (5.1.2)$$

Finally, we assume some control on the derivatives of H at the boundary of the cone, meaning that we assume the existence of an open set $\Omega \subset X_1 \times X_2$ such that

$$\begin{aligned} &\pi^i(\Omega) = X_i, \quad i = 1, 2, \\ &\lim_{r_1 \downarrow 0} \frac{H([x_1, r_1], [x_2, 1]) - H_2(x_2)}{r_1} = -\infty \quad \text{for every } (x_1, x_2) \in \Omega, \\ &\lim_{r_2 \downarrow 0} \frac{H([x_1, 1], [x_2, r_2]) - H_1(x_1)}{r_2} = -\infty \quad \text{for every } (x_1, x_2) \in \Omega. \end{aligned} \qquad (5.1.3)$$

Lemma 5.1.1. *There exists a finite set* $\{x_1^n, x_2^n, r_n\}_{n=1}^N \subset X_1 \times X_2 \times (0, +\infty)$ *such that*

$$\bigcup_n B(x_i^n, r_n) = X_i, \quad i = 1,2 \qquad \bigcup_n \overline{B(x_1^n, r_n)} \times \overline{B(x_2^n, r_n)} \subset \Omega.$$

Moreover there exists a constant $C = C(X_1, X_2, \mu_1, \mu_2, H) > 0$ *such that, if* $(\varphi_1, \varphi_2) \in \Phi_H$ *are such that*

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 \geqslant 0,$$

then for every $n \in \{1, \dots, N\}$ *there exist* $y_1^n \in \overline{B(x_1^n, r_n)}, y_2^n \in \overline{B(x_2^n, r_n)}$ *such that*

$$\varphi_1(y_1^n) \geqslant -C, \quad \varphi_2(y_2^n) \geqslant -C. \tag{5.1.4}$$

Proof. The first claim follows by the compactness of X_i and by the properties of Ω. We claim that

$$\int_{B(x_1^n, r_i)} \varphi_1 \, d\mu_1 \geqslant -(\kappa_1 + \kappa_2 + 1), \quad \int_{B(x_2^n, r_i)} \varphi_2 \, d\mu_2 \geqslant -(\kappa_1 + \kappa_2 + 1) \tag{5.1.5}$$

for every $n = 1, \dots, N$. Indeed, if there exists $i \in \{1, 2\}$ (say $i = 1$) and $n \in \{1, \dots, N\}$ such that

$$\int_{B(x_1^n, r_n)} \varphi_1 \, d\mu_1 < -(\kappa_1 + \kappa_2 + 1),$$

then

$$\begin{aligned} \int_{X_1} \varphi_1 \, d\mu_1 &= \int_{B(x_1^n, r_n)} \varphi_1 \, d\mu_1 + \int_{X_1 \setminus B(x_1^n, r_n)} \varphi_1 \, d\mu_1 \\ &< -(\kappa_1 + \kappa_1 + 1) + \kappa_1 \\ &= -(\kappa_2 + 1). \end{aligned}$$

Thus

$$\kappa_2 \geqslant \int_{X_2} \varphi_2 \, d\mu_2 \geqslant -\int_{X_1} \varphi_1 \, d\mu_1 \geqslant \kappa_2 + 1,$$

a contradiction. Let us set

$$m := \min_{i=1,2} \min_{n=1,\dots,N} \mu_i \left(B(x_i^n, r_n) \right) > 0,$$

since the supports of the measures coincide with the whole spaces X_1 and X_2. By (5.1.5) we have, for every $i = 1, 2$, that

$$\mu_i \left(B(x_i^n, r_n) \right) \sup_{B(x_i^n, r_n)} \varphi_i \geqslant \int_{B(x_i^n, r_n)} \varphi_i \, d\mu_i \geqslant -(\kappa_1 + \kappa_2 + 1),$$

hence

$$\sup_{B(x_i^n, r_n)} \varphi_i \geqslant -\frac{\kappa_1 + \kappa_2 + 1}{m}.$$

This gives the existence of $y_i^n \in \overline{B(x_i^n, r_n)}$ such that

$$\varphi_i(y_i^n) \geqslant -\frac{\kappa_1 + \kappa_2 + 1}{m} =: C.$$

□

Lemma 5.1.2. *There exists* $\varepsilon = \varepsilon(X_1, X_2, \mu_1, \mu_2, H) > 0$ *such that, for every* $(\varphi_1, \varphi_2) \in \Phi_H$ *with*

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 \geqslant 0,$$

it holds

$$\varphi_i(x_i) \leqslant H_i(x_i) - \varepsilon \quad \textit{for every } x_i \in X_i, \quad i = 1, 2.$$

Proof. We prove the statement for $i = 1$, being the other case completely analogous. Suppose by contradiction that there exists $(\varphi_1^j, \varphi_2^j)_j \subset \Phi_H$ with $\int_{X_1} \varphi_1^j \, d\mu_1 + \int_{X_2} \varphi_2^j \, d\mu_2 \geqslant 0$ and $(z_j)_j \subset X_1$ such that

$$H_1(z_j) - \varphi_1^j(z_j) \to 0$$

as $j \to +\infty$. Up to passing to a subsequence, we can assume that

$$0 \leqslant H_1(z_j) - \varphi_1^j(z_j) \leqslant \frac{1}{j} \quad \text{for every } j \in \mathbb{N}$$

and the existence of $z \in X_1$ such that $z_j \to z$. By Lemma 5.1.1, we have that $z \in B(x_1^n, r_n)$ for some $n \in \{1, \dots, N\}$ and we can thus assume, up to passing again to a subsequence, that $z_j \in B(x_1^n, r_n)$ for every $j \in \mathbb{N}$. By Lemma 5.1.1, we can find, $y_j \in \overline{B(x_2^n, r_n)}$ such that $\varphi_2^j(y_j) \geqslant -C$. By compactness of X_2, we can assume that $y_j \to y \in \overline{B(x_2^n, r_n)}$. We have thus proven the existence of $(z_j, y_j) \in \Omega$ such that $(z_j, y_j) \to (z, y) \in \Omega$ with

$$0 \leqslant H_1(z_j) - \varphi_1^j(z_j) \leqslant \frac{1}{j}, \quad \varphi_2^j(y_j) \geqslant -C \quad \text{for every } j \in \mathbb{N}.$$

We have

$$r_1 \left(H_1(z_j) - \frac{1}{j} \right) - Cr_2 \leqslant \varphi_1^j(z_j) r_1 + \varphi_2^j(y_j) r_2 \leqslant H([z_j, r_1], [y_j, r_2])$$

for every $r_1, r_2 \geqslant 0$. Choosing $r_1 = 1$, we get

$$\frac{H([z_j, 1], [y_j, r_2] - H_1(z_j)}{r_2} \geqslant -C - \frac{1}{jr_2} \quad \text{for every } j \in \mathbb{N},\ r_2 > 0.$$

Passing first to the limit as $j \to +\infty$ and then to the limit as $r_2 \downarrow 0$, we obtain

$$\lim_{r_2 \downarrow 0} \frac{H([z, 1], [y, r_2] - H_1(x_\infty)}{r_2} \geqslant -C > -\infty,$$

a contradiction with (5.1.3). □

Definition 5.1.3. Let $(\varphi_1, \varphi_2) \in \Phi_H$. We define the Borel functions $\varphi_1^H : X_2 \to \mathbb{R}$, $\varphi_1^{HH} : X_1 \to \mathbb{R}$ as

$$\varphi_1^H(x_2) := \inf_{x_1 \in X_1} \inf_{\alpha \geqslant 0} \left\{ H([x_1, \alpha], [x_2, 1]) - \alpha \varphi_1(x_1) \right\}, \quad x_2 \in X_2,$$

$$\varphi_1^{HH}(x_1) := \inf_{x_2 \in X_2} \inf_{\alpha \geqslant 0} \left\{ H([x_1, 1], [x_2, \alpha]) - \alpha \varphi_1^H(x_2) \right\}, \quad x_1 \in X_1.$$

Proposition 5.1.4. *There exist constants* $R = R(X_1, X_2, \mu_1, \mu_2, H) > 1$ *and* $M = M(X_1, X_2, \mu_1, \mu_2, H) > 0$ *such that, for every* $(\varphi_1, \varphi_2) \in \Phi_H$ *with*

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 \geqslant 0,$$

it holds

$$\varphi_1^H(x_2) = \inf_{x_1 \in X_1} \inf_{0 \leqslant \alpha \leqslant R} \Big\{ H([x_1, \alpha], [x_2, 1]) - \alpha \varphi_1(x_1) \Big\}, \quad x_2 \in X_2, \tag{5.1.6}$$

$$\varphi_1^{HH}(x_1) = \inf_{x_2 \in X_2} \inf_{0 \leqslant \alpha \leqslant R} \Big\{ H([x_1, 1], [x_2, \alpha]) - \alpha \varphi_1^H(x_2) \Big\}, \quad x_1 \in X_1. \tag{5.1.7}$$

In particular, $(\varphi_1^{HH}, \varphi_1^H) \in \Phi_H$, $\varphi_1^{HH} \geqslant \varphi_1$, $\varphi_1^H \geqslant \varphi_2$, *both of them are uniformly continuous with the same (uniform) modulus of continuity of* H *on* $\mathfrak{C}_R[X_1, X_2]$ *and*

$$\|\varphi_1^H\|_\infty \leqslant M, \quad \|\varphi_1^{H,H}\|_\infty \leqslant M.$$

Proof. Let (φ_1, φ_2) be as in the statement. By Lemma 5.1.2 we know that there exists $\varepsilon > 0$ (not depending on the couple) such that

$$\varphi_1(x_1) \leqslant H_1(x_1) - \varepsilon \quad \text{for every } x_1 \in X_1.$$

Then, by uniform continuity of H on $\mathfrak{C}_1[X_1, X_2]$, we can find $0 < \delta < 1$ such that

$$|H([x_1, 1], [x_2, r_2]) - H_1(x_1)| \leqslant \frac{\varepsilon}{2} \quad \text{for every } 0 \leqslant r_2 \leqslant \delta.$$

If we define $R := 1 + \frac{1}{\delta} + \frac{2}{\varepsilon}(\max_{X_2} H_2 + \max_{X_1} H_1 + 1)$, then, for every $\alpha > R$, we have

$$\begin{aligned} H([x_1, \alpha], [x_2, 1]) - \alpha \varphi_1(x_1) \\ &= H([x_1, \alpha], [x_2, 1]) - H_1(x_1)\alpha + \alpha (H_1(x_1) - \varphi_1(x_1)) \\ &= \alpha (H([x_1, 1], [x_2, 1/\alpha]) - H_1(x_1)) + \alpha (H_1(x_1) - \varphi_1(x_1)) \\ &\geqslant \alpha \frac{\varepsilon}{2} \\ &\geqslant H_2(x_2) + 1. \end{aligned}$$

Thus, for every $x_2 \in X_2$, we get

$$\begin{aligned} \inf_{x_1 \in X_1} \inf_{\alpha > R} \{H([x_1, \alpha], [x_2, 1]) - \alpha \varphi_1(x_1)\} &\geqslant H_2(x_2) + 1 \\ &> \inf_{x_1 \in X_1} \inf_{0 \leqslant \alpha \leqslant R} \{H([x_1, \alpha], [x_2, 1]) - \alpha \varphi_1(x_1)\} \end{aligned}$$

and this proves (5.1.6). The proof of (5.1.7) is analogous.
The fact that $\varphi_1^H \geqslant \varphi_2$, $\varphi_1^{HH} \geqslant \varphi_1$ and

$$\varphi_1^{HH}(x_1) r_1 + \varphi_1^H(x_2) r_2 \leqslant H([x_1, r_1], [x_2, r_2])$$

for every $(x_1, x_2) \in X_1 \times X_2$, $r_1, r_2 \geqslant 0$, follow by the definition of φ_1^H and φ_1^{HH}. It is then clear that φ_1^H (resp. φ_1^{HH}) is bounded from below by $\min_{x_2 \in X_2} \varphi_2$

(resp. $\min_{x_1 \in X_1} \varphi_1$) and by above by $\max_{x_2 \in X_2} H_2$ (resp. $\max_{x_1 \in X_2} H_1$). Let now $x_2, x_2' \in X_2$, then, recalling (3.2.4), we have

$$\begin{aligned} \left|\varphi_1^H(x_2) - \varphi_1^H(x_2')\right| &\leqslant \sup_{x_1 \in X_1} \sup_{0 \leqslant \alpha \leqslant R} \left|H([x_1, \alpha], [x_2, 1]) - H([x_1, \alpha], [x_2', 1])\right| \\ &\leqslant \omega_H^R\left((d_1 \otimes_{\mathfrak{C}} d_2)\left(([x_1, \alpha], [x_2, 1]), ([x_1, \alpha], [x_2', 1])\right)\right) \\ &= \omega_H^R(d_{2,\mathfrak{C}}([x_2, 1], [x_2', 1])) \\ &\leqslant \omega_H^R(d_2(x_2, x_2')), \end{aligned} \tag{5.1.8}$$

where ω_H^R is the (uniform) modulus of continuity of H on $\mathfrak{C}_R[X_1, X_2]$ and we have used that $d_{2,\mathfrak{C}}([x_2, 1], [x_2', 1])) \leqslant d_2(x_2, x_2')$ (see formula (7.5) in [76]). The analogous statement for φ_1^{HH} follows by the same strategy. This proves that φ_1^H and φ_1^{HH} are uniformly continuous with the same (uniform) modulus of continuity of H on $\mathfrak{C}_R[X_1, X_2]$ and concludes the proof that $(\varphi_1^{HH}, \varphi_1^H) \in \Phi_H$. If we define (recalling Lemma 5.1.1)

$$M := C + \omega_H^R(\operatorname{diam} X_1) + \omega_H^R(\operatorname{diam} X_2) + \max_{x_1 \in X_1} H_1 + \max_{x_2 \in X_2} H_2,$$

we have that $\varphi_1^H \leqslant H_2 \leqslant M$ and, by (5.1.8), we get

$$\varphi_1^H(x_2) \geqslant \varphi_1^H(x_2') - \omega_H^R(d_2(x_2, x_2')) \geqslant -C - \omega_H^R(\operatorname{diam} X_2) \geqslant -M,$$

for every $x_2 \in X_2$, where $x_2' \in X_2$ is some point where φ_1^H is larger than $-C$ (whose existence is given by Lemma 5.1.1). The proof for φ_1^{HH} is the same. □

Theorem 5.1.5. *There exists* $(\varphi_1, \varphi_2) \in \Phi_H$ *such that*

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 = \mathscr{U}_H(\mu_1, \mu_2).$$

Proof. If $\mathscr{U}_H(\mu_1, \mu_2) = 0$, we can take φ_1 and φ_2 to be the null functions. We thus assume that $\mathscr{U}_H(\mu_1, \mu_2) > 0$. If this is the case, we can find a maximizing sequence $(\varphi_1^j, \varphi_2^j)_j \subset \Phi_H$ for the dual problem (4.2.4) with

$$\int_{X_1} \varphi_1^j \, d\mu_1 + \int_{X_2} \varphi_2^j \, d\mu_2 \geqslant 0 \quad \text{for every } j \in \mathbb{N}.$$

By Proposition 5.1.4 we have that $(\varphi_1^{j,HH}, \varphi_1^{j,H})_j \subset \Phi_H$ is a maximizing sequence of equi-uniformly continuous and equi-bounded functions. By Arzelà–Ascoli theorem, we can assume, up to passing to a subsequence, that there exists a pair $(\varphi_1, \varphi_2) \in \Phi_H$ such that $(\varphi_1^{j,HH}, \varphi_1^{j,H}) \to (\varphi_1, \varphi_2)$ uniformly on the compact space $X_1 \times X_2$. By dominated convergence, we have

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 = \lim_j \left(\int_{X_1} \varphi_1^{j,HH} \, d\mu_1 + \int_{X_2} \varphi_1^{j,H} \, d\mu_2 \right) = \mathscr{U}_H(\mu_1, \mu_2).$$

□

In the next statement we assume that $(X_1, d_1) = (X_2, d_2) = (K, d_e)$, where $K \subset \mathbb{R}^d$ is a compact convex set with non-empty interior and d_e is the Euclidean distance on K. We say that $H : \mathfrak{C}[K,K] \to [0,+\infty)$ is *partially differentiable* if the limits

$$(\partial_1 H(x_1,r_1;x_2,r_2))_n := \lim_{h\to 0} \frac{H([x_1+he_n,r_1],[x_2,r_2]) - H([x_1,r_1],[x_2,r_2])}{h},$$
$$\partial_2 H(x_1,r_1;x_2,r_2) := \lim_{h\to 0} \frac{H([x_1,r_1+h],[x_2,r_2]) - H([x_1,r_1],[x_2,r_2])}{h}$$

exist for every $n = 1,\dots,d$, $x_1 \in \operatorname{int}(K)$, $r_1 > 0$ and $x_2 \in K$, $r_2 \geqslant 0$, where $(e_n)_{n=1}^d$ is the canonical basis of $\mathbb{R}^d$.

Theorem 5.1.6. *Let $K \subset \mathbb{R}^d$ be a compact and convex set with non-empty interior, let $H : \mathfrak{C}[K,K] \to [0,+\infty)$ be a continuous, 1-homogeneous and convex function which is in addition partially differentiable and $(d_e \otimes_{\mathfrak{C}} d_e)$-Lipschitz continuous on $\mathfrak{C}_R[K,K]$, where $R > 1$ is as in Proposition 5.1.4. Let $\mu_i \in \mathcal{M}_+(K)$ with $\operatorname{supp}\mu_i = K$, $i = 1,2$, and assume that (5.1.2) and (5.1.3) hold true. If μ_1 is absolutely continuous w.r.t. $\mathcal{L}^d|_K$ (the d dimensional Lebesgue measure on K) and*

$$\text{for every } x_1 \in \operatorname{int}(K) \text{ the map}$$
$$\mathfrak{C}[K] \ni [y,q] \mapsto \begin{pmatrix} \partial_1 H(x_1,1;y,q) \\ \partial_2 H(x_1,1;y,q) \end{pmatrix} \in \mathbb{R}^{d+1} \text{ is invertible,} \tag{5.1.9}$$

then there exists a Borel map $T : K \to \mathfrak{C}[K]$ s.t.

$$\mu_2 = (\mathfrak{h} \circ T_\sharp)(\mu_1), \qquad \int_K H([x_1,1],T(x_1))\,d\mu_1(x_1) = \mathscr{U}_H(\mu_1,\mu_2).$$

Proof. By Theorem 5.1.5, we know that there exists a pair $(\varphi_1,\varphi_2) \in \Phi_H$ of Lipschitz continuous functions (see also Proposition 5.1.4) such that

$$\int_K \varphi_1\,d\mu_1 + \int_K \varphi_2\,d\mu_2 = \mathscr{U}_H(\mu_1,\mu_2).$$

If $\alpha \in \mathfrak{H}^1_H(\mu_1,\mu_2)$ (cf. (4.1.6)), then we can find a full α-measure Borel set $\Gamma \subset \mathfrak{C}[K,K]$ such that

$$\varphi_1(x_1)r_1 + \varphi_2(x_2)r_2 = H([x_1,r_1],[x_2,r_2]) \quad \text{for every } ([x_1,r_1],[x_2,r_2]) \in \Gamma.$$

Notice that, since $H(\mathfrak{o},\mathfrak{o}) = 0$, we can assume that $(\mathfrak{o},\mathfrak{o}) \notin \Gamma$. Since φ_1 is Lipschitz continuous, we can find a full μ_1-measure Borel set $U \subset \operatorname{int}(K)$ where φ_1 is differentiable (since K is convex, its boundary has 0 Lebesgue measure). Let $([\bar{x}_1,\bar{r}_1],[\bar{x}_2,\bar{r}_2]) \in \Gamma$ with $\bar{x}_1 \in U$. We have that

$$\begin{gathered}\text{the map } (x_1,r_1) \mapsto H([x_1,r_1],[\bar{x}_2,\bar{r}_2]) - \varphi_1(x_1)r_1 \\ \text{has a minimum at } (x_1,r_1) = (\bar{x}_1,\bar{r}_1).\end{gathered} \tag{5.1.10}$$

We claim that $\bar{r}_1 \neq 0$: if not, we would have that

$$H([x_1,r_1],[\bar{x}_2,\bar{r}_2]) - \varphi_1(x_1)r_1 \geqslant H(\mathfrak{o},[\bar{x}_2,\bar{r}_2]) \quad \text{for every } x_1 \in K,\ r_1 > 0.$$

By (5.1.3), there exists $\tilde{x}_1 \in K$ such that $(\tilde{x}_1,\bar{x}_2) \in \Omega$ so that, if we chose $x_1 = \tilde{x}_1$, we get

$$H([\tilde{x}_1,r_1],[\bar{x}_2,\bar{r}_2]) - \varphi_1(\tilde{x}_1)r_1 \geqslant H(\mathfrak{o},[\bar{x}_2,\bar{r}_2]) \quad \text{for every } r_1 > 0.$$

Since $\bar{r}_2 \neq 0$ (recall that we assumed that $(\mathfrak{o}, \mathfrak{o}) \notin \Gamma$), we can divide by $\bar{r}_2$ and obtain

$$\frac{H([\tilde{x}_1, \frac{r_1}{\bar{r}_2}], [\bar{x}_2, 1]) - H_2(\bar{x}_2)}{r_1} \geqslant \varphi_1(\tilde{x}_1) \quad \text{for every } r_1 > 0,$$

which, passing to the limit as $r_1 \downarrow 0$, leads to a contradiction with (5.1.3).
Thus from (5.1.10) it follows that

$$\begin{cases} \partial_1 H(\bar{x}_1, \bar{r}_1; \bar{x}_2, \bar{r}_2) = \nabla\varphi_1(\bar{x}_1)\bar{r}_1, \\ \partial_2 H(\bar{x}_1, \bar{r}_1; \bar{x}_2, \bar{r}_2) = \varphi_1(\bar{x}_1) \end{cases}$$

which, using the 1-homogeneity of H, can be rewritten as

$$\begin{cases} \partial_1 H(\bar{x}_1, 1; \bar{x}_2, \bar{r}_2/\bar{r}_1) = \nabla\varphi_1(\bar{x}_1), \\ \partial_2 H(\bar{x}_1, 1; \bar{x}_2, \bar{r}_2/\bar{r}_1) = \varphi_1(\bar{x}_1). \end{cases}$$

Let us denote by $I_{x_1} \subset \mathbb{R}^{d+1}$ the image of the map in (5.1.9) and by $f_{x_1} : I_{x_1} \to \mathfrak{C}[X_2]$ the inverse of such map. If we define $T : X_1 \to \mathfrak{C}[X_2]$ as

$$T(x_1) := \begin{cases} f_{x_1}(\nabla\varphi_1(x_1), \varphi_1(x_1)) & \text{if } (\nabla\varphi_1(x_1), \varphi_1(x_1)) \in I_{x_1}, \\ \mathfrak{o} & \text{else }, \end{cases}$$

then T is a Borel map and we have just proven that

$$[\bar{x}_2, \bar{r}_2/\bar{r}_1] = T(\bar{x}_1),$$

meaning that

$$r_1 T(x_1) = [x_2, r_2] \quad \text{for every } ([x_1, r_1], [x_2, r_2] \in \Gamma, \, x_1 \in U.$$

From this and the fact that $\Gamma \cap \mathbf{x}_1^{-1}(U)$ has full $\boldsymbol{\alpha}$-measure (recall that $\{\mathbf{r}_1 = 0\} \cap \Gamma = \emptyset$) it follows that

$$\boldsymbol{\alpha} = ([\mathrm{id}_{X_1}, 1], T)_\sharp \mu_1$$

which leads to the conclusion. □

5.1.2 *The case of* H *finite on a cone*

In this subsection we assume that (X_1, d_1), (X_2, d_2) are compact metric spaces, $\mu_i \in \mathcal{M}_+(X_i)$, $i = 1, 2$ with $\mu_i \neq 0_{X_i}$ for $i = 1, 2$ and $H : \mathfrak{C}[X_1, X_2] \to [0, +\infty]$ is a 1-homogeneous, convex and continuous function.
We suppose the existence of two nonnegative numbers q_1, q_2 with

$$q_1 < \frac{\mu_2(X_2)}{\mu_1(x_1)}, \quad q_2 < \frac{\mu_1(X_1)}{\mu_2(X_2)} \tag{5.1.11}$$

such that the open set

$$U_{q_1 q_2} := \{([x_1, r_1], [x_2, r_2]) \in \mathfrak{C}[X_1, X_2] \mid r_2 > r_1 q_1, \, r_1 > r_2 q_2\}$$

is nonempty and it holds

$$\begin{aligned} \lim_{r_1 \downarrow q_2} \inf_{(x_1,x_2)\in X_1\times X_2} H([x_1,r_1],[x_2,1]) &= +\infty, \\ \lim_{r_2 \downarrow q_1} \inf_{(x_1,x_2)\in X_1\times X_2} H([x_1,1],[x_2,r_2]) &= +\infty, \\ H(\mathfrak{y}_1,\mathfrak{y}_2) < +\infty \quad &\text{for every } (\mathfrak{y}_1,\mathfrak{y}_2) \in U_{q_1q_2}. \end{aligned} \tag{5.1.12}$$

Notice that this implies that $H = +\infty$ on $\mathfrak{C}[X_1, X_2] \setminus (U_{q_1q_2} \cup (\mathfrak{o}, \mathfrak{o}))$. Moreover, if $q_1 = q_2 = 0$, then $U_{q_1q_2}$ is simply the open product cone.

Proposition 5.1.7. *There exists a constant* $C = C(X_1, X_2, \mu_1, \mu_2, H) > 0$ *such that, for every* $(\varphi_1, \varphi_2) \in \Phi_H$ *with*

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 \geqslant 0,$$

it holds

$$\varphi_1(x_1) \leqslant C, \quad \varphi_2(x_2) \leqslant C \quad \textit{for every } (x_1,x_2) \in X_1 \times X_2$$

and there exists $(\bar{x}_1, \bar{x}_2) \in X_1 \times X_2$ *such that* $\varphi_1(\bar{x}_1) \geqslant -C$, $\varphi_2(\bar{x}_2) \geqslant -C$.

Proof. We start from the last claim for φ_1. Assume by contradiction that there exists a sequence $(\varphi_1^j, \varphi_2^j)_j \subset \Phi_H$ with $\int_{X_1} \varphi_1^j \, d\mu_1 + \int_{X_2} \varphi_2^j \, d\mu_2 \geqslant 0$ such that $\max_{x_1 \in X_1} \varphi_1^j(x_1) \to -\infty$. Let $(x_1^j)_j \subset X_1$ be the sequence of points where the maxima are attained. We thus have

$$\varphi_1^j(x_1^j)\mu_1(X_1) + \int_{X_2} \varphi_2^j \, d\mu_2 \geqslant 0 \quad \text{for every } j \in \mathbb{N}$$

so that we can find $(x_2^j)_j \subset X_2$ such that

$$\varphi_2^j(x_2^j) \geqslant -\frac{\varphi_1^j(x_1^j)\mu_1(X_1)}{\mu_2(X_2)} \quad \text{for every } j \in \mathbb{N}.$$

Since $(\varphi_1^j, \varphi_2^j) \in \Phi_H$, we have

$$\varphi_1^j(x_1^j)\left(r_1 - \frac{\mu_1(X_1)}{\mu_2(X_2)} r_2\right) \leqslant \varphi_1^j(x_1^j) r_1 + \varphi_2^j(x_2^j) \leqslant H([x_1^j, r_1], [x_2^j, r_2])$$

for every $r_1, r_2 \geqslant 0$. We can assume, up to passing to a subsequence, that $(x_1^j, x_2^j) \to (x_1, x_2) \in X_1 \times X_2$. Thanks to (5.1.11) and (5.1.12), we can find $\bar{r}_1, \bar{r}_2 > 0$ such that

$$H([x_1, \bar{r}_1], [x_2, \bar{r}_2]) < +\infty, \quad \bar{r}_1 - \frac{\mu_1(X_1)}{\mu_2(X_2)} \bar{r}_2 < 0.$$

We thus have that

$$+\infty \leqslant H([x_1, \bar{r}_1], [x_2, \bar{r}_2],$$

a contradiction with (5.1.12). Since the proof for φ_2 is the same, we have proven that there exists a constant $D > 0$ independent of (φ_1, φ_2) and a point $(\bar{x}_1, \bar{x}_2) \in X_1 \times X_2$ such that

$$\varphi_1(\bar{x}_1) \geqslant -D, \quad \varphi_2(\bar{x}_2) \geqslant -D.$$

Thus, if we set

$$C := D + \max_{(x_1,x_2)\in X_1\times X_2} H([x_1, \bar{r}_1], [x_2, \bar{r}_2]),$$

where $\bar{r}_1$ and $\bar{r}_2$ are as above, we get that

$$\varphi_1(x_1) \leqslant \frac{1}{\bar{r}_1}(H([x_1, \bar{r}_1], [\bar{x}_2, \bar{r}_1]) - \bar{r}_2\varphi_2(\bar{x}_2)) \leqslant C \quad \text{for every } x_1 \in X_1$$

and the corresponding statement for φ_2. □

Proposition 5.1.8. *There exist constants $a_i, a_s, b_i, b_s, M > 0$, depending only on $X_1, X_2, \mu_1, \mu_2, H$, such that, for every $(\varphi_1, \varphi_2) \in \Phi_H$ with*

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 \geqslant 0,$$

it holds

$$\varphi_1^H(x_2) = \inf_{x_1\in X_1} \inf_{a_i\leqslant\alpha\leqslant a_s} \Big\{ H([x_1, \alpha], [x_2, 1]) - \alpha\varphi_1(x_1) \Big\}, \quad x_2 \in X_2, \tag{5.1.13}$$

$$\varphi_1^{HH}(x_1) = \inf_{x_2\in X_2} \inf_{b_i\leqslant\alpha\leqslant b_s} \Big\{ H([x_1, 1], [x_2, \alpha]) - \alpha\varphi_1^H(x_2) \Big\}, \quad x_1 \in X_1. \tag{5.1.14}$$

Moreover the sets

$$\mathfrak{C}_0^1 := \{(\mathfrak{y}_1, \mathfrak{y}_2) \in \mathfrak{C}[X_1, X_2] \mid a_i \leqslant r(\mathfrak{y}_1) \leqslant a_s,\ r(\mathfrak{y}_2) = 1\},$$
$$\mathfrak{C}_0^2 := \{(\mathfrak{y}_1, \mathfrak{y}_2) \in \mathfrak{C}[X_1, X_2] \mid b_i \leqslant r(\mathfrak{y}_2) \leqslant b_s,\ r(\mathfrak{y}_1) = 1\}$$

are compact subsets of $U_{q_1q_2}$, $(\varphi_1^{HH}, \varphi_1^H) \in \Phi_H$, $\varphi_1^{HH} \geqslant \varphi_1$, $\varphi_1^H \geqslant \varphi_2$, φ_1^{HH} is uniformly continuous with the same (uniform) modulus of continuity of H on $\mathfrak{C}_0^2$, φ_1^H is uniformly continuous with the same (uniform) modulus of continuity of H on $\mathfrak{C}_0^1$ and

$$\|\varphi_1^H\|_\infty \leqslant M, \quad \|\varphi_1^{H,H}\|_\infty \leqslant M.$$

Proof. Let (φ_1, φ_2) be as in the statement. Let us set

$$\delta := \frac{1}{2}\frac{1 - q_1q_2}{1 + q_1 + q_2}$$

so that, for every $0 < \varepsilon \leqslant \delta$, we have

$$q_2 + \varepsilon < \frac{1}{q_1 + \varepsilon}.$$

Let us fix a point $\bar{\alpha} \in (q_2 + \delta, \frac{1}{q_1+\delta})$ and let us define

$$m := \max_{(x_1,x_2)\in X_1\times X_2} H([x_1,\bar{\alpha}],[x_2,1]) < +\infty,$$

since $([x_1,\bar{\alpha}],[x_2,1]) \in U_{q_1 q_2}$ for every $(x_1,x_2) \in X_1 \times X_2$.
By (5.1.12), we know that for every $L > 0$, there exists $\varepsilon_L > 0$ such that

$$H([x_1,r_1],[x_2,1] \geqslant L \quad \text{for every } (x_1,x_2) \in X_1\times X_2,\ 0 \leqslant r_1 < q_2 + \varepsilon_L,$$
$$H([x_1,1],[x_2,r_2] \geqslant L \quad \text{for every } (x_1,x_2) \in X_1\times X_2,\ 0 \leqslant r_2 < q_1 + \varepsilon_L.$$

Let

$$L := \max\{m + C(\bar{\alpha} + q_2 + \delta), (q_1+\delta)(m+\bar{\alpha}C) + C\},$$

where C comes from Proposition 5.1.7, and let us take any $a_i, a_s > 0$ such that

$$q_2 < a_i < q_2 + \varepsilon_L \wedge \delta, \quad \frac{1}{q_1 + \varepsilon_L \wedge \delta} < a_s < \frac{1}{q_1},$$

so that $0 < a_i < a_s$, $\mathfrak{C}_0^1 \subset U_{q_1 q_2}$ and $a_i \leqslant \bar{\alpha} \leqslant a_s$.
If $\alpha > a_s$, then, for every $(x_1,x_2) \in X_1 \times X_2$, we have

$$\begin{aligned}
H([x_1,\alpha],[x_2,1]) - \alpha\varphi_1(x_1) &= \alpha\,(H([x_1,1],[x_2,1/\alpha]) - \varphi_1(x_1)) \\
&\geqslant \alpha(L-C) \\
&\geqslant a_s(L-C) \\
&\geqslant m + \bar{\alpha}C \\
&\geqslant H([\bar{x}_1,\bar{\alpha}]),[x_2,1] - \bar{\alpha}\varphi_1(\bar{x}_1),
\end{aligned}$$

where $\bar{x}_1$ comes from Proposition 5.1.7. If $\alpha < a_i$, then, for every $(x_1,x_2) \in X_1 \times X_2$, we have

$$\begin{aligned}
H([x_1,\alpha],[x_2,1]) - \alpha\varphi_1(x_1) &\geqslant L - \alpha C \\
&\geqslant L - C a_i \\
&\geqslant m + \bar{\alpha}C \\
&\geqslant H([\bar{x}_1,\bar{\alpha}],[x_2,1]) - \bar{\alpha}\varphi_1(\bar{x}_1).
\end{aligned}$$

Thus, for every $x_2 \in X_2$, we get

$$\begin{aligned}
&\inf_{x_1\in X_1} \inf_{0\leqslant\alpha<a_i \vee \alpha>a_s} \{H([x_1,\alpha],[x_2,1]) - \alpha\varphi_1(x_1)\} \\
&\qquad \geqslant H([\bar{x}_1,\bar{\alpha}],[x_2,1]) - \bar{\alpha}\varphi_1(\bar{x}_1) \\
&\qquad > \inf_{x_1\in X_1} \inf_{a_i\leqslant\alpha\leqslant a_s} \{H([x_1,\alpha],[x_2,1]) - \alpha\varphi_1(x_1)\}
\end{aligned}$$

and this proves (5.1.13). The proof of (5.1.14) is analogous.
The remaining part of the proof is identical to the one of Proposition 5.1.4. □

By Proposition 5.1.8 we obtain Theorem 5.1.5 also in this setting with exactly the same proof. In the next statement we assume that $(X_1, d_1) = (X_2, d_2) = (K, d_e)$, where $K \subset \mathbb{R}^d$ is a compact convex set with nonempty interior and d_e

is the Euclidean distance on K. We say that $H : \mathfrak{C}[K,K] \to [0,+\infty]$ is *partially differentiable* if the limits

$$(\partial_1 H(x_1,r_1;x_2,r_2))_n := \lim_{h\to 0} \frac{H([x_1+he_n,r_1],[x_2,r_2]) - H([x_1,r_1],[x_2,r_2])}{h},$$
$$\partial_2 H(x_1,r_1;x_2,r_2) := \lim_{h\to 0} \frac{H([x_1,r_1+h],[x_2,r_2]) - H([x_1,r_1],[x_2,r_2])}{h}$$

exist for every $n = 1,\dots,d$, every $([x_1,r_1],[x_2,r_2]) \in U_{q_1q_2}$ with $x_1 \in \operatorname{int}(K)$, where $(e_n)_{n=1}^d$ is the canonical basis of $\mathbb{R}^d$. Notice that, for every $x_1 \in X_1$, the map

$$[y,q] \mapsto \begin{pmatrix} \partial_1 H(x_1,1;y,q) \\ \partial_2 H(x_1,1;y,q) \end{pmatrix}$$

is well defined for those $[y,q]$ in

$$D_{x_1} := \left\{ \mathfrak{y} \in \mathfrak{C}[K] \mid q_1 < r(\mathfrak{y}) < \frac{1}{q_2} \right\}.$$

Theorem 5.1.9. *Let $K \subset \mathbb{R}^d$ be a compact and convex set with non-empty interior, let $H : \mathfrak{C}[K,K] \to [0,+\infty]$ be a continuous, 1-homogeneous and convex function which is in addition partially differentiable and $(d_e \otimes_{\mathfrak{C}} d_e)$-Lipschitz continuous on $\mathfrak{C}_0^2$ (see Proposition 5.1.8). Let $\mu_i \in \mathcal{M}_+(K)$ and assume that (5.1.12) holds true. If μ_1 is absolutely continuous w.r.t. $\mathcal{L}^d|_K$ (the d dimensional Lebesgue measure on K) and*

$$\text{for every } x_1 \in \operatorname{int}(K) \text{ the map}$$
$$D_{x_1} \ni [y,q] \mapsto \begin{pmatrix} \partial_1 H(x_1,1;y,q) \\ \partial_2 H(x_1,1;y,q) \end{pmatrix} \in \mathbb{R}^{d+1} \text{ is invertible,} \tag{5.1.15}$$

then there exists a Borel map $T : K \to \mathfrak{C}[K]$ s.t.

$$\mu_2 = (\mathfrak{h} \circ T_\sharp)(\mu_1), \qquad \int_K H([x_1,1],T(x_1))\,d\mu_1(x_1) = \mathscr{U}_H(\mu_1,\mu_2).$$

Proof. By Theorem 5.1.5, we know that there exists a pair $(\varphi_1,\varphi_2) \in \Phi_H$ with φ_1 Lipschitz continuous (see also Proposition 5.1.8) such that

$$\int_K \varphi_1\,d\mu_1 + \int_K \varphi_2\,d\mu_2 = \mathscr{U}_H(\mu_1,\mu_2).$$

If $\alpha \in \mathfrak{H}_H^1(\mu_1,\mu_2)$ (cf. (4.1.6)), then we can find a full α-measure Borel set $\Gamma \subset \mathfrak{C}[K,K]$ such that

$$\varphi_1(x_1)r_1 + \varphi_2(x_2)r_2 = H([x_1,r_1],[x_2,r_2]) \qquad \text{for every } ([x_1,r_1],[x_2,r_2]) \in \Gamma.$$

Notice that, since $H(\mathfrak{o},\mathfrak{o}) = 0$ and $H = +\infty$ outside $U_{q_1q_2}$, we can assume that $\Gamma \subset U_{q_1q_2}$ and $(\mathfrak{o},\mathfrak{o}) \notin \Gamma$. Since φ_1 is Lipschitz continuous, we can find a full μ_1-measure Borel set $U \subset \operatorname{int}(K)$ where φ_1 is differentiable (since K is convex, its boundary has 0 Lebesgue measure). Let $([\bar{x}_1,\bar{r}_1],[\bar{x}_2,\bar{r}_2]) \in \Gamma$ with $\bar{x}_1 \in U$. We have that

$$\begin{gathered}\text{the map } (x_1,r_1) \mapsto H([x_1,r_1],[\bar{x}_2,\bar{r}_2]) - \varphi_1(x_1)r_1 \\ \text{has a minimum at } (x_1,r_1) = (\bar{x}_1,\bar{r}_1).\end{gathered} \tag{5.1.16}$$

From (5.1.16) it follows that

$$\begin{cases}\partial_1 H(\bar{x}_1,\bar{r}_1;\bar{x}_2,\bar{r}_2) = \nabla\varphi_1(\bar{x}_1)\bar{r}_1, \\ \partial_2 H(\bar{x}_1,\bar{r}_1;\bar{x}_2,\bar{r}_2) = \varphi_1(\bar{x}_1)\end{cases}$$

which, using the 1-homogeneity of H, can be rewritten as

$$\begin{cases}\partial_1 H(\bar{x}_1,1;\bar{x}_2,\bar{r}_2/\bar{r}_1) = \nabla\varphi_1(\bar{x}_1), \\ \partial_2 H(\bar{x}_1,1;\bar{x}_2,\bar{r}_2/\bar{r}_1) = \varphi_1(\bar{x}_1).\end{cases}$$

Let us denote by $I_{x_1} \subset \mathbb{R}^{d+1}$ the image of the map in (5.1.15) and by $f_{x_1} : I_{x_1} \to \mathfrak{C}[K]$ the inverse of such map. If we define $T : X_1 \to \mathfrak{C}[K]$ as

$$T(x_1) := \begin{cases} f(\nabla\varphi_1(x_1),\varphi_1(x_1)) & \text{if } (\nabla\varphi_1(x_1),\varphi_1(x_1) \in I_{x_1}, \\ \mathfrak{o} & \text{else ,}\end{cases}$$

then T is a Borel map and we have just proven that

$$[\bar{x}_2,\bar{r}_2/\bar{r}_1] = T(\bar{x}_1),$$

meaning that

$$r_1 T(x_1) = [x_2,r_2] \quad \text{for every } ([x_1,r_1],[x_2,r_2] \in \Gamma,\ x_1 \in U.$$

From this and the fact that $\Gamma \cap \mathbf{x}_1^{-1}(U)$ has full $\boldsymbol{\alpha}$-measure (recall that $\{\mathbf{r}_1 = 0\} \cap \Gamma = \emptyset$) it follows that

$$\boldsymbol{\alpha} = ([\mathrm{id}_{X_1},1],T)_\sharp \mu_1$$

which leads to the conclusion. □

Remark 5.1.10. Let us suppose that (X_1,d_1), (X_2,d_2) are non compact metric spaces, $\mu_i \in \mathcal{M}_+(X_i)$ for $i = 1,2$ and $H : \mathfrak{C}[X_1,X_2] \to [0,+\infty]$ is a 1-homogeneous, convex and continuous function. We assume that the supports $\tilde{X}_i := \operatorname{supp}\mu_i$, $i = 1,2$ are nonempty compact sets and one between the settings described in (5.1.2), (5.1.3) or (5.1.11), (5.1.12) is satisfied for the restriction $\tilde{H}$ of H to $\mathfrak{C}[\tilde{X}_1,\tilde{X}_2]$ and the restrictions $\tilde{\mu}_i$ of μ_i to $\tilde{X}_i$, for $i = 1,2$. In both situations, using Theorem 5.1.5, we obtain the existence of two continuous and bounded functions $\tilde{\varphi}_i : \tilde{X}_i \to \mathbb{R}$, $i = 1,2$, such that

$$\tilde{\varphi}_1(x_1)r_1 + \tilde{\varphi}_2(x_2) \leqslant H([x_1,r_1],[x_2,r_2]) \tag{5.1.17}$$

for every $(x_1, x_2) \in \tilde{X}_1 \times \tilde{X}_2$ and $r_1, r_2 \geqslant 0$, and

$$\int_{\tilde{X}_1} \tilde{\varphi}_1 \, d\tilde{\mu}_1 + \int_{\tilde{X}_2} \tilde{\varphi}_2 \, d\tilde{\mu}_2 = \min \left\{ \int_{\mathfrak{C}[\tilde{X}_1, \tilde{X}_2]} \tilde{H} \, d\tilde{\boldsymbol{\alpha}} \mid \tilde{\boldsymbol{\alpha}} \in \mathfrak{H}^1(\tilde{\mu}_1, \tilde{\mu}_2) \right\}. \quad (5.1.18)$$

We can thus extend $\tilde{\varphi}_i$ to $\varphi_i : X_i \to \mathbb{R}$ setting it equal to $-\infty$ outside $\tilde{X}_i$, $i = 1, 2$. Observing that the right hand side of (5.1.18) coincides with $\mathscr{U}_H(\mu_1, \mu_2)$ we get that

$$\varphi_1(x_1) r_1 + \varphi_2(x_2) \leqslant H([x_1, r_1], [x_2, r_2])$$

for every $(x_1, x_2) \in X_1 \times X_2$ and $r_1, r_2 \geqslant 0$, and

$$\int_{X_1} \varphi_1 \, d\mu_1 + \int_{X_2} \varphi_2 \, d\mu_2 = \mathscr{U}_H(\mu_1, \mu_2)$$

so that (φ_1, φ_2) is a maximizing pair of (unbounded) upper semicontinuous functions for the dual problem (4.2.4).

5.2 GENERAL OPTIMALITY CONDITIONS

In this section we provide sufficient and necessary conditions for a plan $\boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2)$ to be optimal. In this section X_1 and X_2 are completely regular spaces and $H : \mathfrak{C}[X_1, X_2] \to [0, +\infty]$ is a 1-homogeneous, convex and lower semicontinuous function.

Definition 5.2.1. Let $B \subset \mathbb{R}^2_+$; we define *the convex cone generated by* B as

$$\hat{B} := \left\{ \left(\sum_{i=1}^N \alpha_i r_1^i, \sum_{i=1}^N \alpha_i r_2^i \right) \mid \{(r_1^i, r_2^i)\}_{i=1}^N \subset B, \{\alpha_i\}_{i=1}^N \subset [0, +\infty), N \geqslant 1 \right\}.$$

If $\Gamma \subset \mathfrak{C}[X_1, X_2]$ and $(x_1, x_2) \in X_1 \times X_2$, we define the (x_1, x_2)-*section* of Γ as

$$\Gamma_{x_1, x_2} := \left\{ (r_1, r_2) \in \mathbb{R}^2_+ \mid ([x_1, r_1], [x_2, r_2]) \in \Gamma \right\}$$

and the *convex cone generated by* Γ

$$\hat{\Gamma} := \bigcup_{(x_1, x_2) \in X_1 \times X_2} \left\{ ([x_1, r_1], [x_2, r_2]) \mid (r_1, r_2) \in \hat{\Gamma}_{x_1, x_2} \right\}. \quad (5.2.1)$$

Definition 5.2.2. Let $\Gamma \subset \mathfrak{C}[X_1, X_2]$; we say that Γ is H-*cyclically monotone* if for every finite family of points $\{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=1}^N \subset \Gamma$ and every permutation σ of $\{1, \ldots, N\}$ it holds

$$\sum_{i=1}^N H(\mathfrak{y}_1^i, \mathfrak{y}_2^i) \leqslant \sum_{i=1}^N H(\mathfrak{y}_1^i, \mathfrak{y}_2^{\sigma(i)}).$$

Proposition 5.2.3 (Necessity of cyclical monotonicity). *Let* $\mu_i \in \mathcal{M}_+(X_i)$ *for* $i = 1, 2$, *let* $\boldsymbol{\alpha} \in \mathfrak{H}^1_H(\mu_1, \mu_2)$ *be optimal and suppose that* $\int_{\mathfrak{C}[X_1, X_2]} H \, d\boldsymbol{\alpha} < +\infty$. *Then* $\boldsymbol{\alpha}$ *is concentrated on a Borel subset* $\Gamma \subset \mathfrak{C}[X_1, X_2]$ *s.t.* $\hat{\Gamma}$ *is* H-*cyclically monotone.*

Proof. Let $\{(\varphi_1^k, \varphi_2^k)\}_{k\geqslant 1} \subset \Phi_H$ be a maximizing sequence for the dual problem (4.2.4) and let us define

$$H_k([x_1,r_1],[x_2,r_2]) := H([x_1,r_1];[x_2,r_2]) - \varphi_1^k(x_1)r_1 - \varphi_2^k(x_2)r_2.$$

Then there exist a subsequence $m \mapsto k(m)$ and a Borel subset $\Gamma \subset \mathfrak{C}[X_1,X_2]$ on which α is concentrated s.t. $H_{k(m)} \to 0$ on Γ as $m \to +\infty$. Since $H_{k(m)}$ is a convex function, this convergence takes place also on the set

$$\tilde{\Gamma} := \bigcup_{(x_1,x_2)\in X_1\times X_2} \{([x_1,r_1],[x_2,r_2]) \mid (r_1,r_2) \in \mathrm{co}\,(\Gamma_{x_1,x_2})\} \subset \hat{\Gamma}.$$

Observe that points of $\hat{\Gamma}$ are of the form $(\lambda\mathfrak{y}_1, \lambda\mathfrak{y}_2)$ for some $\lambda \geqslant 0$ and some $(\mathfrak{y}_1,\mathfrak{y}_2) \in \tilde{\Gamma}$. Let now $\{(\mathfrak{y}_1^i,\mathfrak{y}_2^i)\}_{i=1}^N \subset \Gamma$ be a finite family of points and let σ be a permutation of $\{1,\dots,N\}$. We can thus find $\{\lambda_i\}_{i=1}^N \subset [0,+\infty)$ and $\{(\mathfrak{w}_i^1,\mathfrak{w}_i^2)\}_{i=1}^N = \{([x_1^i,r_1^i],[x_2^i,r_2^i])\}_{i=1}^N \subset \tilde{\Gamma}$ such that $(\mathfrak{y}_1^i,\mathfrak{y}_2^i) = (\lambda_i\mathfrak{w}_1^i, \lambda_i\mathfrak{w}_2^i)$ for every $i = 1,\dots,N$. Then

$$\begin{aligned}
\sum_{i=1}^N H(\mathfrak{y}_1^i,\mathfrak{y}_2^{\sigma(i)}) &\geqslant \sum_{i=1}^N \left(\lambda_i r_1^i \varphi_1^{k(m)}(x_1^i) + \lambda_{\sigma(i)} r_2^{\sigma(i)} \varphi_2^{k(m)}(x_2^{\sigma(i)})\right)\\
&= \sum_{i=1}^N \left(\lambda_i r_1^i \varphi_1^{k(m)}(x_1^i) + \lambda_i r_2^i \varphi_2^{k(m)}(x_2^i)\right)\\
&= \sum_{i=1}^N \left(H(\mathfrak{y}_1^i,\mathfrak{y}_2^i) - \lambda_i H_{k(m)}(\mathfrak{w}_1^i,\mathfrak{w}_2^i)\right).
\end{aligned}$$

Letting $m \to +\infty$, we obtain the sought H-cyclical monotonicity of $\hat{\Gamma}$. $\square$

In the next statements, given $(x_1,x_2) \in X_1 \times X_2$, we denote by $\partial H_{x_1,x_2}(r_1,r_2)$ the subdifferential of H_{x_1,x_2} at a point $(r_1,r_2) \in \mathbb{R}_+^2$ of its domain, defined as

$$\partial H_{x_1,x_2}(r_1,r_2) := \left\{ (a,b) \in \mathbb{R}^2 \;\middle|\; \begin{array}{l} H_{x_1,x_2}(s_1,s_2) - H_{x_1,x_2}(r_1,r_2) \geqslant \\ a(s_1-r_1) + b(s_2-r_2) \\ \text{for every } (s_1,s_2) \in \mathbb{R}_+^2 \end{array} \right\}.$$

The domain of the subdifferential is denoted by

$$D(\partial H_{x_1,x_2}) := \left\{(r_1,r_2) \in \mathbb{R}_+^2 \mid \partial H_{x_1,x_2}(r_1,r_2) \neq \emptyset\right\}$$

and its interior part $\mathrm{int}\,(D(\partial H_{x_1,x_2}))$ is taken with respect to the topology of $\mathbb{R}^2$ so that, in particular, if $(r_1,r_2) \in \mathrm{int}\,(D(\partial H_{x_1,x_2}))$, then $r_1 > 0, r_2 > 0$.

Proposition 5.2.4. *Let $\Gamma \subset \mathfrak{C}[X_1,X_2]$ be such that $\hat{\Gamma}$ is H-cyclically monotone and suppose that there exists $([x_1,r_1],[x_2,r_2]) \in \hat{\Gamma}$ such that $(r_1,r_2) \in \mathrm{int}\,(D(\partial H_{x_1,x_2}))$. Then there exists $(\bar{\varphi}_1^+,\bar{\varphi}_1^-) \in \mathbb{R}^2$ such that*

$$H([x_1,r],\mathfrak{y}_2^N) - H(\mathfrak{y}_1^N,\mathfrak{y}_2^N) + \sum_{i=0}^{N-1} \left(H(\mathfrak{y}_1^{i+1},\mathfrak{y}_2^i) - H(\mathfrak{y}_1^i,\mathfrak{y}_2^i)\right) + \bar{\varphi}_1^+ r(\mathfrak{y}_1^0) \geqslant \bar{\varphi}_1^- r \tag{5.2.2}$$

for every $r > 0$ and every finite family of points $\{(\mathfrak{y}_1^i,\mathfrak{y}_2^i)\}_{i=0}^N \subset \hat{\Gamma}$ with $(\mathfrak{y}_1^0,\mathfrak{y}_2^0) \in \{(\lambda[x_1,r_1],\lambda[x_2,r_2]) \mid \lambda \geqslant 0\}$.

Proof. Let $r > 0$ and $\{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N \subset \hat{\Gamma}$ be as in the statement and let us set $\bar{H} := H_{x_1,x_2}$, $q := \frac{r_2}{r_1} > 0$ and $\mathfrak{y}_1^{N+1} := [x_1, r]$. For every $k \in \mathbb{N}$, we define $\vartheta_k := q\left(\frac{r(\mathfrak{y}_1^0)}{r}\right)^{1/k}$ and we consider the points

$$\mathfrak{y}_1^{N+1+n} := \left[x_1, r\left(\frac{\vartheta_k}{q}\right)^n\right], \quad \mathfrak{y}_2^{N+n} := \left[x_1, rq\left(\frac{\vartheta_k}{q}\right)^{n-1}\right], \quad n = 1, \dots, k.$$

Notice that

$$\mathfrak{y}_1^{N+k+1} = [x_1, r] = \mathfrak{y}_1^0, \quad (\mathfrak{y}_1^{N+n}, \mathfrak{y}_2^{N+n}) = (\rho_{k,n}[x_1, r_1], \rho_{k,n}[x_2, r_2]) \in \hat{\Gamma}$$

for every $n = 1, \dots, k$, where $\rho_{k,n} = \frac{r}{r_1}\left(\frac{\vartheta_k}{q}\right)^n$. We define the two quantities

$$A := \sum_{i=0}^N \left(H(\mathfrak{y}_1^{i+1}, \mathfrak{y}_2^i) - H(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\right), \quad B := \sum_{n=1}^k \left(H(\mathfrak{y}_1^{N+n+1}, \mathfrak{y}_2^{N+n}) - H(\mathfrak{y}_1^{N+n}, \mathfrak{y}_2^{N+n})\right)$$

and we notice that, by H-cyclical monotonicity, we have that $A + B \geqslant 0$. Hence, in order to bound A from below, it is enough to bound $-B$ from below:

$$\begin{aligned}
-B &= -\sum_{n=1}^k \left(H(\mathfrak{y}_1^{N+n+1}, \mathfrak{y}_2^{N+n}) - H(\mathfrak{y}_1^{N+n}, \mathfrak{y}_2^{N+n})\right) \\
&= \sum_{n=1}^k \left(\bar{H}\left(r\left(\frac{\vartheta_k}{q}\right)^{n-1}, rq\left(\frac{\vartheta_k}{q}\right)^{n-1}\right)\right. \\
&\qquad \left. - \bar{H}\left(r\left(\frac{\vartheta_k}{q}\right)^{n-1}\frac{\vartheta_k}{q}, rq\left(\frac{\vartheta_k}{q}\right)^{n-1}\right)\right) \\
&= \sum_{n=1}^k r\left(\frac{\vartheta_k}{q}\right)^{n-1}\left(\bar{H}(1,q) - \bar{H}\left(\frac{\vartheta_k}{q}, q\right)\right) \\
&\geqslant \sum_{n=1}^k r\left(\frac{\vartheta_k}{q}\right)^{n-1}\left(1 - \frac{\vartheta_k}{q}\right)\bar{\varphi}_1^k \\
&= r\bar{\varphi}_1^k\left(1 - \frac{r(\mathfrak{y}_1^0)}{r}\right) \\
&= \bar{\varphi}_1^k(r - r(\mathfrak{y}_1^0))
\end{aligned}$$

where $(\bar{\varphi}_1^k, \bar{\varphi}_2^k) \in \partial\bar{H}\left(\frac{\theta_k}{q}, q\right) \neq \emptyset$ for k sufficiently large, since $\left(\frac{\theta_k}{q}, q\right) \to (1, q)$ as $k \to +\infty$ and $\partial\bar{H}(1, q) = \partial\bar{H}(r_1, r_2)$ with $(r_1, r_2) \in \operatorname{int}\left(D(\partial\bar{H})\right)$. We have proven that, for $k \in \mathbb{N}$ large enough, there exists $(\bar{\varphi}_1^k, \bar{\varphi}_2^k) \in \partial\bar{H}\left(\frac{\theta_k}{q}, q\right)$ such that

$$A + \bar{\varphi}_1^k r(\mathfrak{y}_1^0) \geqslant \bar{\varphi}_1^k r.$$

Let us define

$$\bar{\varphi}_1^+ := \lim_{h \downarrow 0} \frac{\bar{H}(r_1 + h, r_2) - \bar{H}(r_1, r_2)}{h},$$

$$\bar{\varphi}_1^- := \lim_{h \uparrow 0} \frac{\bar{H}(r_1 + h, r_2) - \bar{H}(r_1, r_2)}{h}.$$

Up to passing to a subsequence, we can assume that $(\bar{\varphi}_1^k, \bar{\varphi}_2^k) \to (\tilde{\varphi}_1, \tilde{\varphi}_2) \in \partial\bar{H}(r_1, r_2)$. Observe that, if $r(\mathfrak{y}_1^0) \geqslant r$, we have $\tilde{\varphi}_1 = \bar{\varphi}_1^+$, while, if $r(\mathfrak{y}_1^0) \leqslant r$, we have that $\tilde{\varphi}_1 = \bar{\varphi}_1^-$ so that

$$A + \bar{\varphi}_1^+ r(\mathfrak{y}_1^0) \geqslant A + \tilde{\varphi}_1 r(\mathfrak{y}_1^0) \geqslant \tilde{\varphi}_1 r \geqslant \bar{\varphi}_1^- r.$$

□

Theorem 5.2.5 (Sufficiency of cyclical monotonicity)**.** *Let $\mu_i \in \mathcal{M}_+(X_i)$ for $i = 1, 2$, let $\boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2)$ be an admissible plan concentrated on a Borel subset Γ such that $\hat{\Gamma} \subset D(H)$ and suppose that the effective domain of H is independent of $(x_1, x_2) \in X_1 \times X_2$, meaning that $D(H_{x_1,x_2}) = D(H_{y_1,y_2})$ for every $(x_1, x_2), (y_1, y_2) \in X_1 \times X_2$. Suppose moreover that the following conditions are satisfied:*

1. *there exists $([\bar{x}_1, \bar{r}_1], [\bar{x}_2, \bar{r}_2]) \in \hat{\Gamma}$ such that $(\bar{r}_1, \bar{r}_2) \in \operatorname{int}(D(\partial H_{\bar{x}_1,\bar{x}_2}))$;*
2. *there exist positive constants a_i, b_i, $i = 1, 2$ s.t.*

$$\begin{aligned} &\mu_1\left(\left\{x_1 \in X_1 \mid \int_{X_2} H([x_1, a_1]; [x_2, b_1])\, d\mu_2(x_2) < +\infty\right\}\right) > 0, \\ &\mu_2\left(\left\{x_2 \in X_2 \mid \int_{X_1} H([x_1, a_2]; [x_2, b_2])\, d\mu_1(x_1) < +\infty\right\}\right) > 0. \end{aligned} \tag{5.2.3}$$

If $\hat{\Gamma}$ is H-cyclically monotone, then $\boldsymbol{\alpha}$ is optimal, $\int_{\mathfrak{C}[X_1,X_2]} H\, d\boldsymbol{\alpha} < +\infty$ and there exists a maximizing pair $(\varphi_1, \varphi_2) \in L^1(X_1, \mu_1; \bar{\mathbb{R}}) \times L^1(X_2, \mu_2; \bar{\mathbb{R}})$ for the dual problem i.e.

$$\int_{X_1} \varphi_1\, d\mu_1 + \int_{X_2} \varphi_2\, d\mu_2 = \mathscr{U}_H(\mu_1, \mu_2).$$

Proof. Let us define

$$S_1 := \mathfrak{o} \times (\mathfrak{C}[X_2] \setminus \{\mathfrak{o}\}), \quad S_2 := (\mathfrak{C}[X_1] \setminus \{\mathfrak{o}\}) \times \mathfrak{o}.$$

Since the effective domain of H is independent of $(x_1, x_2) \in X_1 \times X_2$ and H is 1-homogeneous, there are only two possibilities for the value of H on S_i for $i = 1, 2$: either H is infinite on the whole S_i or it is finite on the whole S_i, for every $i = 1, 2$. There are thus four possible cases, but since the statement does not depend on the order of X_1 and X_2, we have to deal actually with only three possibilities:

(i) $H(\mathfrak{y}_1, \mathfrak{o}) = H(\mathfrak{o}, \mathfrak{y}_2) = +\infty$ for every $\mathfrak{y}_1 \in \mathfrak{C}[X_1] \setminus \{\mathfrak{o}\}$ and every $\mathfrak{y}_2 \in \mathfrak{C}[X_2] \setminus \{\mathfrak{o}\}$,

(ii) $H(\mathfrak{o}, \mathfrak{y}_2) = +\infty$ for every $\mathfrak{y}_2 \in \mathfrak{C}[X_2] \setminus \{\mathfrak{o}\}$ and $H(\mathfrak{y}_1, \mathfrak{o}) < +\infty$ for every $\mathfrak{y}_1 \in \mathfrak{C}[X_1]$

(iii) $H(\mathfrak{y}_1, \mathfrak{y}_2) < +\infty$ for every $(\mathfrak{y}_1, \mathfrak{y}_2) \in \mathfrak{C}[X_1, X_2]$,

since, by convexity, if H is finite both on S_1 and S_2, it is finite everywhere. We group these three possibilities in two cases:

(a) $H(\mathfrak{o}, \mathfrak{y}_2) = +\infty$ for every $\mathfrak{y}_2 \in \mathfrak{C}[X_2] \setminus \{\mathfrak{o}\}$,

(b) $H(\mathfrak{y}_1, \mathfrak{y}_2) < +\infty$ for every $(\mathfrak{y}_1, \mathfrak{y}_2) \in \mathfrak{C}[X_1, X_2]$.

Notice that in case (a) $(\mathfrak{o}, \mathfrak{y}_2) \notin \hat{\Gamma}$ if $\mathfrak{y}_2 \neq \mathfrak{o}$.
We will carry out the proof under assumption (a) and point out at the end the changes that have to be done to deal with case (b).
Step 1 (Definition of Φ): For every finite family of points $\{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N \subset \hat{\Gamma}$, we define

$$\Theta\left(\mathfrak{y}_1; \{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N\right) := H(\mathfrak{y}_1, \mathfrak{y}_2^N) - H(\mathfrak{y}_1^N, \mathfrak{y}_2^N) + \sum_{i=0}^{N-1} \left(H(\mathfrak{y}_1^{i+1}, \mathfrak{y}_2^i) - H(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\right) + \bar{\varphi}_1^+ r(\mathfrak{y}_1^0),$$

for every $\mathfrak{y}_1 \in \mathfrak{C}[X_1]$, where $\bar{\varphi}_1^+$ is given by Proposition 5.2.4 with the choice $([x_1, r_1], [x_2, r_2]) := ([\bar{x}_1, \bar{r}_1], [\bar{x}_2, \bar{r}_2])$ and we use the convention that, whenever $N = 0$, the summation is equal to 0. Notice that there is no ambiguity in the definition of Θ, since H is finite on $\hat{\Gamma}$. We define $\Phi : \mathfrak{C}[X_1] \to [-\infty, +\infty]$ as

$$\Phi(\mathfrak{y}_1) := \inf\left\{ \Theta\left(\mathfrak{y}_1; \{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N\right) \;\middle|\; \begin{array}{l} \{\mathfrak{y}_1^i, \mathfrak{y}_2^i\}_{i=0}^N \subset \hat{\Gamma}, \\ (\mathfrak{y}_1^0, \mathfrak{y}_2^0) \in \{(\lambda[\bar{x}_1, \bar{r}_1], \lambda[\bar{x}_2, \bar{r}_2]) \mid \lambda \geq 0\}, \\ N \in \mathbb{N}. \end{array} \right\},$$

for every $\mathfrak{y}_1 \in \mathfrak{C}[X_1]$. For every $\mathfrak{y}_1 \in \mathfrak{C}[X_1]$, we can choose $N = 0$ and $(\mathfrak{y}_1^0, \mathfrak{y}_2^0) = (\frac{r(\mathfrak{y}_1)}{\bar{r}_1}[\bar{x}_1, \bar{r}_1], \frac{r(\mathfrak{y}_1)}{\bar{r}_1}[\bar{x}_2, \bar{r}_2])$, obtaining that

$$\begin{aligned} \Phi(\mathfrak{y}_1) &\leq \Theta\left(\mathfrak{y}_1; \{(\mathfrak{y}_1^0, \mathfrak{y}_2^0)\}\right) \\ &= \frac{r(\mathfrak{y}_1)}{\bar{r}_1} H([x(\mathfrak{y}_1), \bar{r}_1], [\bar{x}_2, \bar{r}_2]) - H(\mathfrak{y}_1^0, \mathfrak{y}_2^0) + \bar{\varphi}_1^+ r(\mathfrak{y}_1^0) < +\infty \end{aligned} \tag{5.2.4}$$

since the domain of H is independent of $(x_1, x_2) \in X_1 \times X_2$ and $H([\bar{x}_1, \bar{r}_1], [\bar{x}_2, \bar{r}_2]) < +\infty$, being $[\bar{x}_1, \bar{r}_1], [\bar{x}_2, \bar{r}_2] \in \hat{\Gamma}$. In particular, for every $r_1 \geq 0$, we can choose $\mathfrak{y}_1 = [\bar{x}_1, r_1]$ in (5.2.4) obtaining that

$$\Phi([\bar{x}_1, r_1]) \leq \bar{\varphi}_1^+ r_1.$$

Choosing again $N = 0$ and $(\mathfrak{y}_1^0, \mathfrak{y}_2^0) = (\mathfrak{o}, \mathfrak{o})$ we also get

$$\Phi(\mathfrak{y}_1) \leq H(\mathfrak{y}_1, \mathfrak{o}) - H(\mathfrak{o}, \mathfrak{o}) = H(\mathfrak{y}_1, \mathfrak{o}), \quad \text{for every } \mathfrak{y}_1 \in \mathfrak{C}[X_1]. \tag{5.2.5}$$

Finally, again by Proposition 5.2.4, we have that

$$\Phi([\bar{x}_1, r_1]) \geq \bar{\varphi}_1^- r_1 \quad \text{for every } r_1 > 0. \tag{5.2.6}$$

Since for every finite family of points $\{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N \subset \hat{\Gamma}$ we have that

$$\Theta\left(\lambda\mathfrak{y}_1; \{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N\right) = \lambda\Theta\left(\mathfrak{y}_1; \{(\lambda^{-1}\mathfrak{y}_1^i, \lambda^{-1}\mathfrak{y}_2^i)\}_{i=0}^N\right) \quad \text{for every } \mathfrak{y}_1 \in \mathfrak{C}[X_1], \lambda > 0$$

and $\hat{\Gamma}$ is invariant by dialations, we get that

$$\Phi(\lambda\mathfrak{y}_1) = \lambda\Phi(\mathfrak{y}_1)$$

for every $\mathfrak{y}_1 \in \mathfrak{C}[X_1]$ and $\lambda > 0$. This shows in particular that $\Phi(\mathfrak{o})$ can be only equal to 0 or $-\infty$. Arguing as in [5, Step 1 of Therem 6.14] we can see that

$$\Phi(\mathfrak{y}_1) = \lim_p \lim_m \lim_l \Phi_{p,m,l}(\mathfrak{y}_1) \quad \text{for every } \mathfrak{y}_1 \in \mathfrak{C}[X_1],$$

where $\Phi_{p,m,l}$ are suitable upper semicontinuous functions so that Φ is a Borel function.
Given $\mathfrak{y}_1' \in \mathfrak{C}[X_1]$, $(\mathfrak{y}_1, \mathfrak{y}_2) \in \hat{\Gamma}$ and any finite family of points $\{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N \subset \hat{\Gamma}$ with $(\mathfrak{y}_1^0, \mathfrak{y}_2^0) \in \{(\lambda[\bar{x}_1, \bar{r}_1], \lambda[\bar{x}_2, \bar{r}_2]) \mid \lambda \geqslant 0\}$, we have, if we set $(\mathfrak{y}_1^{N+1}, \mathfrak{y}_2^{N+1}) := (\mathfrak{y}_1, \mathfrak{y}_2)$, that the finite family of points $\{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^{N+1}$ is still contained in $\hat{\Gamma}$ and of course satisfies $(\mathfrak{y}_1^0, \mathfrak{y}_2^0) \in \{(\lambda[\bar{x}_1, \bar{r}_1], \lambda[\bar{x}_2, \bar{r}_2]) \mid \lambda \geqslant 0\}$; thus

$$\Phi(\mathfrak{y}_1') \leqslant H(\mathfrak{y}_1', \mathfrak{y}_2) - H(\mathfrak{y}_1, \mathfrak{y}_2) + \Theta\left(\mathfrak{y}_1; \{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N\right). \tag{5.2.7}$$

If $\mathfrak{y}_1 \in \pi^{\mathfrak{C}_1}(\hat{\Gamma}) \setminus \{\mathfrak{o}\}$, we can find $\mathfrak{y}_2 \in \mathfrak{C}[X_2]$ such that $(\mathfrak{y}_1, \mathfrak{y}_2) \in \hat{\Gamma}$ and plug it into (5.2.7) with $\mathfrak{y}_1' = [\bar{x}_1, r(\mathfrak{y}_1)]$. Passing then to the infimum among the admissible finite family of points, we get, also using (5.2.6), that

$$\bar{\varphi}_1^- r(\mathfrak{y}_1) \leqslant \Phi([\bar{x}_1, r(\mathfrak{y}_1)]) \leqslant H([\bar{x}_1, r(\mathfrak{y}_1)], \mathfrak{y}_2) - H(\mathfrak{y}_1, \mathfrak{y}_2) + \Phi(\mathfrak{y}_1).$$

Thus, noticing that $H([\bar{x}_1, r(\mathfrak{y}_1)], \mathfrak{y}_2) < +\infty$ thanks to the independence of the effective domain of H from $(x_1, x_2) \in X_1 \times X_2$, we get that $\Phi(\mathfrak{y}_1) \in \mathbb{R}$.
Summarizing the first step, we have proven that there exists a Borel function $\Phi : \mathfrak{C}[X_1] \to [-\infty, +\infty)$ such that

$$\Phi([\bar{x}_1, r_1]) \geqslant \bar{\varphi}_1^- r_1 \qquad \text{for every } r_1 > 0, \tag{5.2.8}$$
$$\Phi(\mathfrak{y}_1) \leqslant H(\mathfrak{y}_1, \mathfrak{o}) \qquad \text{for every } \mathfrak{y}_1 \in \mathfrak{C}[X_1], \tag{5.2.9}$$
$$\Phi(\lambda\mathfrak{y}_1) = \lambda\Phi(\mathfrak{y}_1) \qquad \text{for every } \mathfrak{y}_1 \in \mathfrak{C}[X_1],\ \lambda > 0, \tag{5.2.10}$$
$$\Phi(\mathfrak{y}_1) \in \mathbb{R} \qquad \text{for every } \mathfrak{y}_1 \in \pi^{\mathfrak{C}_1}(\hat{\Gamma}) \setminus \{\mathfrak{o}\} \tag{5.2.11}$$

and for every finite family of points $\{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N \subset \hat{\Gamma}$ with

$$(\mathfrak{y}_1^0, \mathfrak{y}_2^0) \in \{(\lambda[\bar{x}_1, \bar{r}_1], \lambda[\bar{x}_2, \bar{r}_2]) \mid \lambda \geqslant 0\}$$

it holds

$$\Phi(\mathfrak{y}_1') \leqslant H(\mathfrak{y}_1', \mathfrak{y}_2) - H(\mathfrak{y}_1, \mathfrak{y}_2) + \Theta\left(\mathfrak{y}_1; \{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N\right) \tag{5.2.12}$$

for every $\mathfrak{y}_1' \in \mathfrak{C}[X_1]$ and every $(\mathfrak{y}_1, \mathfrak{y}_2) \in \hat{\Gamma}$.
Step 2 (Definition of Ψ): We define $\Psi : \mathfrak{C}[X_2] \to [-\infty, +\infty]$ as

$$\Psi(\mathfrak{y}_2) := \inf_{\mathfrak{y}_1 \in \mathfrak{C}[X_1]} \{H(\mathfrak{y}_1, \mathfrak{y}_2) - \Phi(\mathfrak{y}_1)\}.$$

It is clear from the definition that

$$\Psi(\mathfrak{y}_2) \leqslant H(\mathfrak{y}_1, \mathfrak{y}_2) - \Phi(\mathfrak{y}_1) \quad \text{for every } (\mathfrak{y}_1, \mathfrak{y}_2) \in \mathfrak{C}[X_1, X_2]$$

and by (5.2.11) we deduce that

$$\Psi(\mathfrak{y}_2) + \Phi(\mathfrak{y}_1) \leqslant H(\mathfrak{y}_1, \mathfrak{y}_2) \tag{5.2.13}$$

for every $\mathfrak{y}_1 \in \pi^{\mathfrak{C}_1}(\hat{\Gamma}) \setminus \{\mathfrak{o}\}$ and every $\mathfrak{y}_2 \in \mathfrak{C}[X_2]$. By the definition of Ψ and (5.2.10) it also easily follows that

$$\Psi(\lambda \mathfrak{y}_2) = \lambda \Psi(\mathfrak{y}_2) \quad \text{for every } \mathfrak{y}_2 \in \mathfrak{C}[X_2],\ \lambda > 0$$

so that $\Psi(\mathfrak{o}) \in \{-\infty, 0, +\infty\}$. Moreover, if $\mathfrak{y}_2 \in \mathfrak{C}[X_2] \setminus \{\mathfrak{o}\}$, we have that

$$\Psi(\mathfrak{y}_2) \leqslant H\left(\frac{r(\mathfrak{y}_2)}{\bar{r}_2}[\bar{x}_1, \bar{r}_1], \mathfrak{y}_2\right) - \bar{\varphi}_1^- \frac{r(\mathfrak{y}_2)}{\bar{r}_2}\bar{r}_1 < +\infty, \tag{5.2.14}$$

where we have used (5.2.8) and the fact that $([\bar{x}_1, \bar{r}_1], [\bar{x}_2, \bar{r}_2]) \in D(H)$ together with the independence of the effective domain of H from $(x_1, x_2) \in X_1 \times X_2$. If $(\mathfrak{y}_1, \mathfrak{y}_2) \in \hat{\Gamma} \setminus \{(\mathfrak{o}, \mathfrak{o})\}$, it must be that $\mathfrak{y}_1 \neq \mathfrak{o}$, since $[\mathfrak{o}, \mathfrak{y}_2] \notin \hat{\Gamma}$ if $\mathfrak{y}_2 \neq \mathfrak{o}$. We thus know by (5.2.11) that $\Phi(\mathfrak{y}_1) \in \mathbb{R}$; if $\Psi(\mathfrak{y}_2) = +\infty$ the inequality

$$H(\mathfrak{y}_1, \mathfrak{y}_2) \leqslant \Phi(\mathfrak{y}_1) + \Psi(\mathfrak{y}_2)$$

is trivially satisfied. If $\Psi(\mathfrak{y}_2) < +\infty$ we can write (5.2.12) for those $\mathfrak{y}_1' \in \mathfrak{C}[X_1]$ and finite families of points $\{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N \subset \hat{\Gamma}$ with $(\mathfrak{y}_1^0, \mathfrak{y}_2^0) \in \{(\lambda[\bar{x}_1, \bar{r}_1], \lambda[\bar{x}_2, \bar{r}_2]) \mid \lambda \geqslant 0\}$ such that $H(\mathfrak{y}_1', \mathfrak{y}_2) - \Phi(\mathfrak{y}_1') < +\infty$ and $\Theta\left(\mathfrak{y}_1; \{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N\right) \in \mathbb{R}$ obtaining

$$H(\mathfrak{y}_1, \mathfrak{y}_2) \leqslant H(\mathfrak{y}_1', \mathfrak{y}_2) - \Phi(\mathfrak{y}_1') + \Theta\left(\mathfrak{y}_1; \{(\mathfrak{y}_1^i, \mathfrak{y}_2^i)\}_{i=0}^N\right).$$

Passing to the infimum w.r.t.$\mathfrak{y}_1'$ and the families of points we get again

$$H(\mathfrak{y}_1, \mathfrak{y}_2) \leqslant \Phi(\mathfrak{y}_1) + \Psi(\mathfrak{y}_2).$$

This, together with (5.2.13), proves that

$$H(\mathfrak{y}_1, \mathfrak{y}_2) = \Psi(\mathfrak{y}_2) + \Phi(\mathfrak{y}_1) \quad \text{for every } (\mathfrak{y}_1, \mathfrak{y}_2) \in \hat{\Gamma} \setminus \{(\mathfrak{o}, \mathfrak{o})\}$$

and also gives that $\Psi(\mathfrak{y}_2) \in \mathbb{R}$ for every $\mathfrak{y}_2 \in \pi^{\mathfrak{C}_2}(\hat{\Gamma}) \setminus \{\mathfrak{o}\}$. The Borel measurability of Ψ can be checked as in [5, Step 2 of Theorem 6.14].
Summarizing the second step, we have proven that there exists a Borel function $\Psi : \mathfrak{C}[X_2] \to [-\infty, +\infty]$ such that

$$\Psi \leqslant H - \Phi \quad \text{on } \mathfrak{C}[X_1, X_2], \tag{5.2.15}$$
$$\Psi(\lambda \mathfrak{y}_2) = \lambda \Psi(\mathfrak{y}_2) \quad \text{for every } \mathfrak{y}_2 \in \mathfrak{C}[X_2],\ \lambda > 0, \tag{5.2.16}$$
$$\Psi(\mathfrak{y}_2) < +\infty \quad \text{for every } \mathfrak{y}_2 \in \mathfrak{C}[X_2] \setminus \{\mathfrak{o}\}, \tag{5.2.17}$$
$$H(\mathfrak{y}_1, \mathfrak{y}_2) = \Psi(\mathfrak{y}_2) + \Phi(\mathfrak{y}_1) \quad \text{for every } (\mathfrak{y}_1, \mathfrak{y}_2) \in \hat{\Gamma} \setminus \{(\mathfrak{o}, \mathfrak{o})\}, \tag{5.2.18}$$
$$\Psi(\mathfrak{y}_2) \in \mathbb{R} \quad \text{for every } \mathfrak{y}_2 \in \pi^{\mathfrak{C}_2}(\hat{\Gamma}) \setminus \{\mathfrak{o}\}. \tag{5.2.19}$$

Step 3 (Definition of φ_1 and φ_2): Let us define $\varphi_1 : \mathfrak{C}[X_1] \to [-\infty, +\infty)$ and $\varphi_2 : \mathfrak{C}[X_2] \to [-\infty, +\infty]$ as

$$\varphi_1(x_1) := \Phi([x_1, 1]), \quad \varphi_2(x_2) := \Psi([x_2, 1]), \quad x_1 \in X_1,\ x_2 \in X_2.$$

Notice that φ_1, φ_2 are Borel functions and $\varphi(x_2) < +\infty$ for every $x_2 \in X_2$ thanks to (5.2.17). We claim that

$$\varphi_1(x_1)r_1 + \varphi_2(x_2)r_2 \leqslant H([x_1, r_1], [x_2, r_2]), \quad ([x_1, r_1], [x_2, r_2]) \in \mathfrak{C}[X_1, X_2], \tag{5.2.20}$$

$$\varphi_1(x_1)r_1 + \varphi_2(x_2)r_2 = H([x_1, r_1], [x_2, r_2]), \quad ([x_1, r_1], [x_2, r_2]) \in \hat{\Gamma}, \tag{5.2.21}$$

where we are adopting again the convention that $0 \cdot \infty = 0$. We start from (5.2.20) and we distinguish four cases for a general point $([x_1, r_1], [x_2, r_2]) \in \mathfrak{C}[X_1, X_2]$:

(i) if $([x_1, r_1], [x_2, r_2]) = (\mathfrak{o}, \mathfrak{o})$, both sides are equal to 0;

(ii) if $[x_i, r_i] \neq \mathfrak{o}$ for $i = 1, 2$, then $\Phi([x_1, r_1]) = r_1 \varphi_1(x_1)$ and $\Psi([x_2, r_2]) = r_2 \varphi_2(x_2)$ by (5.2.10) and (5.2.16). Moreover $\Psi([x_2, r_2]) < +\infty$ by (5.2.17) so that (5.2.15) becomes (5.2.20);

(iii) if $[x_1, r_1] \neq \mathfrak{o}$ and $[x_2, r_2] = \mathfrak{o}$, (5.2.20) is exactly (5.2.9) since, by (5.2.10), we have that $\Phi([x_1, r_1]) = r_1 \varphi_1(x_1)$;

(iv) if $[x_1, r_1] = \mathfrak{o}$ and $[x_2, r_2] \neq \mathfrak{o}$ there is nothing to prove, since $H(\mathfrak{o}, [x_2, r_2]) = +\infty$;

To prove (5.2.21) we argue in the same way taking a point $([x_1, r_1], [x_2, r_2]) \in \hat{\Gamma}$, distinguishing in three cases (the case $(\mathfrak{o}, [x_2, r_2]) \in \hat{\Gamma}$ with $r_2 > 0$ is impossible due to (a)):

(i) if $([x_1, r_1], [x_2, r_2]) = (\mathfrak{o}, \mathfrak{o})$, both sides are equal to 0;

(ii) if $[x_i, r_i] \neq \mathfrak{o}$ for $i = 1, 2$, then $\Phi([x_1, r_1]) = r_1 \varphi_1(x_1)$ and $\Psi([x_2, r_2]) = r_2 \varphi_2(x_2)$ by (5.2.10) and (5.2.16). Hence (5.2.21) follows by (5.2.18);

(iii) if $[x_1, r_1] \neq \mathfrak{o}$ and $[x_2, r_2] = \mathfrak{o}$, by (5.2.10), we have that $\Phi([x_1, r_1]) = r_1 \varphi_1(x_1)$ and $r_1 \varphi_1(x_1) \in \mathbb{R}$ by (5.2.11). Thus, since H is finite on $\hat{\Gamma}$, (5.2.18) forces $\Psi(\mathfrak{o}) = 0$ and thus gives (5.2.21).

Notice that, if $x_1 \in \mathsf{x}_1(\hat{\Gamma} \setminus (\mathfrak{o} \times \mathfrak{C}[X_2]))$, then there exists $r_1 > 0$ such that $[x_1, r_1] \in \pi^{\mathfrak{C}_1}(\hat{\Gamma}) \setminus \{\mathfrak{o}\}$ so that by (5.2.10) and (5.2.11), we get that $\varphi_1(x_1) \in \mathbb{R}$. Analogously, if $x_2 \in \mathsf{x}_2(\hat{\Gamma} \setminus (\mathfrak{C}[X_1] \times \mathfrak{o}))$, then there exists $r_2 > 0$ such that $[x_2, r_2] \in \pi^{\mathfrak{C}_2}(\hat{\Gamma}) \setminus \{\mathfrak{o}\}$ so that by (5.2.16) and (5.2.19), we get that $\varphi_2(x_2) \in \mathbb{R}$. Thus, to prove that $\varphi_i(x_i) \in \mathbb{R}$ for μ_i-a.e. $x_i \in X_i$ for $i = 1, 2$, it is enough to show that

$$\mu_i \left(X_i \setminus \mathsf{x}_i \left(\hat{\Gamma} \setminus \{\mathfrak{y}_i = \mathfrak{o}\} \right) \right) = 0, \quad i = 1, 2$$

and this is a consequence of the following chain of inequalities:

$$\begin{aligned}
\mu_i \left(X_i \setminus \mathsf{x}_i \left(\hat{\Gamma} \setminus \{\mathfrak{y}_i = \mathfrak{o}\} \right) \right) &= \mathfrak{h}_i^1(\boldsymbol{\alpha}) \left(X_i \setminus \mathsf{x}_i \left(\hat{\Gamma} \setminus \{\mathfrak{y}_i = \mathfrak{o}\} \right) \right) \\
&= (\mathsf{r}_i \boldsymbol{\alpha}) \left(\mathsf{x}_i^{-1} \left(X_i \setminus \mathsf{x}_i \left(\hat{\Gamma} \setminus \{\mathfrak{y}_i = \mathfrak{o}\} \right) \right) \right) \\
&= (\mathsf{r}_i \boldsymbol{\alpha}) \left(\mathfrak{C}[X_1, X_2] \setminus \mathsf{x}_i^{-1} \left(\mathsf{x}_i \left(\hat{\Gamma} \setminus \{\mathfrak{y}_i = \mathfrak{o}\} \right) \right) \right) \\
&\leqslant (\mathsf{r}_i \boldsymbol{\alpha}) \left(\mathfrak{C}[X_1, X_2] \setminus \left(\hat{\Gamma} \setminus \{\mathfrak{y}_i = \mathfrak{o}\} \right) \right) \\
&\leqslant (\mathsf{r}_i \boldsymbol{\alpha}) \left(\mathfrak{C}[X_1, X_2] \setminus \hat{\Gamma} \right) + (\mathsf{r}_i \boldsymbol{\alpha}) \left(\hat{\Gamma} \cap \{\mathfrak{y}_i = \mathfrak{o}\} \right) \\
&\leqslant \int_{\mathfrak{C}[X_1, X_2] \setminus \Gamma} \mathsf{r}_i \, d\boldsymbol{\alpha} + \int_{\{\mathsf{r}_i = 0\}} \mathsf{r}_i \, d\boldsymbol{\alpha} \\
&= 0.
\end{aligned}$$

Summarizing the third step, we have proven that there exist two Borel functions $\varphi_i : X_i \to [-\infty, +\infty)$, $i = 1, 2$ such that $\varphi_i(x_i) \in \mathbb{R}$ for μ_i-a.e. $x_i \in X_i$, $i = 1, 2$ and satisfying

$$\begin{aligned}
\varphi_1(x_1) r_1 + \varphi_2(x_2) r_2 &\leqslant H([x_1, r_1], [x_2, r_2]), \quad ([x_1, r_1], [x_2, r_2]) \in \mathfrak{C}[X_1, X_2], \\
\varphi_1(x_1) r_1 + \varphi_2(x_2) r_2 &= H([x_1, r_1], [x_2, r_2]), \quad ([x_1, r_1], [x_2, r_2]) \in \hat{\Gamma}.
\end{aligned}$$

Step 4 (Conclusion): due to the third step and (5.2.3), we can find some $x_1 \in X_1$ such that $\varphi_1(x_1) \in \mathbb{R}$ and $\int_{X_2} H([x_1, a_2], [x_2, b_2])\, d\mu_2(x_2) < +\infty$; then, still by the third step, we get that

$$\varphi_1(x_1)a_2 + \varphi_2(x_2)b_2 \leqslant H([x_1, a_2], [x_2, b_2]) \quad \text{for every } x_2 \in X_2$$

so that

$$\varphi_2^+(x_2) \leqslant \frac{H([x_1, a_2], [x_2, b_2]) - \varphi_1^-(x_1)a_2}{b_2} \quad \text{for every } x_2 \in X_2,$$

where we denoted by u^+ and u^- the positive and negative part respectively of an extended real number $u \in [-\infty, +\infty]$. This gives that $\varphi_2^+ \in L^1(X_2, \mu_2; \bar{\mathbb{R}})$. The argument for φ_1 is the same. We can thus conclude that

$$\int_{\mathfrak{C}[X_1,X_2]} (\varphi_1(x_1)r_1 + \varphi_2(x_2)r_2)\, d\tilde{\boldsymbol{\alpha}}([x_1, r_1], [x_2, r_2]) = \int_{X_1} \varphi_1\, d\mu_1 + \int_{X_2} \varphi_2\, d\mu_2$$

belongs to $\mathbb{R} \cup \{-\infty\}$ for every $\tilde{\boldsymbol{\alpha}} \in \mathfrak{H}^1(\mu_1, \mu_2)$. Choosing $\tilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha}$ we get that $\int_{\mathfrak{C}[X_1,X_2]} H\, d\boldsymbol{\alpha} < +\infty$ and $\varphi_i \in L^1(X_i, \mu_i; \bar{\mathbb{R}})$ for $i = 1, 2$. Finally, for every $\tilde{\boldsymbol{\alpha}} \in \mathfrak{H}^1(\mu_1, \mu_2)$, we have

$$\begin{aligned}
\int_{\mathfrak{C}[X_1,X_2]} H\, d\tilde{\boldsymbol{\alpha}} &\geqslant \int_{\mathfrak{C}[X_1,X_2]} (\varphi_1(x_1)r_1 + \varphi_2(x_2)r_2)\, d\tilde{\boldsymbol{\alpha}}([x_1, r_1], [x_2, r_2]) \\
&= \int_{X_1} \varphi_1\, d\mu_1 + \int_{X_2} \varphi_2\, d\mu_2 \\
&= \int_{\mathfrak{C}[X_1,X_2]} (\varphi_1(x_1)r_1 + \varphi_2(x_2)r_2)\, d\boldsymbol{\alpha}([x_1, r_1], [x_2, r_2]) \\
&= \int_{\mathfrak{C}[X_1,X_2]} H\, d\boldsymbol{\alpha},
\end{aligned}$$

showing both that $\boldsymbol{\alpha}$ is optimal and that

$$\int_{X_1} \varphi_1\, d\mu_1 + \int_{X_2} \varphi_2\, d\mu_2 = \mathscr{U}_H(\mu_1, \mu_2).$$

We briefly summarize case (b) i.e. when $H < +\infty$. The definition of Θ is the same while we take $(\mathfrak{y}_1^0, \mathfrak{y}_2^0) = (\mathfrak{o}, \mathfrak{o})$ in the definition of Φ. By H-cyclical monotonicity it follows that $\Phi(\mathfrak{o}) \geqslant 0$. Choosing $N = 0$ and $(\mathfrak{y}_1^0, \mathfrak{y}_2^0)$ in the definition of Θ, we get (5.2.5) which now shows that $\Phi(\mathfrak{y}_1) < +\infty$ for every $\mathfrak{y}_1 \in \mathfrak{C}[X_1]$. The positive 1-homogeneity of Φ follows by the same argument and shows again that $\Phi(\mathfrak{o}) \in \{-\infty, 0\}$. However, in this case, we have shown that $\Phi(\mathfrak{o}) \geqslant 0$ so that we conclude that $\Phi(\mathfrak{o}) = 0$. The measurability of Φ follows by the same argument. (5.2.7) is obtained with the same proof and shows that $\Phi(\mathfrak{y}_1) \in \mathbb{R}$ for every $\mathfrak{y}_1 \in \pi^{\mathfrak{C}_1}(\hat{\Gamma})$ (choosing for example $\mathfrak{y}_1' = \mathfrak{o}$).
The function Ψ is defined in the same way and, choosing $\mathfrak{y}_1 = \mathfrak{o}$ one immediately sees that $\Psi(\mathfrak{y}_2) < +\infty$ for every $\mathfrak{y}_2 \in \mathfrak{C}[X_2]$ so that $\Phi + \Psi \leqslant H$ on the whole $\mathfrak{C}[X_1, X_2]$. From (5.2.7) we get that $\Phi + \Psi \geqslant H$ on $\hat{\Gamma}$, so that we get equality on $\hat{\Gamma}$.
The measurability of Φ is obtained in the same way.
The functions φ_i for $i = 1, 2$ are defined in the same way and (5.2.20) and (5.2.21) follow immediately. The conclusion is the same. □

5.3 THE METRIC VIEWPOINT

In this section we study the metric and topological properties of $\mathscr{U}_H$ under suitable hypotheses on H. We fix a completely regular space X, $p \in [1, +\infty)$ and we assume that $H : \mathfrak{C}[X, X] \to [0, +\infty]$ is a lower semicontinuous and 1-homogenous function which is the p-th power of an extended distance on $\mathfrak{C}[X]$ whose induced topology is stronger than the topology of $\mathfrak{C}[X]$.

Definition 5.3.1. We define $\mathcal{D}_{H,p} : \mathcal{M}_+(X) \times \mathcal{M}_+(X) \to [0, +\infty]$ and $W_{H,p} : \mathcal{P}(\mathfrak{C}[X]) \times \mathcal{P}(\mathfrak{C}[X]) \to [0, +\infty]$ as

$$\mathcal{D}_{H,p}(\mu_1, \mu_2) := \mathscr{U}_H^{1/p}(\mu_1, \mu_2) = \left(\inf \left\{ \int_{\mathfrak{C}[X,X]} H \, d\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in \mathfrak{H}^1(\mu_1, \mu_2) \right\} \right)^{1/p},$$

$$W_{H,p}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) := \left(\inf \left\{ \int_{\mathfrak{C}[X,X]} H \, d\boldsymbol{\gamma} \mid \boldsymbol{\gamma} \in \Gamma(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) \right\} \right)^{1/p},$$

where $\Gamma(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$ is the set of transport plans from $\boldsymbol{\alpha}_1$ to $\boldsymbol{\alpha}_2$ defined as

$$\Gamma(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) := \left\{ \boldsymbol{\gamma} \in \mathfrak{C}[X, X] \mid \pi_\sharp^{\mathfrak{C}_1} \boldsymbol{\gamma} = \boldsymbol{\alpha}_1, \, \pi_\sharp^{\mathfrak{C}_2} \boldsymbol{\gamma} = \boldsymbol{\alpha}_2 \right\}.$$

Finally we set

$$\mathcal{P}_{H,p}(\mathfrak{C}[X]) := \left\{ \boldsymbol{\alpha} \in \mathcal{P}(\mathfrak{C}[X] \mid \int_{\mathfrak{C}[X]} H(\mathfrak{y}, \mathfrak{o}) \, d\boldsymbol{\alpha}(\mathfrak{y}) < +\infty \right\},$$

$$\mathcal{M}_{H,p}(X) := \left\{ \mu \in \mathcal{M}_+(X) \mid \int_X H([x, 1], \mathfrak{o}) \, d\mu(x) < +\infty \right\}.$$

Remark 5.3.2. If $\mu \in \mathcal{M}_{H,p}(x)$ then every $\boldsymbol{\alpha} \in \mathcal{P}(\mathfrak{C}[X])$ such that $\mathfrak{h}^1(\boldsymbol{\alpha}) = \mu$ is an element of $\mathcal{P}_{H,p}(\mathfrak{C}[X])$.

Remark 5.3.3. Let us consider the map $T^p : \mathfrak{C}[X, X] \to \mathfrak{C}[X, X]$ defined as

$$T^p([x_1, r_1], [x_2, r_2]) := ([x_1, r_1^{1/p}], [x_2, r_2^{1/p}]), \quad [x_1, r_1], [x_2, r_2] \in \mathfrak{C}[X, X].$$

It can be easily checked that, for every $\mu_1, \mu_2 \in \mathcal{M}_+(X)$, $T_\sharp^p : \mathfrak{H}^1(\mu_1, \mu_2) \to \mathfrak{H}^p(\mu_1, \mu_2)$ is a bijection so that

$$\mathscr{U}_{H \circ T^p}(\mu_1, \mu_2) = \inf \left\{ \int_{\mathfrak{C}[X,X]} H \, d\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in \mathfrak{H}^p(\mu_1, \mu_2) \right\}.$$

Thus, if $H : \mathfrak{C}[X, X] \to [0, +\infty]$ is a lower semicontinuous p-homogenous function which is the p-th power of an extended distance on $\mathfrak{C}[X]$ whose induced topology is stronger than the topology of $\mathfrak{C}[X]$, then $H \circ T^p : \mathfrak{C}[X, X] \to [0, +\infty]$ is a lower semicontinuous 1-homogenous function which is the p-th power of an extended distance on $\mathfrak{C}[X]$ whose induced topology is stronger than the topology of $\mathfrak{C}[X]$. This allows us to treat only the 1-homogenous case also in this metric setting without loss of generality. For example the case $H = d_{\mathfrak{C}}^2$ fits into this setting (and induces the Helliger Kantorovich distance on nonnegative measures).

Of course $(\mathcal{P}(\mathfrak{C}[X], W_{H,p})$ is an extended metric space while $(\mathcal{P}_{H,p}(\mathfrak{C}[X]), W_{H,p})$ is a metric space.
The following two Lemmas are the analogue of [76, Corollary 7.7, Corollary 7.13] and the proof is identical and thus omitted.

Lemma 5.3.4. *For every $\mu_1, \mu_2 \in \mathcal{M}_+(X)$, it holds*

$$\mathcal{D}_{H,p}(\mu_1, \mu_2) = \min\left\{W_{H,p}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) \mid \boldsymbol{\alpha}_i \in \mathcal{P}(\mathfrak{C}[X]),\, \mathfrak{h}^1(\boldsymbol{\alpha}_i) = \mu_i,\ i = 1, 2\right\}.$$

Remark 5.3.5. In particular we have $\mathcal{D}_{H,p}(\mu_1, \mu_2) = W_{H,p}(\pi_\sharp^{\mathfrak{C}_1}\boldsymbol{\alpha}, \pi_\sharp^{\mathfrak{C}_2}\boldsymbol{\alpha})$ for every $\boldsymbol{\alpha} \in \mathfrak{H}^1_H(\mu_1, \mu_2)$.

Lemma 5.3.6. *If $(\mu_i)_{i=1}^N \subset \mathcal{M}_+(X)$ with $N \geqslant 2$, then there exist $(\boldsymbol{\alpha}_i)_{i=1}^N \in \mathcal{P}(\mathfrak{C}[X])$ such that*

$$\mathfrak{h}^1(\boldsymbol{\alpha}_i) = \mu_i, \quad \mathcal{D}_{H,p}(\mu_{i-1}, \mu_i) = W_{H,p}(\boldsymbol{\alpha}_{i-1}, \boldsymbol{\alpha}_i) \quad \textit{for every } i \in \{2, \ldots, N\}.$$

The next two theorems are slight generalizations of [76, Corollary 7.14, Theorem 7.15] and, although the proofs are very similar, there are some small modifications to be taken into account so that we report them.

Theorem 5.3.7. *The pair $(\mathcal{M}_+(X), \mathcal{D}_{H,p})$ is an extended metric space. If H finite, then $(\mathcal{M}_{H,p}(X), \mathcal{D}_{H,p})$ is a metric space.*

Proof. The map $T : \mathfrak{C}[X, X] \to \mathfrak{C}[X, X]$ defined as

$$T(\mathfrak{y}_1, \mathfrak{y}_2) = (\mathfrak{y}_2, \mathfrak{y}_1), \quad (\mathfrak{y}_1, \mathfrak{y}_2) \in \mathfrak{C}[X, X]$$

is such that, for every $\mu_1, \mu_2 \in \mathcal{M}_+(X)$, the map $T_\sharp : \mathfrak{H}^1(\mu_1, \mu_2) \to \mathfrak{H}^1(\mu_2, \mu_1)$ is a bijection satisfying

$$\int_{\mathfrak{C}[X,X]} H \, d\boldsymbol{\alpha} = \int_{\mathfrak{C}[X,X]} H \, dT_\sharp \boldsymbol{\alpha}$$

by the symmetry of H. This gives that $\mathcal{D}_{H,p}(\mu_1, \mu_2) = \mathcal{D}_{H,p}(\mu_2, \mu_1)$ for every $\mu_1, \mu_2 \in \mathcal{M}_+(X)$.
If $\mu \in \mathcal{M}_+(X)$ and we define

$$\boldsymbol{\alpha} = ((\mathrm{id}_{\mathfrak{C}[X]}, \mathrm{id}_{\mathfrak{C}[X]}) \circ \mathfrak{p})_\sharp(\mu \otimes \delta_1) \in \mathfrak{H}^1(\mu, \mu),$$

we obtain that

$$\mathcal{D}^p_{H,p}(\mu, \mu) \leqslant \int_{\mathfrak{C}[X,X]} H \, d\boldsymbol{\alpha} = \int_{\mathfrak{C}[X,X]} H(\mathfrak{y}, \mathfrak{y}) \, d(\mathfrak{p}_\sharp(\mu \otimes \delta_1))(\mathfrak{y}) = 0.$$

If, on the other hand, $\mu_1, \mu_2 \in \mathcal{M}_+(X)$ are s.t. $\mathcal{D}_{H,p}(\mu_1, \mu_2) = 0$ and $\boldsymbol{\alpha} \in \mathfrak{H}^1_H(\mu_1, \mu_2)$, we get that $\boldsymbol{\alpha}$ is concentrated on $\{(\mathfrak{y}, \mathfrak{y}) \mid \mathfrak{y} \in \mathfrak{C}[X]\}$, so that $\mu_1 = \mathfrak{h}^1_1(\boldsymbol{\alpha}) = \mathfrak{h}^1_2(\boldsymbol{\alpha}) = \mu_2$. This proves that $\mathcal{D}_{H,p}(\mu_1, \mu_2) = 0$ if and only if $\mu_1 = \mu_2$.
Finally if $\mu_1, \mu_2, \mu_3 \in \mathcal{M}_+(X)$, we can find, thanks to Lemma 5.3.6, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3 \in \mathcal{P}(\mathfrak{C}[X])$ such that

$$\mathfrak{h}^1(\boldsymbol{\alpha}_i) = \mu_i, \quad \mathcal{D}_{H,p}(\mu_{i-1}, \mu_i) = W_{H,p}(\boldsymbol{\alpha}_{i-1}, \boldsymbol{\alpha}_i) \quad i = 2, 3.$$

Then, using Lemma 5.3.4, we have

$$\begin{aligned}\mathcal{D}_{H,p}(\mu_1,\mu_3) &\leqslant W_{H,p}(\alpha_1,\alpha_3)\\ &\leqslant W_{H,p}(\alpha_1,\alpha_2)+W_{H,p}(\alpha_2,\alpha_3)\\ &= \mathcal{D}_{H,p}(\mu_1,\mu_2)+\mathcal{D}_{H,p}(\mu_2,\mu_3).\end{aligned}$$

This proves that $\mathcal{D}_{H,p}$ satisfies the triangular inequality and concludes the proof that $(\mathcal{M}_+(X),\mathcal{D}_{H,p})$ is an extended metric space.
If H is finite, $\mu \in \mathcal{M}_{H,p}(X)$ and $\alpha \in \mathcal{P}(\mathfrak{C}[X])$ is s.t. $\mathfrak{h}^1(\alpha)=\mu$, then

$$\int_{\mathfrak{C}[X]} H(\mathfrak{y},\mathfrak{o})\,d\alpha(\mathfrak{y}) = \int_{\mathfrak{C}[X]} rH([x,1],\mathfrak{o})\,d\alpha([x,r]) = \int_X H([x,1],\mathfrak{o})\,d\mu(x) < +\infty,$$

so that $\alpha \in \mathcal{P}_{H,p}(\mathfrak{C}[X])$. Then, again from Lemma 5.3.4, if $\mu_1,\mu_2 \in \mathcal{M}_{H,p}(X)$ we can find $\alpha_1,\alpha_2 \in \mathcal{P}_{H,p}(\mathfrak{C}[X])$ s.t. $\mathfrak{h}^1(\alpha_i)=\mu_i$ for $i=1,2$ so that

$$\mathcal{D}_{H,p}(\mu_1,\mu_2) \leqslant W_{H,p}(\alpha_1,\alpha_2) < +\infty.$$

□

Theorem 5.3.8. *If $(\mu_n)_n \subset \mathcal{M}_{H,p}(X)$ and $\mu \in \mathcal{M}_{H,p}(X)$, then*

$$\lim_{n\to+\infty} \mathcal{D}_{H,p}(\mu_n,\mu)=0 \quad\Longleftrightarrow\quad \begin{cases}\mu_n \rightharpoonup \mu, \\ \int_X H([x,1];\mathfrak{o})\,d\mu(x) \to \int_X H([x,1];\mathfrak{o})\,d\mu(x).\end{cases}.$$

In particular, if X is separable, also $(\mathcal{M}_{H,p}(X),\mathcal{D}_{H,p})$ is separable.

Proof. We claim that

$$\text{there exists } a>0 \text{ such that } H([x,1],\mathfrak{o}) \geqslant a \quad \text{for every } x\in X. \tag{5.3.1}$$

If not, we could find a sequence $(x_n)_n \subset X$ such that

$$H([x_n,1],\mathfrak{o}) \to 0$$

which implies, since the topology induced by H is stronger than the topology of $\mathfrak{C}[X]$, that $[x_n,1] \to \mathfrak{o}$ in the topology of $\mathfrak{C}[X]$, which is a contradiction. This proves (5.3.1).
We first prove the $\Rightarrow$ implication. Notice that

$$\mathcal{D}_{H,p}(\nu,0_X) = \int_X H([x,1],\mathfrak{o})\,d\nu(x) \quad \text{for every } \nu \in \mathcal{M}_{H,p}(X) \tag{5.3.2}$$

so that, by triangle inequality, we get

$$\int_X H([x,1],\mathfrak{o})\,d\mu_n(x) = \mathcal{D}_{H,p}(\mu_n,0_X) \to \mathcal{D}_{H,p}(\mu,0_X) = \int_X H([x,1],\mathfrak{o})\,d\mu(x).$$

We show that $\mu_n \rightharpoonup \mu$ by contradiction: assume that there exist $\xi \in C_b(X)$ and a (unrelabeled) subsequence s.t.

$$\inf_n \left|\int_X \xi\,d\mu_n - \int_X \xi\,d\mu\right| > 0. \tag{5.3.3}$$

Observe that

$$\mu_n(X) \leqslant \frac{1}{a}\int_X H([x,1],\mathfrak{o})\,d\mu_n(x) \to \frac{1}{a}\int_X H([x,1],\mathfrak{o})\,d\mu(x) < +\infty,$$

so that $R := (\sup_n \mu_n(X) + \mu(X)) < +\infty$. By Proposition 4.1.3 and Lemma 3.2.10, we can find $(\boldsymbol{\alpha}_n)_n \subset \mathcal{P}(\mathfrak{C}_R[X,X])$ such that $\boldsymbol{\alpha}_n \in \mathfrak{H}^1_H(\mu_n,\mu)$ for every $n \in \mathbb{N}$. Let us define $\boldsymbol{\alpha}^1_n := \pi^{\mathfrak{C}_1}_\sharp \boldsymbol{\alpha}_n$, $\boldsymbol{\alpha}^2_n := \pi^{\mathfrak{C}_2}_\sharp \boldsymbol{\alpha}_n$, $n \in \mathbb{N}$. Since $\mathfrak{h}^1(\boldsymbol{\alpha}^2_n) = \mu$ for every $n \in \mathbb{N}$, we obtain (see the proof of Lemma 3.2.6) the existence of a subsequence $k \mapsto n(k)$ and $\boldsymbol{\alpha}_2 \in \mathcal{P}(\mathfrak{C}_R[X])$ with $\mathfrak{h}^1(\boldsymbol{\alpha}_2) = \mu$ such that $\boldsymbol{\alpha}^2_{n(k)} \rightharpoonup \boldsymbol{\alpha}_2$. Moreover

$$\int_{\mathfrak{C}[X]} H(\mathfrak{y},\mathfrak{o})\,d\boldsymbol{\alpha}^2_n = \int_X H([x,1],\mathfrak{o})\,d\mu(x) \quad \text{for every } n \in \mathbb{N},$$

giving that (see e.g. [5, Proposition 7.1.5]) $W_{H,p}(\boldsymbol{\alpha}^2_{n(k)},\boldsymbol{\alpha}_2) \to 0$. Then

$$\begin{aligned} W_{H,p}(\boldsymbol{\alpha}^1_{n(k)},\boldsymbol{\alpha}_2) &\leqslant W_{H,p}(\boldsymbol{\alpha}^1_{n(k)},\boldsymbol{\alpha}^2_{n(k)}) + W_{H,p}(\boldsymbol{\alpha}^2_{n(k)},\boldsymbol{\alpha}_2) \\ &= \mathcal{D}_{H,p}(\mu_{n(k)},\mu) + W_{H,p}(\boldsymbol{\alpha}^2_n,\boldsymbol{\alpha}_2) \to 0, \end{aligned}$$

where we used Remark 5.3.5. Thus $W_{H,p}(\boldsymbol{\alpha}^1_{n(k)},\boldsymbol{\alpha}_2) \to 0$ and, in particular, $\boldsymbol{\alpha}^1_{n(k)} \rightharpoonup \boldsymbol{\alpha}_2$ so that

$$\int_X \xi\,d\mu_{n(k)} = \int_{\mathfrak{C}[X]} \xi(x) r\,d\boldsymbol{\alpha}^1_{n(k)}([x,r]) \to \int_{\mathfrak{C}[X]} \xi(x) r\,d\boldsymbol{\alpha}_2([x,r]) = \int_X \xi\,d\mu, \quad (5.3.4)$$

where we used that the map

$$[x,r] \mapsto r\xi(x)$$

belongs to $C_b(\mathfrak{C}_R[X])$ and $\boldsymbol{\alpha}^1_{n(k)}$ is concentrated on $\mathfrak{C}_R[X]$ for every $k \in \mathbb{N}$. Since (5.3.4) is a contradiction with (5.3.3), this concludes the proof of the $\Rightarrow$ implication.

Let us prove the $\Leftarrow$ implication. If $\mu = 0_X$, we have already by (5.3.2) that $\mathcal{D}_{H,p}(\mu_n,\mu) \to 0$. Let us then assume that $m := \mu(X) > 0$. Up to passing to a (unrelabeled) subsequence, we can assume that $m_n := \mu_n(X) \geqslant m/2 > 0$ for every $n \in \mathbb{N}$. Let us define $\boldsymbol{\alpha}_n, \boldsymbol{\alpha} \in \mathcal{P}(\mathfrak{C}[X])$ as

$$\boldsymbol{\alpha} := \mathfrak{p}_\sharp\left(m^{-1}\mu \otimes \delta_m\right), \quad \boldsymbol{\alpha}_n := \mathfrak{p}_\sharp\left(m_n^{-1}\mu_n \otimes \delta_{m_n}\right) \quad n \in \mathbb{N}.$$

It is easy to check that $\mathfrak{h}^1(\boldsymbol{\alpha}_n) = \mu_n$, $n \in \mathbb{N}$, $\mathfrak{h}^1(\boldsymbol{\alpha}) = \mu$ and $\boldsymbol{\alpha}_n \rightharpoonup \boldsymbol{\alpha}$. To conclude is then enough to show that $W_{H,p}(\boldsymbol{\alpha}_n,\boldsymbol{\alpha}) \to 0$ and then apply Lemma 5.3.4. Since

$$\begin{aligned} \int_{\mathfrak{C}[X]} H(\mathfrak{y},\mathfrak{o})\,d\boldsymbol{\alpha}_n([x,r]) &= \int_X H([x,1],\mathfrak{o})\,d\mu_n(x) \\ &\to \int_X H([x,1],\mathfrak{o})\,d\mu(x) \\ &= \int_{\mathfrak{C}[X]} H(\mathfrak{y},\mathfrak{o})\,d\boldsymbol{\alpha}([x,r]), \end{aligned}$$

we get that $W_{H,p}(\boldsymbol{\alpha}_n,\boldsymbol{\alpha}) \to 0$ applying [5, Proposition 7.1.5]. □

Part II

DISSIPATIVE EVOLUTIONS IN KANTOROVICH-WASSERSTEIN SPACES

We introduce and investigate a notion of multivalued dissipative operator (called Multivalued Probability Vector Field - MPVF) in the 2-Wasserstein space of Borel probability measures on a (possibly infinite dimensional) separable Hilbert space. Taking inspiration from the theories of dissipative operators in Hilbert spaces and of Wasserstein gradient flows, we study the well-posedness for evolutions driven by such MPVFs, and we characterize them by a suitable Evolution Variational Inequality (EVI), following the Bénilan notion of integral solutions to dissipative evolutions in Banach spaces. Our approach to prove the existence of such EVI-solutions is twofold: on one side, under an abstract stability condition, we build a measure-theoretic version of the Explicit Euler scheme showing novel convergence results with optimal error estimates; on the other hand, under a suitable discrete approximation assumption on the MPVF, we recast the EVI-solution as the evolving law of the solution trajectory of an appropriate dissipative evolution in an L^2 space of random variables.

6 PRELIMINARY MATERIAL

The aim of this chapter is to introduce some of the technical tools that are used in the rest of Part II. In particular, in Section 6.1 we synthetize the theory of dissipative evolutions in Hilbert spaces and we present a few results connected to maximal dissipative extensions; Section 6.2 deals with Borel partitions, the parametrization of measures through random variables and the approximation of Optimal Transport plans with maps; Section 6.3 presents a crucial notion of topology for measures in product spaces that will be applied to the case of $\mathcal{P}(\mathsf{T}\mathbb{H}) \cong \mathcal{P}(\mathbb{H} \times \mathbb{H})$; Section 6.4 contains a few results related to *triplans* (i.e. probabilities in $\mathcal{P}(\mathbb{H}^3)$) and the interpolation of measures they generate; finally Section 6.5 presents the local optimality of couplings between discrete measures.

This Chapter is the result of a collaboration with Giulia Cavagnari and Giuseppe Savaré and Section 6.3 appeared in [34].

6.1 DISSIPATIVE EVOLUTIONS IN HILBERT SPACES

Let H be a Hilbert space. Given a set $\mathbf{B} \subset \mathsf{H} \times \mathsf{H}$ we will set $\mathbf{B}[x] := \{v \in \mathsf{H} : (x, v) \in \mathbf{B}\}$ and $\mathrm{D}(\mathbf{B}) := \{x \in \mathsf{H} : \mathbf{B}[x] \neq \emptyset\}$. A set $\mathbf{B} \subset \mathsf{H} \times \mathsf{H}$ is dissipative if

$$\langle v - w, x - y\rangle \leqslant 0 \quad \text{for every } (x, v),\ (y, w) \in \mathbf{B}. \tag{6.1.1}$$

A dissipative set $\mathbf{B}$ is maximal if [26, Chap. II, Def. 2.2]

$$(x, v) \in \mathsf{H} \times \mathsf{H}, \quad \langle v - w, x - y\rangle \leqslant 0 \quad \text{for every } (y, w) \in \mathbf{B} \quad \Rightarrow \quad (x, v) \in \mathbf{B}. \tag{6.1.2}$$

$\mathbf{B}$ is maximal if and only if for every $x \in \mathsf{H}$ and every $\tau > 0$ there exists a unique $x_\tau \in \mathrm{D}(\mathbf{B})$ solving [26, Cap. II, Prop. 2.2]

$$x_\tau - x \in \tau \mathbf{B}[x_\tau]. \tag{6.1.3}$$

In particular the resolvent operator $\mathbf{J}_\tau := (I - \tau\mathbf{B})^{-1}$ is an everywhere defined contraction in H.

Given $E \subset \mathsf{H}$, we denote by $\operatorname{co}(E)$ the convex hull of E and by $\overline{\operatorname{co}}(E)$ its closure.

If $\mathbf{B}$ is a dissipative operator, then there exists a maximal extension $\tilde{\mathbf{B}}$ of $\mathbf{B}$ whose domain is included in $\overline{\operatorname{co}}(\mathrm{D}(\mathbf{B}))$ [26, Chap. II, Cor. 2.1].

If $\mathbf{B}$ is maximal then $\overline{\mathrm{D}(\mathbf{B})}$ is convex and, for every $x \in \mathrm{D}(\mathbf{B})$, $\mathbf{B}[x]$ is a closed convex subset of H, whose element of minimal norm is denoted by $\mathbf{B}^\circ(x)$. The map $\mathbf{B}^\circ : \mathrm{D}(\mathbf{B}) \to \mathsf{H}$ is also called minimal selection of $\mathbf{B}$ and satisfies the following property (see [26, Chap. II, Prop. 2.7])

$$(x, v) \in \overline{\mathrm{D}(\mathbf{B})} \times \mathsf{H}, \quad \langle v - \mathbf{B}^\circ(y), x - y\rangle \leqslant 0 \quad \text{for every } y \in \mathrm{D}(\mathbf{B}) \quad \Rightarrow \quad (x, v) \in \mathbf{B}.$$

(6.1.4)

Moreover, by [26, Chap. II,Prop. 2.6(iii)] we have

$$\mathbf{B}^\circ(x) = \lim_{\tau\downarrow 0} \frac{J_\tau(x) - x}{\tau}.$$

The following proposition is a slight generalization of [10, Lemma 2.3] but we report its proof for the reader's convenience.

Proposition 6.1.1. *Let* $\mathbf{B} \subset H \times H$ *be maximal and let* $\mathbf{G} \subset \mathbf{B}$ *be s.t.* $D(\mathbf{G})$ *is dense in* $D(\mathbf{B})$. *Then for every* $x \in \operatorname{int}(D(\mathbf{B}))$ *it holds*

$$\mathbf{B}[x] = \overline{\mathrm{co}}\left(\{v \in H \mid \exists (x_n, v_n) \in \mathbf{G} \text{ s.t. } x_n \to x,\ v_n \rightharpoonup v\}\right). \tag{6.1.5}$$

Proof. Let $x \in \operatorname{int}(D(\mathbf{B}))$ and let us define

$$\mathbf{M}[x] := \overline{\mathrm{co}}\left(\{v \in H \mid \exists (x_n, v_n) \in \mathbf{G} \text{ s.t. } x_n \to x,\ v_n \rightharpoonup v\}\right).$$

If $(x_n, v_n) \in \mathbf{G} \subset \mathbf{B}$ with $x_n \to x$ and $v_n \rightharpoonup v$, by dissipativity of $\mathbf{B}$, we have that

$$\langle v_n - w, x_n - y\rangle \leqslant 0 \quad \forall (y, w) \in \mathbf{B}.$$

Passing to the limit we get

$$\langle v - w, x - y\rangle \leqslant 0 \quad \forall (y, w) \in \mathbf{B},$$

meaning that $v \in \mathbf{B}[x]$. This, together with the closure and convexity of $\mathbf{B}[x]$, proves that $\mathbf{M}[x] \subset \mathbf{B}[x]$. Let us prove the other inclusion by contradiction: suppose that there is some $v \in \mathbf{B}[x]$ s.t. $v \notin \mathbf{M}[x]$. The sets $\{v\}$ and $\mathbf{M}[x]$ are disjoint, closed, convex and $\{v\}$ is also compact. By Hahn-Banach theorem we can find some $z \in H$ with $|z| = 1$ s.t.

$$\langle v, z\rangle > \langle u, z\rangle \quad \forall u \in \mathbf{M}[x]. \tag{6.1.6}$$

Since $x \in \operatorname{int}(D(\mathbf{B}))$, if we define $z_n := x + z/n$, we have that $z_n \in \operatorname{int}(D(\mathbf{B}))$ for n sufficiently large. We can thus find $x_n \in D(\mathbf{G})$ s.t. $|x_n - z_n| < n^{-2}$. Clearly $x_n \to x$ and it is easy to check that $(x_n - x)/|x_n - x| \to z$. Since $x_n \in D(\mathbf{G})$, we can find $v_n \in \mathbf{G}(x_n)$. Since $\mathbf{B}$ is maximal, it is locally bounded ([26, Prop. 2.9]) at x. Being $\mathbf{G} \subset \mathbf{B}$ and being $x_n \to x$, the sequence (v_n) is bounded so that, up to an unrelabeled subsequence, it converges weakly to some point $u \in H$. By dissipativity of $\mathbf{B}$ we have

$$\langle v - v_n, x - x_n\rangle \leqslant 0 \quad \forall n \in \mathbb{N},$$

so that, dividing by $|x_n - x|$ and passing to the limit, we obtain

$$\langle v - u, z\rangle \leqslant 0,$$

a contradiction with (6.1.6) since, obviously, $u \in \mathbf{M}[x]$. □

The following proposition is an immediate consequence of [98, Theorem 1].

Proposition 6.1.2. *Let* $\mathbf{B} \subset H \times H$ *be dissipative with open non empty convex domain. Then there exists a unique maximal* $\widetilde{\mathbf{B}} \supset \mathbf{B}$ *with* $D(\widetilde{\mathbf{B}}) \subset \overline{D(\mathbf{B})}$ *and it is characterized by*

$$\widetilde{\mathbf{B}} = \left\{ (x, v) \in \overline{D(\mathbf{B})} \times H \mid \langle v - w, x - y \rangle \leqslant 0 \quad \forall (y, w) \in \mathbf{B} \right\}.$$

As a consequence of Propositions 6.1.1 and 6.1.2 we can prove the following.

Theorem 6.1.3. *Let* $\mathbf{B} \subset H \times H$ *be dissipative with*

$$C := \overline{D(\mathbf{B})} \textit{ is convex}, \quad \operatorname{int}(D(\mathbf{B})) \neq \emptyset.$$

Then there exists a unique maximal $\widetilde{\mathbf{B}} \supset \mathbf{B}$ *with* $D(\widetilde{\mathbf{B}}) \subset C$ *and it is characterized by*

$$\widetilde{\mathbf{B}} = \{(x, v) \in C \times H \mid \langle v - w, x - y \rangle \leqslant 0 \quad \forall (y, w) \in \mathbf{B}\}. \tag{6.1.7}$$

Moreover, for every $x \in \operatorname{int}\left(D(\widetilde{\mathbf{B}})\right)$ *it holds*

$$\widetilde{\mathbf{B}}[x] = \overline{\mathrm{co}}\left(\{v \in H \mid \exists (x_n, v_n) \in \mathbf{B} \textit{ s.t. } x_n \to x,\, v_n \rightharpoonup v\}\right). \tag{6.1.8}$$

Finally

$$\operatorname{int}(C) = \operatorname{int}\left(D(\widetilde{\mathbf{B}})\right) \subset D(\widetilde{\mathbf{B}}) \subset \overline{D(\widetilde{\mathbf{B}})} = C. \tag{6.1.9}$$

Proof. Let $\mathbf{B}'$ be a maximal extension of $\mathbf{B}$ with $D(\mathbf{B}') \subset C$; by dissipativity of $\mathbf{B}'$ and since $\mathbf{B} \subset \mathbf{B}'$, then $\mathbf{B}' \subset \widetilde{\mathbf{B}}$, where $\widetilde{\mathbf{B}}$ is defined as in (6.1.7). We need to prove the other inclusion.
Since $D(\mathbf{B}) \subset D(\mathbf{B}') \subset C$, we have that $\overline{D(\mathbf{B}')} = C$. Moreover, being $\mathbf{B}'$ maximal and being the interior of its domain nonempty, we have (see [26, Proposition 2.9]) that

$$\operatorname{int}\left(D(\mathbf{B}')\right) \text{ is convex}, \quad \operatorname{int}\left(D(\mathbf{B}')\right) = \operatorname{int}\left(\overline{D(\mathbf{B}')}\right) = \operatorname{int}(C).$$

It is then clear that $\mathbf{B}_0 := \mathbf{B}' \cap (\operatorname{int}\left(D(\mathbf{B}')\right) \times H)$ is dissipative with open and nonempty convex domain so that, by Proposition 6.1.2, there exists a unique maximal $\mathbf{B}'' \supset \mathbf{B}_0$ with $D(\mathbf{B}'') \subset \overline{D(\mathbf{B}_0)} = \overline{\operatorname{int}\left(D(\mathbf{B}')\right)} = \overline{\operatorname{int}(C)} = C$ (C is convex) and it is characterized by

$$\mathbf{B}'' = \{(x, v) \in C \times H \mid \langle v - w, x - y \rangle \leqslant 0 \quad \forall (y, w) \in \mathbf{B}_0\}. \tag{6.1.10}$$

Since $\mathbf{B}' \supset \mathbf{B}_0$, $\mathbf{B}'$ is maximal and $D(\mathbf{B}') \subset C$, it must be that $\mathbf{B}' = \mathbf{B}''$.
By (6.1.10), we need to prove that

$$\widetilde{\mathbf{B}} \subset \{(x, v) \in C \times H \mid \langle v - w, x - y \rangle \leqslant 0 \quad \forall (y, w) \in \mathbf{B}_0\}. \tag{6.1.11}$$

To this aim we apply Proposition 6.1.1 to the maximal $\mathbf{B}'$ and its subset $\mathbf{B}$ noticing that $D(\mathbf{B})$ is dense in $D(\mathbf{B}')$. In this way, we obtain that

$$\mathbf{B}_0[y] = \overline{\mathrm{co}}\left(\overline{\mathbf{B}}[y]\right), \quad y \in D(\mathbf{B}_0), \tag{6.1.12}$$

where

$$\overline{\mathbf{B}}[y] = \{u \in H \mid \exists (y_n, u_n) \in \mathbf{B} \text{ s.t. } y_n \to y,\, u_n \rightharpoonup u\}.$$

If $(x, v) \in \widetilde{\mathbf{B}}$ and $(y, w) \in D(\mathbf{B}_0) \times H$ is such that $w \in \overline{\mathbf{B}}[y]$, we can find a sequence $(y_n, u_n) \in \mathbf{B}$ s.t. $y_n \to y$ and $u_n \rightharpoonup w$; then, by the very definition of $\widetilde{\mathbf{B}}$, we have

$$\langle v - u_n, x - y_n \rangle \leqslant 0 \quad \forall n \in \mathbb{N},$$

so that, passing to the limit, we get

$$\langle v - w, x - y \rangle \leqslant 0.$$

This proves that, if $(x, v) \in \widetilde{\mathbf{B}}$, then

$$\langle v - w, x - y \rangle \leqslant 0 \quad \forall w \in \overline{\mathbf{B}}[y], \quad \forall y \in D(\mathbf{B}_0). \tag{6.1.13}$$

Finally, if $(x, v) \in \widetilde{\mathbf{B}}$ and $(y, w) \in \mathbf{B}_0$, we can find a sequence $(N_n)_n \subset \mathbb{N}$, numbers $(\alpha_i^n)_{i=1}^{N_n} \subset [0, 1]$ and points $(w_i^n)_{i=1}^{N_n} \subset \overline{\mathbf{B}}[y]$ s.t.

$$\sum_{i=1}^{N_n} \alpha_i^n = 1 \ \forall n \in \mathbb{N}, \qquad \lim_{n\to+\infty} \sum_{i=1}^{N_n} \alpha_i^n w_i^n = w.$$

By (6.1.13)

$$\langle v - w_i^n, x - y \rangle \leqslant 0 \quad \forall i = 1, \ldots, N_n, \quad \forall n \in \mathbb{N},$$

so that, multiplying by α_i^n and summing up w.r.t. i, we obtain

$$\langle v - \sum_{i=1}^{N_n} \alpha_i^n w_i^n, x - y \rangle \leqslant 0 \quad \forall n \in \mathbb{N}.$$

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

$$\langle v - w, x - y \rangle \leqslant 0,$$

so that (6.1.11) holds. Finally notice that (6.1.8) is already stated in (6.1.12). □

As a consequence of (6.1.10), since we have proven that $\mathbf{B}'' = \widetilde{\mathbf{B}}$, we have the following corollary.

Corollary 6.1.4. *Let* $\mathbf{B} \subset H \times H$ *be as in Theorem 6.1.3 and let* $\mathbf{G} : \operatorname{int}(C) \to H$ *be a single valued selection of the maximal extension* $\widetilde{\mathbf{B}}$ *of* $\mathbf{B}$*. Then the unique maximal extension* $\widetilde{\mathbf{G}}$ *of* $\mathbf{G}$ *coincides with* $\widetilde{\mathbf{B}}$ *and in particular*

$$(x, v) \in \widetilde{\mathbf{B}} \Leftrightarrow x \in C, \quad \langle v - \mathbf{G}[y], x - y \rangle \leqslant 0 \quad \forall y \in \operatorname{int}(C). \tag{6.1.14}$$

Let us consider a different situation when $D(\mathbf{B})$ does not contain interior points but $\mathbf{B}$ satisfies

$$\begin{gathered}\text{there exists } D \supset D(\mathbf{B}) \text{ s.t. for every } x \in D,\ \tau > 0 \\ \text{there exists a unique } x_\tau =: \mathbf{J}_\tau x \in D(\mathbf{B}) : \quad x_\tau - x \in \tau \mathbf{B}[x_\tau].\end{gathered} \tag{6.1.15}$$

Lemma 6.1.5. *If* $\mathbf{B} \subset H \times H$ *is dissipative, satistifies* (6.1.15) *and* $C := \overline{D(\mathbf{B})}$ *is convex, then* $\mathbf{B}$ *admits a unique maximal extension* $\widetilde{\mathbf{B}}$ *with* $D(\widetilde{\mathbf{B}}) \subset C$ *characterized by*

$$\widetilde{\mathbf{B}} = \big\{(x, v) \in C \times H \mid \langle v - \tau^{-1}(\mathbf{J}_\tau y - y), x - \mathbf{J}_\tau y\rangle \leqslant 0 \quad \forall y \in D(\mathbf{B}),\ \tau > 0\big\}. \tag{6.1.16}$$

If moreover D *is dense in* H *we have*

$$\widetilde{\mathbf{B}} = \overline{\mathbf{B}} := \Big\{(x, v) \in H \times H : \exists\, (x_n, v_n) \in \mathbf{B} : x_n \to x, v_n \to v \quad \textit{as } n \to \infty\Big\}. \tag{6.1.17}$$

Proof. Let $\mathbf{B}'$ be any maximal extension of $\mathbf{B}$ with domain included in C and let $\mathbf{J}'_\tau$ be the resolvent associated with $\mathbf{B}'$. By dissipativity of $\mathbf{B}'$ and since $\mathbf{B} \subset \mathbf{B}'$, $\mathbf{B}' \subset \widetilde{\mathbf{B}}$ defined as in (6.1.16). We need to prove the other inclusion.
Clearly, the restriction of $\mathbf{J}'_\tau$ to $D \supset D(\mathbf{B})$ coincides with $\mathbf{J}_\tau$; since $\mathbf{J}'_\tau$ is a contraction, it is the only 1-Lipschitz extension of $\mathbf{J}_\tau$ to $\overline{D} \supset C$.

If $(x, v) \in \widetilde{\mathbf{B}}$, (6.1.16) yields by density that

$$\langle v - \tau^{-1}(\mathbf{J}'_\tau y - y), x - \mathbf{J}'_\tau y\rangle \leqslant 0 \quad \forall y \in D(\mathbf{B}'), \quad \forall \tau > 0, \tag{6.1.18}$$

and passing to the limit as $\tau \downarrow 0$ we obtain that

$$\langle v - \mathbf{B}'^{\circ}(y), x - y\rangle \leqslant 0 \quad \forall y \in D(\mathbf{B}'). \tag{6.1.19}$$

We can then apply (6.1.4) and conclude that $(x, v) \in \mathbf{B}'$.

Let us now prove (6.1.17) in the case D is dense in H: since $\overline{\mathbf{B}} \subset \widetilde{\mathbf{B}}$, it is sufficient to prove the opposite inclusion $\widetilde{\mathbf{B}} \subset \overline{\mathbf{B}}$. Let $(x, v) \in \widetilde{\mathbf{B}}$ and set $y := x - v$. Clearly $\mathbf{J}'_1 y = x$; since D is dense in H, there exists a sequence $(y_n)_n \subset D$ converging to y as $n \to \infty$. Setting $x_n := \mathbf{J}'_1 y_n$ and $v_n := x_n - y_n \in \mathbf{B}(x_n)$ we clearly have $\lim_{n\to\infty} x_n = x$, $\lim_{n\to\infty} v_n = v$. □

6.2 BOREL PARTITIONS

In this section we list some useful results concerning Borel isomorphisms and partitions of standard Borel spaces.

Definition 6.2.1. A *standard Borel space* $(\Omega, \mathcal{B})$ is a measurable space that is isomorphic (as a measure space) to a Polish space. Equivalently, there exists a Polish topology τ on Ω such that the Borel sigma algebra generated by τ coincides with $\mathcal{B}$. We say that a positive finite measure $\mathfrak{m}$ on $(\Omega, \mathcal{B})$ is diffuse if $\mathfrak{m}(\{\omega\}) = 0$ for every $\omega \in \Omega$ (notice that $\{\omega\} \in \mathcal{B}$ since it is compact in any Polish topology on Ω). In this case, we call the triplet $(\Omega, \mathcal{B}, \mathfrak{m})$ a *standard Borel measure space* (resp. standard Borel probability space, if $\mathfrak{m}$ is a probability).

We start with the following fundamental result that follows by e.g. [104, Theorem 9, Chapter 15].

Theorem 6.2.2. *Let* $(\Omega, \mathcal{B}, \mathfrak{m})$ *and* $(\Omega', \mathcal{B}', \mathfrak{m}')$ *be standard Borel measure spaces such that* $\mathfrak{m}(\Omega) = \mathfrak{m}'(\Omega')$. *Then there exist two measurable functions* $\varphi : \Omega \to \Omega'$ *and* $\psi : \Omega' \to \Omega$ *such that*

$$\psi \circ \varphi = \mathfrak{i}_\Omega \ \mathfrak{m}\text{-a.e. in } \Omega, \quad \varphi \circ \psi = \mathfrak{i}_{\Omega'} \ \mathfrak{m}'\text{-a.e. in } \Omega', \quad \varphi_\sharp \mathfrak{m} = \mathfrak{m}', \quad \psi_\sharp \mathfrak{m}' = \mathfrak{m}. \tag{6.2.1}$$

Corollary 6.2.3. *Let* $(\Omega, \mathcal{B}, \mathfrak{m})$ *be a standard Borel measure space and* $(\Omega', \mathcal{B}')$ *be a standard Borel space. Then for every positive measure* μ *on* $(\Omega', \mathcal{B}')$ *such that* $\mu(\Omega') = \mathfrak{m}(\Omega)$*, there exists a measurable map* $X : \Omega \to \Omega'$ *such that* $X_\sharp \mathfrak{m} = \mu$.

Definition 6.2.4. If $(\Omega, \mathcal{B})$ is a standard Borel space and $N \in \mathbb{N}$, a family of subsets $\mathfrak{P}_N = \{\Omega_{N,k}\}_{k \in I_N} \subset \mathcal{B}$, where $I_N := \{0, \dots, N-1\}$, is called a N-*partition of* $(\Omega, \mathcal{B})$ if

$$\bigcup_{k \in I_N} \Omega_{N,k} = \Omega, \quad \Omega_{N,k} \cap \Omega_{N,h} = \emptyset \text{ if } h, k \in I_N,\ h \neq k.$$

If $(\Omega, \mathcal{B}, \mathfrak{m})$ and $(\Omega', \mathcal{B}', \mathfrak{m}')$ are standard Borel measure spaces such that $\mathfrak{m}(\Omega) = \mathfrak{m}'(\Omega')$ and $\mathfrak{P}_N = \{\Omega_{N,k}\}_{k \in I_N}$ and $\mathfrak{P}'_N = \{\Omega'_{N,k}\}_{k \in I_N}$ are N-partitions of $(\Omega, \mathcal{B})$ and $(\Omega', \mathcal{B}')$ respectively, we say that $\mathfrak{P}_N$ and $\mathfrak{P}'_N$ are $\mathfrak{m} - \mathfrak{m}'$ *compatible* if

$$\mathfrak{m}(\Omega_{N,k}) = \mathfrak{m}'(\Omega'_{N,k}) \quad \forall k \in I_N.$$

Lemma 6.2.5. *Let* $(\Omega, \mathcal{B}, \mathfrak{m})$ *and* $(\Omega', \mathcal{B}', \mathfrak{m}')$ *be standard Borel measure spaces such that* $\mathfrak{m}(\Omega) = \mathfrak{m}'(\Omega')$ *and let* $\mathfrak{P}_N = \{\Omega_{N,k}\}_{k \in I_N}$ *and* $\mathfrak{P}'_N = \{\Omega'_{N,k}\}_{k \in I_N}$ *be two* $\mathfrak{m} - \mathfrak{m}'$ *compatible* N*-partitions of* $(\Omega, \mathcal{B})$ *and* $(\Omega', \mathcal{B}')$ *respectively, for some* $N \in \mathbb{N}$*. Then there exist two functions* $\varphi : \Omega \to \Omega'$ *and* $\psi : \Omega' \to \Omega$ *such that*

1. φ *is* $\mathcal{B}$-$\mathcal{B}'$ *measurable and* $\sigma(\mathfrak{P}_N)$-$\sigma(\mathfrak{P}'_N)$ *measurable;*
2. ψ *is* $\mathcal{B}'$-$\mathcal{B}$ *measurable and* $\sigma(\mathfrak{P}'_N)$-$\sigma(\mathfrak{P}_N)$ *measurable;*
3. *for every* $k \in I_N$ *it holds*
$$\varphi(\Omega_{N,k}) \subset \Omega'_{N,k}, \quad \psi(\Omega'_{N,k}) \subset \Omega_{N,k}; \tag{6.2.2}$$
4. *for every* $I \subset I_N$ *it holds*
$$\begin{aligned} &\psi_I \circ \varphi_I = \mathfrak{i}_{\Omega_I} \ \mathfrak{m}_I\text{-a.e. in } \Omega_I,\\ &\varphi_I \circ \psi_I = \mathfrak{i}_{\Omega'_I} \ \mathfrak{m}'_I\text{-a.e. in } \Omega'_I,\\ &(\varphi_I)_\sharp \mathfrak{m}_I = \mathfrak{m}'_I,\\ &(\psi_I)_\sharp \mathfrak{m}'_I = \mathfrak{m}_I, \end{aligned}$$
where the subscript I *denotes the restriction to* $\cup_{k \in I} \Omega_{N,k}$ *or* $\cup_{k \in I} \Omega'_{N,k}$.

Proof. Applying Theorem 6.2.2 to the standard Borel measures spaces $(\Omega_{\{k\}}, \mathcal{B}_{\{k\}}, \mathfrak{m}_{\{k\}})$ and $(\Omega'_{\{k\}}, \mathcal{B}'_{\{k\}}, \mathfrak{m}'_{\{k\}})$ for every $k \in I_N$, we obtain the existence of measurable functions φ_k, ψ_k satisfying (6.2.1) for each pair $\Omega_{N,k}$, $\Omega'_{N,k}$. It is then enough to define

$$\varphi(\omega) := \varphi_k(\omega) \quad \text{if } \omega \in \Omega_{N,k}, \quad \psi(\omega') := \psi_k(\omega') \quad \text{if } \omega' \in \Omega'_{N,k}.$$

Notice that (6.2.2) is satisfied by construction. □

If $(\Omega, \mathcal{B}, \mathfrak{m})$ is a standard Borel measure space, we denote by $S(\Omega, \mathcal{B}, \mathfrak{m})$ the class of $\mathcal{B}$-$\mathcal{B}$-measurable maps $g : \Omega \to \Omega$ which are essentially injective and measure preserving, meaning that there exists a full $\mathfrak{m}$-measure set $\Omega_0 \in \mathcal{B}$ such that g is injective on Ω_0 and $g_\sharp \mathfrak{m} = \mathfrak{m}$. If $\mathcal{A} \subset \mathcal{B}$ is a sigma algebra on Ω we denote by $S(\Omega, \mathcal{B}, \mathfrak{m}; \mathcal{A})$ the subset of $S(\Omega, \mathcal{B}, \mathfrak{m})$ of $\mathcal{A} - \mathcal{A}$ measurable maps. Finally $\mathrm{Sym}(I_N)$ denotes the set of permutations of I_N i.e. bijective maps $\sigma : I_N \to I_N$.

Corollary 6.2.6. *Let* $(\Omega, \mathcal{B}, \mathfrak{m})$ *be a standard Borel measure space and let* $\mathfrak{P}_N = \{\Omega_{N,k}\}_{k \in I_N}$ *be a* N*-partition of* $(\Omega, \mathcal{B})$ *for some* $N \in \mathbb{N}$ *such that* $\mathfrak{m}(\Omega_{N,k}) = \mathfrak{m}(\Omega)/N$ *for every* $k \in I_N$. *If* $\sigma \in \mathrm{Sym}(I_N)$, *there exists a measure preserving map* $g \in S(\Omega, \mathcal{B}, \mathfrak{m}; \sigma(\mathfrak{P}_N))$ *such that*

$$(g_k)_\sharp \mathfrak{m}|_{\Omega_{N,k}} = \mathfrak{m}|_{\Omega_{N,\sigma(k)}} \quad \forall k \in I_N,$$

where g_k *is the restriction of* g *to* $\Omega_{N,k}$.

Proof. It is enough to apply Lemma 6.2.5 to the standard Borel measure spaces $(\Omega, \mathcal{B}, \mathfrak{m})$ and $(\Omega', \mathcal{B}', \mathfrak{m}')$=$(\Omega, \mathcal{B}, \mathfrak{m})$ together with the N-partitions $\mathfrak{P}_N$ and $\mathfrak{P}'_N = \{\Omega_{N,\sigma(k)}\}_{k \in I_N}$ respectively. □

Corollary 6.2.7. *Let* $(\Omega, \mathcal{B}, \mathfrak{m})$ *be a standard Borel measure space and let* $\Omega_0, \Omega_1 \in \mathcal{B}$ *be such that* $\mathfrak{m}(\Omega_0) = \mathfrak{m}(\Omega_1) > 0$ *and* $\Omega_0 \cap \Omega_1 = \emptyset$. *Then there exists a measure preserving map* $g \in S(\Omega, \mathcal{B}, \mathfrak{m})$ *such that*

$$(g_0)_\sharp \mathfrak{m}|_{\Omega_0} = \mathfrak{m}|_{\Omega_1}, \quad (g_1)_\sharp \mathfrak{m}|_{\Omega_1} = \mathfrak{m}|_{\Omega_0}, \quad g(\omega) = \omega \text{ in } \Omega \setminus (\Omega_0 \cup \Omega_1),$$

where g_i *is the restriction of* g *to* Ω_k, $k = 0, 1$.

Proof. Applying Corollary 6.2.6 to $(\Omega_0 \cup \Omega_1, \mathcal{B}|_{\Omega_0 \cup \Omega_1}, \mathfrak{m}|_{\Omega_0 \cup \Omega_1})$ with the 2-Borel partition $\mathfrak{P}_2 = \{\Omega_k\}_{k=0,1}$ and σ sending 0 to 1, we obtain the existence of a measure preserving map $\tilde{g} \in S(\Omega_0 \cup \Omega_1, \mathcal{B}|_{\Omega_0 \cup \Omega_1}, \mathfrak{m}|_{\Omega_0 \cup \Omega_1})$ such that

$$(\tilde{g}_0)_\sharp \mathfrak{m}|_{\Omega_0} = \mathfrak{m}|_{\Omega_1}, \quad (\tilde{g}_1)_\sharp \mathfrak{m}|_{\Omega_1} = \mathfrak{m}|_{\Omega_0},$$

where $\tilde{g}_i$ is the restriction of $\tilde{g}$ to Ω_k, $k = 0, 1$. It is then enough to define $g : \Omega \to \Omega$ as

$$g(\omega) = \begin{cases} \tilde{g}(\omega) & \text{if } \omega \in \Omega_0 \cup \Omega_1, \\ \omega & \text{if } \omega \in \Omega \setminus (\Omega_0 \cup \Omega_1). \end{cases}$$

□

The next result follows by [112, Theorem 6.1.12], we refer also to [33, Appendix D] for a partial result. We recall that a filtration on $(\Omega, \mathcal{B})$ is a sequence $(\mathcal{F}_n)_{n \in \mathbb{N}}$ of sub-sigma algebras of $\mathcal{B}$ such that $\mathcal{F}_n \subset \mathcal{F}_{n+1}$.

Theorem 6.2.8. *Let* $(\Omega, \mathcal{B}, \mathfrak{m})$ *be a standard Borel measure space,* $(\mathcal{F}_n)_{n \in \mathbb{N}}$ *be a filtration on* $(\Omega, \mathcal{B})$ *such that* $\sigma(\{\mathcal{F}_n \mid n \in \mathbb{N}\}) = \mathcal{B}$ *and let* $\mathbb{H}$ *be a separable Hilbert space. Then, given* $X \in L^2((\Omega, \mathcal{B}, \mathfrak{m}); \mathbb{H})$, *the discrete time martingale*

$$X_n := \mathbb{E}_\mathfrak{m}[X \mid \mathcal{F}_n], \quad n \in \mathbb{N},$$

satisfies

$$\lim_{n \to +\infty} X_n = X \tag{6.2.3}$$

both $\mathfrak{m}$*-a.e. and in* $L^2((\Omega, \mathcal{B}, \mathfrak{m}); \mathbb{H})$.

We consider the partial order on $\mathbb{N}$ given by

$$m \prec n \quad \Leftrightarrow \quad m \mid n, \tag{6.2.4}$$

where $m \mid n$ means that $n/m \in \mathbb{N}$. We write $m \precneqq n$ if $m \prec n$ and $m \neq n$.

Definition 6.2.9. Let $(\Omega, \mathcal{B}, \mathfrak{m})$ be a standard Borel measure space and let $\mathfrak{N} \subset \mathbb{N}$ be an unbounded directed set w.r.t. $\prec$. We say that a collection of partitions $(\mathfrak{P}_N)_{N\in\mathfrak{N}}$ of Ω, with corresponding sigma algebras $\mathcal{B}_N := \sigma(\mathfrak{P}_N)$, is a $\mathfrak{N}$*-segmentation of* $(\Omega, \mathcal{B}, \mathfrak{m})$ if

1. $\mathfrak{P}_N = \{\Omega_{N,k}\}_{k\in I_N}$ is a N-partition of $(\Omega, \mathcal{B})$ for every $N \in \mathfrak{N}$,
2. $\mathfrak{m}(\Omega_{N,k}) = \mathfrak{m}(\Omega)/N$ for every $k \in I_N$ and every $N \in \mathfrak{N}$,
3. if $M \mid N = KM$ then $\bigcup_{k=0}^{K-1} \Omega_{N,mK+k} = \Omega_{M,m}$, $m \in I_M$,
4. $\sigma(\{\mathcal{B}_N \mid N \in \mathfrak{N}\}) = \mathcal{B}$.

In this case we call $(\Omega, \mathcal{B}, \mathfrak{m}, (\mathfrak{P}_N)_{N\in\mathfrak{N}})$ a $\mathfrak{N}$*-refined standard Borel measure space*.

Remark 6.2.10. It is clear that, if $M \mid N$, then $\mathcal{B}_M \subset \mathcal{B}_N$.

Example 6.2.11. The canonical example of $\mathfrak{N}$-refined standard Borel measure space is

$$([0,1), \mathcal{B}([0,1)), \lambda^c, (\mathfrak{I}_N)_{N\in\mathfrak{N}}),$$

where λ^c is the one dimensional Lebesgue measure restricted to $[0,1)$ and weighted by a constant $c > 0$ and $\mathfrak{I}_N = (I_{N,k})_{k\in I_N}$ with $I_{N,k} := [k/N, (k+1)/N)$, $k \in I_N$ and $N \in \mathfrak{N}$.

Lemma 6.2.12. *For any standard Borel measure space* $(\Omega, \mathcal{B}, \mathfrak{m})$ *and any unbounded directed set* $\mathfrak{N} \subset \mathbb{N}$ *w.r.t.* $\prec$*, there exists a* $\mathfrak{N}$*-segmentation of* $(\Omega, \mathcal{B}, \mathfrak{m})$.

Proof. Let $([0,1), \mathcal{B}([0,1)), \lambda^c, (\mathfrak{I}_N)_{N\in\mathfrak{N}})$ be the $\mathfrak{N}$-refined standard Borel measure space of Example 6.2.11 with $c = \mathfrak{m}(\Omega)$. Since $([0,1), \mathcal{B}([0,1)), \lambda^c)$ is a standard Borel measure space such that $\mathfrak{m}(\Omega) = \lambda^c([0,1))$, by Theorem 6.2.2 we can find measurable maps $\varphi : [0,1) \to \Omega$, $\psi : \Omega \to [0,1)$ and two subsets $\Omega_0 \in \mathcal{B}$, $U \in \mathcal{B}([0,1))$ such that $\mathfrak{m}(\Omega_0) = \lambda^c(U) = 0$, $\varphi \circ \psi = \mathbf{i}_{\Omega\setminus\Omega_0}$, $\psi \circ \varphi = \mathbf{i}_{[0,1)\setminus U}$, $\varphi_\sharp \lambda^c = \mathfrak{m}$ and $\psi_\sharp \mathfrak{m} = \lambda^c$. We can thus define

$$\Omega_{N,0} = \varphi(I_{N,0} \setminus U) \cup \Omega_0, \quad \Omega_{N,k} = \varphi(I_{N,k} \setminus U), \quad k \in I_N \setminus \{0\},\ N \in \mathfrak{N}.$$

Setting $\mathfrak{P}_N := \{\Omega_{N,k}\}_{k\in I_N}$ for every $N \in \mathfrak{N}$, it is easy to check that $(\mathfrak{P}_N)_{N\in\mathfrak{N}}$ is a $\mathfrak{N}$-segmentation of $(\Omega, \mathcal{B}, \mathfrak{m})$. □

In general, the collection of sigma-algebras given by $(\mathcal{B}_N)_{N\in\mathfrak{N}}$ is not a filtration since it fails to be ordered by inclusion. However, it is always possible to extract from $(\mathcal{B}_N)_{N\in\mathfrak{N}}$ a filtration still satisfying item (4) in Definition 6.2.9. More precisely we have the following result.

Lemma 6.2.13. *Let $\mathfrak{N} \subset \mathbb{N}$ be an unbounded directed subset w.r.t. $\prec$. Then there exists a sequence $(b_n)_n \subset \mathfrak{N}$ such that $b_n \precnsim b_{n+1}$ for every $n \in \mathbb{N}$ and for every $\mathfrak{N}$-refined standard Borel measure space $(\Omega, \mathcal{B}, \mathfrak{m}, (\mathfrak{P}_N)_{N\in\mathfrak{N}})$ it holds that*

$$\textit{for every } N \in \mathfrak{N} \textit{ there exists } n \in \mathbb{N} \textit{ such that } \mathcal{B}_N \subset \mathcal{B}_{b_n}. \tag{6.2.5}$$

In particular, $(\mathcal{B}_{b_n})_{n\in\mathbb{N}}$ is a filtration on $(\Omega, \mathcal{B})$ such that $\sigma(\{\mathcal{B}_{b_n} \mid n \in \mathbb{N}\}) = \mathcal{B}$, so that for every separable Hilbert space $\mathbb{H}$ we have that

$$\bigcup_{N\in\mathfrak{N}} L^2((\Omega, \mathcal{B}_N, \mathfrak{m}); \mathbb{H}) \textit{ is dense in } L^2((\Omega, \mathcal{B}, \mathfrak{m}); \mathbb{H}). \tag{6.2.6}$$

Proof. Since $\mathfrak{N}$ is unbounded and directed, for every finite subset $\mathfrak{M} \subset \mathfrak{N}$ the quantity

$$\operatorname{succ}(\mathfrak{M}) := \min\{N \in \mathfrak{N} \mid M \precnsim N \, \forall M \in \mathfrak{M}\}$$

is well defined. Let $(a_n)_n \subset \mathbb{N}$ be an enumeration of $\mathfrak{N}$ and consider the following sequence defined by induction

$$b_0 = a_0, \quad b_{n+1} = \operatorname{succ}(\{a_{n+1}, b_n\}), \quad n \in \mathbb{N}.$$

Then $b_n \precnsim b_{n+1}$ for every $n \in \mathbb{N}$ and (6.2.5) holds for $(b_n)_n$ and any $\mathfrak{N}$-refined standard Borel measure space $(\Omega, \mathcal{B}, \mathfrak{m}, (\mathfrak{P}_N)_{N\in\mathfrak{N}})$. □

In the next Lemma we show that, given two distinct points $\omega, \omega'' \in \Omega$, they can always be separated by some partition $\mathfrak{P}_N$ for $N \in \mathfrak{N}$ sufficiently large.

Lemma 6.2.14. *Let $(\Omega, \mathcal{B}, \mathfrak{m}, (\mathfrak{P}_N)_{N\in\mathfrak{N}})$ be a $\mathfrak{N}$-refined standard Borel measure space such that $\mathfrak{N} = (b_n)_{n\in\mathbb{N}}$, where $(b_n)_{n\in\mathbb{N}}$ is a $\prec$-increasing sequence. Then there exists $\Omega_0 \in \mathcal{B}$ with $\mathfrak{m}(\Omega_0) = 0$ such that for every $\omega', \omega'' \in \Omega \setminus \Omega_0$, $\omega' \neq \omega''$ there exists $M \in \mathbb{N}$ such that for every $n \geqslant M$ there are $k', k'' \in I_{b_n}$, $k' \neq k''$ with $\omega' \in \Omega_{b_n,k'}$ and $\omega'' \in \Omega_{b_n,k''}$.*

Proof. Let τ be a Polish topology on Ω such that $\mathcal{B}$ coincides with the Borel sigma algebra generated by τ. By [20, Proposition 6.5.4] there exists a countable family $\mathcal{F}$ of τ-continuous functions $f : \Omega \to [0,1]$ separating the points of Ω, meaning that for every $\omega', \omega'' \in \Omega$, $\omega' \neq \omega''$ there exists $f \in \mathcal{F}$ such that $f(\omega') \neq f(\omega'')$. Since $\mathcal{F} \subset L^2((\Omega, \mathcal{B}, \mathfrak{m}); \mathbb{R})$, by Theorem 6.2.8 with $\mathcal{F}_n := \mathcal{B}_{b_n}$, for every $f \in \mathcal{F}$ there exists a $\mathfrak{m}$-negligible set Ω_f such that

$$\lim_{n\to+\infty} \mathbb{E}_{\mathfrak{m}}[f \mid \sigma(\mathfrak{P}_{b_n})](\omega) = f(\omega) \quad \forall \omega \in \Omega \setminus \Omega_f.$$

Let $\Omega_0 := \cup_{f\in\mathcal{F}} \Omega_f$ and let $\omega', \omega'' \in \Omega \setminus \Omega_0$, $\omega' \neq \omega''$. We can find $f \in \mathcal{F}$ such that $f(\omega') \neq f(\omega'')$. Thus there exists $M \in \mathbb{N}$ such that

$$\mathbb{E}_{\mathfrak{m}}[f \mid \sigma(\mathfrak{P}_{b_n})](\omega') \neq \mathbb{E}_{\mathfrak{m}}[f \mid \sigma(\mathfrak{P}_{b_n})](\omega'') \quad \forall n \geqslant M.$$

Since $\mathbb{E}_{\mathfrak{m}}[f \mid \sigma(\mathfrak{P}_{b_n})]$ is constant on every $\Omega_{b_n,k}$, $k \in I_{b_n}$, we conclude that for every $n \geqslant M$ the points ω' and ω'' belong to different elements of $\mathfrak{P}_{b_n}$. □

Proposition 6.2.15. *Let* $(\Omega, \mathcal{B}, \mathfrak{m}, (\mathfrak{P}_N)_{N\in\mathfrak{N}})$ *and* $(\Omega', \mathcal{B}', \mathfrak{m}', (\mathfrak{P}'_N)_{N\in\mathfrak{N}})$ *be* $\mathfrak{N}$*-refined standard Borel measure spaces such that* $\mathfrak{m}(\Omega) = \mathfrak{m}'(\Omega')$*. Then there exist two measurable functions* $\varphi : \Omega \to \Omega'$ *and* $\psi : \Omega' \to \Omega$ *such that for every* $N \in \mathfrak{N}$ *and every* $I \subset I_N$ *it holds*

$$\psi_I \circ \varphi_I = \mathfrak{i}_{\Omega_I} \ \mathfrak{m}_I\text{-a.e. in } \Omega_I, \qquad \varphi_I \circ \psi_I = \mathfrak{i}_{\Omega'_I} \ \mathfrak{m}'_I\text{-a.e. in } \Omega'_I,$$
$$(\varphi_I)_\sharp \mathfrak{m}_I = \mathfrak{m}'_I, \qquad (\psi_I)_\sharp \mathfrak{m}'_I = \mathfrak{m}_I,$$

where the subscript I *denotes the restriction to* $\cup_{k\in I}\Omega_{N,k}$ *or* $\cup_{k\in I}\Omega'_{N,k}$.

Proof. By Lemma 6.2.13, it is enough to prove the statement in case $\mathfrak{N} = (b_n)_n$, where $(b_n)_n \subset \mathbb{N}$ is a strictly $\prec$-increasing sequence and $(\Omega', \mathcal{B}', \mathfrak{m}', (\mathfrak{P}'_N)_{N\in\mathfrak{N}})$ is given by $([0,1), \mathcal{B}([0,1)), \lambda^c, (\mathfrak{I}_N)_{N\in\mathfrak{N}})$ as in Example 6.2.11 with $c = \mathfrak{m}(\Omega)$. By Lemma 6.2.5, we can find for every $n \in \mathbb{N}$ two measurable maps $\varphi_n : \Omega \to [0,1)$ and ψ_n satisfying the thesis of Lemma 6.2.5 for the standard Borel measure spaces $(\Omega, \mathcal{B}, \mathfrak{m})$ and $([0,1), \mathcal{B}([0,1)), \lambda^c)$ and the $\mathfrak{m} - \lambda^c$ compatible b_n-partitions of $(\Omega, \mathcal{B})$ and $([0,1), \mathcal{B}([0,1)))$ given by $\mathfrak{P}_{b_n}$ and $\mathfrak{I}_{b_n}$, where we recall from Example 6.2.11 that $\mathfrak{I}_{b_n} = (I_{b_n,k})_{k\in I_{b_n}}$ with $I_{b_n,k} = [k/b_n, (k+1)/b_n)$. Since $\sum_n b_n^{-1} < +\infty$, for every $\omega \in \Omega$ the sequence $(\varphi_n(\omega))_n \subset [0,1)$ is Cauchy, hence converges. We thus have the existence of a measurable map $\varphi : \Omega \to [0,1)$ such that

$$\varphi(\omega) = \lim_n \varphi_n(\omega) \quad \forall \omega \in \Omega.$$

If $n \in \mathbb{N}$, $k \in I_{b_n}$ and $\phi \in C_b(I_{b_n,k})$ then

$$\begin{aligned}\int_{I_{b_n,k}} \phi \, d\varphi_\sharp \mathfrak{m} &= \int_{\Omega_{b_n,k}} \phi(\varphi(\omega)) \, d\mathfrak{m}(\omega) = \lim_m \int_{\Omega_{b_n,k}} \phi(\varphi_m(\omega)) \, d\mathfrak{m}(\omega) \\ &= \lim_m \int_{I_{b_n,k}} \phi \, d\lambda^c = \int_{I_{b_n,k}} \phi \, d\lambda^c,\end{aligned}$$

since for m sufficiently large $(\varphi_m)_\sharp \mathfrak{m}|_{\Omega_{b_n,k}} = \lambda^c|_{I_{b_n,k}}$ by Lemma 6.2.5. This shows that $\varphi_\sharp \mathfrak{m}|_{\Omega_{b_n,k}} = \lambda^c|_{I_{b_n,k}}$ for every $k \in I_{b_n}$ and every $n \in \mathbb{N}$. To conclude it is enough to show that φ is $\mathfrak{m}$-essentially injective. Let $\Omega_0 \subset \Omega$ be the $\mathfrak{m}$-negligible subset of Ω given by Lemma 6.2.14 and let $\Omega_1 := \varphi^{-1}(J)$, where

$$J := \{k/b_n \mid k \in I_{b_n}, n \in \mathbb{N}\} \subset [0,1).$$

Since $\lambda^c(J) = 0$, then $\mathfrak{m}(\Omega_1) = 0$; let $\omega', \omega'' \in \Omega \setminus (\Omega_0 \cup \Omega_1)$. Then there exists $M \in \mathbb{N}$ such that ω' and ω'' belong to different elements of $\mathfrak{P}_{b_n}$ for every $n \geqslant M$. By (6.2.2) and Lemma 6.2.14, we can find $k', k'' \in I_{b_M}$ with $k \neq k'$ such that, $\varphi_n(\omega') \in I_{b_M,k'}$ and $\varphi_n(\omega'') \in I_{b_M,k''}$ for every $n \geqslant M$. Thus $\varphi(\omega') \in \overline{I_{b_M,k'}}$ and $\varphi(\omega') \in \overline{I_{b_M,k''}}$; however, since

$$\overline{I_{b_M,k'}} \cap \overline{I_{b_M,k''}} \subset J,$$

it must be that $\varphi(\omega') \neq \varphi(\omega'')$. □

The following result is an application of [24, Theorem 1.1].

Theorem 6.2.16. *Let $(\Omega, \mathcal{B}, \mathbb{P}, (\mathfrak{P}_N)_{N\in\mathfrak{N}})$ be a $\mathfrak{N}$-refined standard Borel probability space such that $\mathfrak{N} = (b_n)_{n\in\mathbb{N}}$, where $(b_n)_{n\in\mathbb{N}}$ is a strictly $\prec$-increasing sequence. Then for every $\boldsymbol{\gamma} \in \Gamma(\mathbb{P},\mathbb{P})$ there exist a strictly increasing sequence $(N_n)_n \subset \mathbb{N}$ and maps $g_n \in S(\Omega, \mathcal{B}, \mathbb{P}; \mathcal{B}_{b_{N_n}})$ such that, for every separable Hilbert space $\mathbb{H}$ and every $X, Y \in L^2((\Omega, \mathcal{B}, \mathbb{P}); \mathbb{H})$, it holds*

$$(X,Y)_\sharp(\mathbf{i}_\Omega, g_n)_\sharp\mathbb{P} \to (X,Y)_\sharp\boldsymbol{\gamma} \textit{ in } \mathcal{P}_2(\mathbb{H}^2). \tag{6.2.7}$$

Proof. We divide the proof in several steps.
(1) *Let $([0,1), \mathcal{B}([0,1)), \lambda^1, (\mathfrak{I}_N)_{N\in\mathfrak{N}})$ be the $\mathfrak{N}$-refined standard Borel probability space of Example 6.2.11 with $c = 1$. Then for every $\boldsymbol{\gamma} \in \Gamma(\lambda^1, \lambda^1)$, there exist a strictly increasing sequence $(N_n)_n \subset \mathbb{N}$ and maps $g_n \in S([0,1), \mathcal{B}([0,1)), \lambda^1; \sigma(\mathfrak{I}_{b_{N_n}}))$ such that*

$$(\mathbf{i}_{[0,1)}, g_n)_\sharp\lambda^1 \to \boldsymbol{\gamma} \textit{ in } \mathcal{P}([0,1)\times[0,1)).$$

Let $\bar{\mathcal{L}}$ be the one dimensional Lebesgue measure restricted to $[0,1]$ and let $\boldsymbol{\gamma} \in \Gamma(\lambda^1, \lambda^1)$. Let $\boldsymbol{\mu} \in \mathcal{P}([0,1]\times[0,1])$ be an extension of $\boldsymbol{\gamma}$ to $[0,1]\times[0,1]$ such that $\boldsymbol{\mu} \in \Gamma(\bar{\mathcal{L}}, \bar{\mathcal{L}})$. In [24, Theorem 1.1] it is proven that it is possible to find a strictly increasing sequence $(N_n)_n \subset \mathbb{N}$ and maps $(f_n)_n \subset S([0,1], \mathcal{B}([0,1]), \bar{\mathcal{L}})$ such that for every $n \in \mathbb{N}$ there exists $\sigma_n \in \mathrm{Sym}(I_{2^{N_n}})$ such that

$$f_n(x) = x - x_{N_n,k} + x_{N_n,\sigma_n(k)}, \quad x \in I_{2^{N_n},k}, \quad k \in I_{2^{N_n}} \tag{6.2.8}$$

satisfying

$$(\mathbf{i}_{[0,1]}, f_n)_\sharp\bar{\mathcal{L}} \to \boldsymbol{\mu} \text{ in } \mathcal{P}([0,1]\times[0,1]). \tag{6.2.9}$$

If we call g_n the restriction of f_n to $[0,1)$, $n \in \mathbb{N}$, we get that g_n is an element of $S([0,1), \mathcal{B}([0,1)), \lambda^1; \sigma(\mathfrak{I}_{b_{N_n}}))$ for every $n \in \mathbb{N}$ and

$$(\mathbf{i}_{[0,1)}, g_n)_\sharp\lambda^1 \to \boldsymbol{\gamma} \text{ in } \mathcal{P}([0,1)\times[0,1)).$$

This proves the first step only in case $b_n = 2^n$. However, it can be easily checked that the proof of [24, Theorem 1.1] does not depend on the specific choice of the sequence b_n but it is enough that $b_n \precnsim b_{n+1}$ for every $n \in \mathbb{N}$ so that the length of the interval $[k/b_n, (k+1)/b_n]$ goes to 0 faster than 2^{-n} as $n \to +\infty$. This concludes the proof of the first claim.
(2) *Let $([0,1), \mathcal{B}([0,1)), \lambda^1, (\mathfrak{I}_N)_{N\in\mathfrak{N}})$ be the $\mathfrak{N}$-refined standard Borel probability space of Example 6.2.11 with $c = 1$. Then for every $\boldsymbol{\gamma} \in \Gamma(\lambda^1, \lambda^1)$, there exist a strictly increasing sequence $(N_n)_n \subset \mathbb{N}$ and maps $g_n \in S([0,1), \mathcal{B}([0,1)), \lambda^1; \sigma(\mathfrak{I}_{b_{N_n}}))$ such that, for every separable Hilbert space $\mathbb{H}$ and every $X, Y \in L^2(([0,1), \mathcal{B}([0,1)), \lambda^1); \mathbb{H})$, it holds*

$$(X,Y)_\sharp(\mathbf{i}_{[0,1)}, g_n)_\sharp\lambda^1 \to (X,Y)_\sharp\boldsymbol{\gamma} \textit{ in } \mathcal{P}_2(\mathbb{H}^2).$$

Let $\boldsymbol{\gamma} \in \Gamma(\lambda^1, \lambda^1)$ and let $(g_n)_n$ be the sequence given by (1) for $\boldsymbol{\gamma}$. Let $\mathbb{H}$ be any separable Hilbert space and let $X, Y \in L^2(([0,1), \mathcal{B}([0,1)), \lambda^1); \mathbb{H})$. Observe that for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset [0,1)$ such that the restrictions of X and Y to K_ε are continuous in K_ε and $\lambda^1([0,1)\setminus K_\varepsilon) < \varepsilon$, so that, setting $\gamma_n :=$

$(\boldsymbol{i}_{[0,1)}, g_n)_\sharp\lambda^1$, $n \in \mathbb{N}$, we have that $\gamma_n([0,1)^2 \setminus K^2_\varepsilon) \leqslant 2\varepsilon$ for every $n \in \mathbb{N}$. By [5, Proposition 5.1.10] and (1), $(X,Y)_\sharp(\boldsymbol{i}_{[0,1)}, g_n)_\sharp\lambda^1 \to (X,Y)_\sharp\boldsymbol{\gamma}$ in $\mathcal{P}(\mathbb{H}^2)$; moreover

$$\int \left(|x|^2 + |y|^2\right) d(X,Y)_\sharp(\boldsymbol{i}_{[0,1)}, g_n)_\sharp\lambda^1 = \mathbb{E}_{\lambda^1}[|X|^2 + |Y|^2] = \int \left(|x|^2 + |y|^2\right) d(X,Y)_\sharp\boldsymbol{\gamma},$$

hence the conclusion by (2.4.1).

(3) *Conclusion*. Let $\boldsymbol{\gamma} \in \Gamma(\mathbb{P},\mathbb{P})$ and let $\varphi : \Omega \to [0,1)$ and $\psi : [0,1) \to \Omega$ be the maps given by Proposition 6.2.15 for the $\mathfrak{N}$-refined standard Borel probability spaces $(\Omega, \mathcal{B}, \mathbb{P}, (\mathfrak{P}_N)_{N\in\mathfrak{N}})$ and $([0,1), \mathcal{B}([0,1)), \lambda^1, (\mathfrak{I}_N)_{N\in\mathfrak{N}})$, where the latter is as in Example 6.2.11 with $c = 1$. If we define $\boldsymbol{\gamma}' := (\varphi,\varphi)_\sharp\boldsymbol{\gamma}$, we have that $\boldsymbol{\gamma}' \in \Gamma(\lambda^1,\lambda^1)$ so that we can find a strictly increasing sequence $(N_n)_n \subset \mathbb{N}$ and maps $g'_n \in S([0,1), \mathcal{B}([0,1)), \lambda^1; \sigma(\mathfrak{I}_{b_{N_n}}))$ as in step (2). Let us define

$$g_n := \psi \circ g'_n \circ \varphi, \quad n \in \mathbb{N}.$$

Then, up to change each g_n on a $\mathbb{P}$-negligible set of points, we can assume that $g_n \in S(\Omega, \mathcal{B}, \mathbb{P}; \mathcal{B}_{b_{N_n}})$. Let $\mathbb{H}$ be a separable Hilbert space and let $X, Y \in L^2(\Omega, \mathcal{B}, \mathbb{P}); \mathbb{H})$. If we define $X' := X \circ \psi$ and $Y' := Y \circ \psi$, we get that $X', Y' \in L^2([0,1), \mathcal{B}([0,1)), \lambda^1); \mathbb{H})$. By step (2) we thus get

$$(X',Y')_\sharp(\boldsymbol{i}_{[0,1)}, g'_n)_\sharp\lambda^1 \to (X',Y')_\sharp\boldsymbol{\gamma}' \text{ in } \mathcal{P}_2(\mathbb{H}^2)$$

which is equivalent to (6.2.7). □

Corollary 6.2.17. *Let $(\Omega, \mathcal{B}, \mathbb{P}, (\mathfrak{P}_N)_{N\in\mathfrak{N}})$ be a $\mathfrak{N}$-refined standard Borel probability space such that $\mathfrak{N} = (b_n)_{n\in\mathbb{N}}$, where $(b_n)_{n\in\mathbb{N}}$ is a strictly $\prec$-increasing sequence. Then for every $\boldsymbol{\gamma} \in \Gamma(\mathbb{P},\mathbb{P})$ there exist a strictly increasing sequence $(N_n)_n \subset \mathbb{N}$ and maps $g_n \in S(\Omega, \mathcal{B}, \mathbb{P}; \mathcal{B}_{b_{N_n}})$ such that, for every Polish topology τ on Ω generating $\mathcal{B}$, it holds*

$$(\boldsymbol{i}_\Omega, g_n)_\sharp\mathbb{P} \to \boldsymbol{\gamma} \text{ in } \mathcal{P}(\Omega \times \Omega, \tau \otimes \tau),$$

where $\tau \otimes \tau$ is the product topology on $\Omega \times \Omega$.

Proof. By Theorem 6.2.16 we have the existence of a strictly increasing sequence $(N_n)_n \subset \mathbb{N}$ and maps $g_n \in S(\Omega, \mathcal{B}, \mathbb{P}; \mathcal{B}_{b_{N_n}})$ such that, choosing the separable Hilbert space $\mathbb{R}$, we get

$$(\varphi_1,\varphi_2)_\sharp(\boldsymbol{i}_\Omega, g_n)_\sharp\mathbb{P} \to (\varphi_1,\varphi_2)_\sharp\boldsymbol{\gamma} \text{ in } \mathcal{P}_2(\mathbb{R}^2)$$

for every $\varphi_1, \varphi_2 \in C_b(\Omega,\tau) \subset L^2((\Omega, \mathcal{B}, \mathbb{P}); \mathbb{R})$. By the $\mathcal{P}_2(\mathbb{R}^2)$ convergence we get (see e.g.[5, Proposition 7.1.5, Lemma 5.1.7]) that

$$\int_{\Omega\times\Omega} h(\varphi_1(\omega_1), \varphi_2(\omega_2))\, d\boldsymbol{\gamma}_n(\omega_1,\omega_2) \to \int_{\Omega\times\Omega} h(\varphi_1(\omega_1), \varphi_2(\omega_2))\, d\boldsymbol{\gamma}(\omega_1,\omega_2)$$

for every continuous function $h : \mathbb{R}^2 \to \mathbb{R}$ with less than 2-growth, where $\boldsymbol{\gamma}_n = (\boldsymbol{i}_\Omega, g_n)_\sharp\mathbb{P}$, $n \in \mathbb{N}$. Choosing $h(x,y) := xy$, we get that

$$\int_{\Omega\times\Omega} \varphi_1(\omega_1)\varphi_2(\omega_2)\, d\boldsymbol{\gamma}_n(\omega_1,\omega_2) \to \int_{\Omega\times\Omega} \varphi_1(\omega_1)\varphi_2(\omega_2)\, d\boldsymbol{\gamma}(\omega_1,\omega_2).$$

(6.2.10)

for every $\varphi_1, \varphi_2 \in C_b(\Omega, \tau)$. Let $\mathcal{A} \subset C_b(\Omega, \tau)$ be a unital subalgebra whose induced initial topology on Ω coincides with τ (e.g. the subset of d-Lipschitz continuous and bounded functions for a complete distance d inducing τ). It is easy to check that

$$\mathcal{A} \otimes \mathcal{A} := \left\{ \sum_{i=1}^{n} \varphi_1^i \otimes \varphi_2^i \mid (\varphi_1^i)_{i=1}^n, (\varphi_2^i)_{i=1}^n \subset \mathcal{A}, n \in \mathbb{N} \right\} \subset C_b(\Omega \times \Omega, \tau \otimes \tau)$$

is a unital subalgebra whose induced initial topology on $\Omega \times \Omega$ coincides with $\tau \otimes \tau$. By (6.2.10) we thus have that

$$\int_{\Omega \times \Omega} \varphi \, d\gamma_n \to \int_{\Omega \times \Omega} \varphi \, d\gamma \quad \forall \varphi \in \mathcal{A} \otimes \mathcal{A}.$$

We conclude by Lemma 3.1.6. □

The following result is a consequence of [97, Theorem B] (see also [3, Theorem 2.1, Theorem 9.3]).

Proposition 6.2.18. *Let $(\Omega, \mathcal{B}, \mathbb{P})$ be a standard Borel probability space, let $\mathbb{H}$ be a separable Hilbert space and let us denote $\mathcal{H} := L^2((\Omega, \mathcal{B}, \mathbb{P}); \mathbb{H})$. If $\mu, \nu \in \mathcal{P}_2(\mathbb{H})$ and $X \in \mathcal{H}$ is s.t. $X_\sharp \mathbb{P} = \mu$, then, for every $\varepsilon > 0$, there exists $Y \in \mathcal{H}$ s.t. $Y_\sharp \mathbb{P} = \nu$ and*

$$|X - Y|_{\mathcal{H}} \leqslant W_2(\mu, \nu) + \varepsilon.$$

Proof. Let $\gamma \in \Gamma_o(\mu, \nu)$; we split μ into its atomless and atomic parts, μ_c and μ_d respectively. Hence, there exists a sequence $(x_n)_n \subset \mathbb{H}$ s.t.

$$\mu_d = \sum_{n \in \mathbb{N}} \mu_n, \quad \mu_n := a_n \delta_{x_n}, \quad a_n := \mu(\{x_n\})$$

and μ_c is atomless (i.e. $\mu_c(\{x\}) = 0$ for every $x \in \mathbb{H}$). Let

$$\Omega_n := X^{-1}(\{x_n\}), \ n \in \mathbb{N}, \quad \Omega_c := \Omega \setminus \bigcup_n \Omega_n.$$

We define the Borel functions

$$X_n := X|_{\Omega_n}, \ n \in \mathbb{N}, \quad X_c := X|_{\Omega_c}$$

and the nonnegative Borel measures

$$\mathbb{P}_n := \mathbb{P}|_{\Omega_n}, \ n \in \mathbb{N}, \quad \mathbb{P}_c := \mathbb{P}|_{\Omega_c}.$$

It is clear that $(X_n)_\sharp \mathbb{P}_n = \mu_n$ for every $n \in \mathbb{N}$ and that $(X_c)_\sharp \mathbb{P}_c = \mu_c$. Let $\{\gamma_x\}_{x \in \mathbb{H}} \subset \mathcal{P}(\mathbb{H})$ be the disintegration of γ w.r.t. μ, we define the nonnegative Borel measures

$$\gamma_n := \mu_n \otimes \gamma_{x_n}, \ n \in \mathbb{N}, \quad \gamma_c := \int_{\mathbb{H}} \gamma_x \, d\mu_c(x)$$

with second marginals $\nu_n = a_n \gamma_{x_n}$, $n \in \mathbb{N}$ and ν_c, respectively. By Corollary 6.2.3, we can find functions $Y_n \in L^2(\Omega_n, \mathbb{P}_n; \mathbb{H})$ s.t. $(Y_n)_\sharp \mathbb{P}_n = \nu_n$ so that $(X_n, Y_n)_\sharp \mathbb{P}_n = \boldsymbol{\gamma}_n$ for every $n \in \mathbb{N}$. This gives

$$\int_{\Omega_n} |X_n - Y_n|^2 \, d\mathbb{P}_n = \int_{\mathbb{H}\times\mathbb{H}} |x-y|^2 \, d\boldsymbol{\gamma}_n(x,y). \tag{6.2.11}$$

On the other hand, we can apply [97, Theorem B] to the Polish space $\mathbb{H}$, the atomless measure μ_c, the measure ν_c and the plan $\boldsymbol{\gamma}_c \in \Gamma(\mu_c, \nu_c)$ so that, for any $\varepsilon > 0$, we can find $t := t_\varepsilon : \mathbb{H} \to \mathbb{H}$ s.t. $t_\sharp \mu_c = \nu_c$ and

$$\int_{\mathbb{H}} |x - t(x)|^2 \, d\mu_c(x) \leqslant \int_{\mathbb{H}\times\mathbb{H}} |x-y|^2 \, d\boldsymbol{\gamma}_c(x,y) + \varepsilon^2.$$

If we define $Y_c \in L^2(\Omega_c, \mathbb{P}_c; \mathbb{H})$ as $Y_c := t \circ X_c$, we get

$$\int_{\Omega_c} |X_c - Y_c|^2 \, d\mathbb{P}_c \leqslant \int_{\mathbb{H}\times\mathbb{H}} |x-y|^2 \, d\boldsymbol{\gamma}_c(x,y) + \varepsilon^2 \tag{6.2.12}$$

with $(Y_c)_\sharp \mathbb{P}_c = \nu_c$. Finally we define $Y \in L^2(\Omega, \mathbb{P}; \mathbb{H})$ as

$$Y(\omega) := \begin{cases} Y_n(\omega) & \text{if } \omega \in \Omega_n, \; n \in \mathbb{N}, \\ Y_c(\omega) & \text{if } \omega \in \Omega_c. \end{cases}$$

By (6.2.11) and (6.2.12) we have that

$$\begin{aligned} |X - Y|^2_{\mathcal{H}} = \int_\Omega |X-Y|^2 \, d\mathbb{P} &= \sum_n \int_{\Omega_n} |X_n - Y_n|^2 \, d\mathbb{P}_n + \int_{\Omega_c} |X_c - Y_c|^2 \, d\mathbb{P}_c \\ &\leqslant \sum_n \int_{\mathbb{H}\times\mathbb{H}} |x-y|^2 \, d\boldsymbol{\gamma}_n(x,y) + \int_{\mathbb{H}\times\mathbb{H}} |x-y|^2 \, d\boldsymbol{\gamma}_c(x,y) + \varepsilon^2 \\ &= \int_{\mathbb{H}\times\mathbb{H}} |x-y|^2 \, d\boldsymbol{\gamma}(x,y) + \varepsilon^2 = W_2^2(\mu,\nu) + \varepsilon^2. \end{aligned}$$

Hence the conclusion, noting that $Y_\sharp \mathbb{P} = \nu$. □

6.3 A STRONG-WEAK TOPOLOGY FOR PROBABILITIES IN PRODUCT SPACES

In this short section we denote by X, Y two separable Hilbert spaces. On the product space $X \times Y$ we consider the product Hilbert norm and $\mathcal{P}_2(X \times Y)$ with the 2-Wasserstein distance. We will endow $\mathcal{P}_2(X \times Y)$ with a weaker topology which is connected to the strong-weak topology on $X \times Y$, i.e. the product topology of $X^s \times Y^w$, where the superscript s (resp. w) means that we are considering the strong (resp. weak) topology on the corresponding Hilbert space. The proof of the results presented here can be found in [86].

Let us consider the space $C_2^{sw}(X \times Y)$ of functions $\zeta : X \times Y \to \mathbb{R}$ such that

$$\zeta \text{ is sequentially continuous in } X^s \times Y^w, \tag{6.3.1}$$

$$\forall\, \varepsilon > 0 \; \exists\, A_\varepsilon \geqslant 0 : |\zeta(x,y)| \leqslant A_\varepsilon(1 + |x|_X^2) + \varepsilon |y|_Y^2 \quad \forall\, (x,y) \in X \times Y. \tag{6.3.2}$$

Notice that, if $\zeta \in C_2^{sw}(X \times Y)$, then it has quadratic growth. On $C_2^{sw}(X)$ we consider the norm

$$\|\zeta\|_{C_2^{sw}(X)} := \sup_{(x,y)\in X} \frac{|\zeta(x,y)|}{1+|x|_X^2+|y|_Y^2}.$$

Lemma 6.3.1. *$(C_2^{sw}(X \times Y), \|\cdot\|_{C_2^{sw}(X\times Y)})$ is a Banach space.*

Definition 6.3.2 (Topology of $\mathcal{P}_2^{sw}(X \times Y)$, [86])**.** We denote by $\mathcal{P}_2^{sw}(X \times Y)$ the space $\mathcal{P}_2(X \times Y)$ endowed with the coarsest topology which makes the following functions continuous

$$\mu \mapsto \int \zeta(x,y)\, d\mu(x,y), \quad \zeta \in C_2^{sw}(X \times Y).$$

The topology of $\mathcal{P}_2(X \times Y)$ is finer than the topology of $\mathcal{P}_2^{sw}(X \times Y)$, and the latter is finer than the topology of $\mathcal{P}(X^s \times Y^w)$. Notice that, if $B : X \times Y \to \mathbb{R}$ is a bounded bilinear form, then it belongs to $C_2^{sw}(X \times Y)$, hence for every net $(\mu_\alpha)_{\alpha\in\mathbb{A}} \subset \mathcal{P}(X \times Y)$ indexed by a directed set $\mathbb{A}$, we have

$$\lim_{\alpha\in\mathbb{A}} \mu_\alpha = \mu \quad \text{in } \mathcal{P}_2^{sw}(X \times Y) \quad \Rightarrow \quad \lim_{\alpha\in\mathbb{A}} \int B\, d\mu_\alpha = \int B\, d\mu. \tag{6.3.3}$$

The following result presents many useful properties of the $\mathcal{P}_2^{sw}(X \times Y)$-topology.

Proposition 6.3.3.

1. *Assume that $(\mu_\alpha)_{\alpha\in\mathbb{A}} \subset \mathcal{P}_2(X \times Y)$ is a net indexed by the directed set $\mathbb{A}$, $\mu \in \mathcal{P}_2(X \times Y)$ and they satisfy*
   a) $\mu_\alpha \to \mu$ *in* $\mathcal{P}(X^s \times Y^w)$,
   b) $\lim_{\alpha\in\mathbb{A}} \int |x|_X^2\, d\mu_\alpha(x,y) = \int |x|_X^2\, d\mu(x,y)$,
   c) $\sup_{\alpha\in\mathbb{A}} \int |y|_Y^2\, d\mu_\alpha(x,y) < \infty$,

   then $\mu_\alpha \to \mu$ in $\mathcal{P}_2^{sw}(X \times Y)$. The converse property holds for sequences: if $\mathbb{A} = \mathbb{N}$ and $\mu_n \to \mu$ in $\mathcal{P}_2^{sw}(X \times Y)$ as $n \to \infty$ then properties (a), (b), (c) hold.

2. *For every compact set $\mathcal{K} \subset \mathcal{P}_2(X^s)$ and every constant $c < \infty$ the sets*

$$\mathcal{K}_c := \left\{\mu \in \mathcal{P}_2(X \times Y) : \pi^X_\sharp \mu \in \mathcal{K}, \quad \int |y|_Y^2\, d\mu(x,y) \leqslant c\right\}$$

   are compact and metrizable in $\mathcal{P}_2^{sw}(X \times Y)$ (in particular they are sequentially compact).

Notice that the topology $\mathcal{P}_2^{ws}(X \times Y)$ is strictly weaker than $\mathcal{P}_2(X \times Y)$ also in case Y has finite dimension. Indeed, the function $(x,y) \mapsto |y|_Y^2$ is not an element of $C_2^{sw}(X \times Y)$, so that the convergence of the quadratic moment w.r.t. y is not guaranteed.

6.4 TRIPLANS AND INJECTIVITY OF INTERPOLATION MAPS

In this section, we state and prove the following preliminary results of independent interest. Their importance for our study will be clear when proving Lemma 7.8.2 and more specifically in Section 9.4, since they will be involved in the proof of one of the main results of the Section: Theorem 9.4.16. In this section $\mathbb{H}$ is a separable Hilbert space and we define $x^{t,\theta}$ as

$$x^{t,\theta}(x_1,x_2,x_3) := (1-\theta)\,[(1-t)x_1 + tx_2] + \theta x_3, \quad t,\theta \in [0,1]. \tag{6.4.1}$$

Proposition 6.4.1. *Let $\boldsymbol{\gamma} \in \mathcal{P}(\mathbb{H}^3)$ be such that $\pi^{1,2}_\sharp\boldsymbol{\gamma}$ and $\pi^{2,3}_\sharp\boldsymbol{\gamma}$ are optimal. Then, for every $t,\theta \in (0,1)$, $(x^{t,\theta},\pi^2)_\sharp\boldsymbol{\gamma}$ is the unique element of $\Gamma_o(x^{t,\theta}_\sharp\boldsymbol{\gamma},\pi^2_\sharp\boldsymbol{\gamma})$ and it is induced by a map, i.e. there exists a Borel map $g:\mathbb{H}\to\mathbb{H}$ such that*

$$(x^{t,\theta},\pi^2)_\sharp\boldsymbol{\gamma} = (\boldsymbol{i}_{\mathbb{H}}, g)_\sharp(x^{t,\theta}_\sharp\boldsymbol{\gamma}).$$

Proof. Let $t,\theta \in (0,1)$ be fixed. By Theorem 2.4.4, to prove the statement it is enough to show that

$$(x^{t,\theta},\pi^2)_\sharp\boldsymbol{\gamma} = (x^q,x^1)_\sharp\boldsymbol{\alpha} \tag{6.4.2}$$

for some optimal plan $\boldsymbol{\alpha} \in \mathcal{P}(\mathbb{H}^2)$ and $q \in (0,1)$. Indeed, this implies that $(x^{t,\theta},\pi^2)_\sharp\boldsymbol{\gamma}$ induces the restriction of a constant speed geodesic to $[q,1]$ and thus it is optimal, unique and concentrated on a map. It is easy to check that (6.4.2) holds with $q := (1-\theta)t$ and $\boldsymbol{\alpha} := (x^p \circ \pi^{1,3},\pi^2)_\sharp\boldsymbol{\gamma}$, where $p := \frac{\theta}{1-(1-\theta)t}$. We are left to show that $\boldsymbol{\alpha}$ is optimal and, thanks to (2.4.9), it is enough to show that it is concentrated on a monotone set. This immediately follows if we show that

$$\langle y_1 - y_1', x_1 - x_1'\rangle \geqslant 0 \quad \forall (y_1,x_1),(y_1',x_1') \in (x^p\circ\pi^{1,3},\pi^2)(\operatorname{supp}\boldsymbol{\gamma}).$$

Let $(y_1,x_2),(y_1',x_2') \in (x^p\circ\pi^{1,3},\pi^2)(\operatorname{supp}\boldsymbol{\gamma})$; we can find $x_1,x_3,x_1',x_3' \in \mathbb{H}$ such that $y_1 = x^p(x_1,x_3)$, $y_1' = x^p(x_1',x_3')$ and (x_1,x_2,x_3), $(x_1',x_2',x_3') \in \operatorname{supp}\boldsymbol{\gamma}$. Then

$$\langle y_1 - y_1', x_1 - x_1'\rangle = (1-p)\langle x_1 - x_1', x_2 - x_2'\rangle + p\langle x_3 - x_3', x_2 - x_2'\rangle \geqslant 0,$$

where we have used the monotonicity of the supports of $\pi^{1,2}_\sharp\boldsymbol{\gamma}$ and $\pi^{2,3}_\sharp\boldsymbol{\gamma}$ coming from their optimality, thanks again to (2.4.9). □

Given two pairs of points (a',b') and (a'',b'') in $\mathbb{H}^2$ it is easy to check that

$$\begin{aligned}(1-t)a' + tb' \neq (1-t)a''+tb'' &\quad \text{for every } t \in (0,1)\\ &\Leftrightarrow \\ b''-b' \notin \Big\{-s(a''-a') &: s>0\Big\}.\end{aligned} \tag{6.4.3}$$

In particular, given a set $A \subset \mathbb{H}$ we consider the set of directions

$$\operatorname{dir}(A) := \Big\{s(a'-a'') : s\in\mathbb{R},\ a',a''\in A\Big\} = \bigcup_{s\in\mathbb{R}} s(A-A) \tag{6.4.4}$$

If $B \subset \mathbb{H}$ satisfies

$$(B-B)\cap\operatorname{dir}(A) = \{0\} \tag{6.4.5}$$

then for every $t\in(0,1)$ the map $x^t:\mathbb{H}^2\to\mathbb{H}$ is injective on $A\times B$.

Proposition 6.4.2. *Let $\boldsymbol{\gamma} \in \mathcal{P}(\mathbb{H}^3)$ be such that $\pi^{1,2}_\sharp \boldsymbol{\gamma}$ and $\pi^{2,3}_\sharp \boldsymbol{\gamma}$ are optimal. Set*

$$A_1 := \pi^1(\operatorname{supp} \boldsymbol{\gamma}), \quad A_3 := \pi^3(\operatorname{supp} \boldsymbol{\gamma})$$

and suppose that $(A_3 - A_3) \cap \operatorname{dir}(A_1) = \{0\}$. Then, for every $t, \theta \in (0,1)$, $x^{t,\theta}$ is injective on $\operatorname{supp} \boldsymbol{\gamma}$.

Proof. Let $t, \theta \in (0,1)$ and let $(x_1, x_2, x_3), (x_1', x_2', x_3') \in \operatorname{supp} \boldsymbol{\gamma}$ be such that $x^{t,\theta}(x_1, x_2, x_3) = x^{t,\theta}(x_1', x_2', x_3')$. Setting $y_t := (1-t)x_1 + tx_2$, $y_t' := (1-t)x_1' + tx_2'$, we obtain

$$(1-\theta)y_t + \theta x_3 = (1-\theta)y_t' + \theta x_3'$$

so that

$$y_t - y_t' = -\frac{\theta}{1-\theta}(x_3 - x_3'). \tag{6.4.6}$$

Since $\pi^{1,2}_\sharp \boldsymbol{\gamma}$ and $\pi^{2,3}_\sharp \boldsymbol{\gamma}$ are optimal, their supports are monotone by (2.4.9), thus

$$\langle x_1 - x_1', x_2 - x_2' \rangle \geqslant 0, \quad \langle x_3 - x_3', x_2 - x_2' \rangle \geqslant 0$$

which gives

$$\langle y_t - y_t', x_2 - x_2' \rangle \geqslant t|x_2 - x_2'|^2.$$

Then

$$0 \geqslant -\frac{\theta}{1-\theta}\langle x_3 - x_3', x_2 - x_2' \rangle \geqslant t|x_2 - x_2'|^2$$

so that $x_2 = x_2'$. Inserting this in (6.4.6), we obtain

$$x_3 - x_3' = -\frac{(1-\theta)(1-t)}{\theta}(x_1 - x_1').$$

Since by assumption $(A_3 - A_3) \cap \operatorname{dir}(A_1) = \{0\}$, then we conclude that $x_1 = x_1'$ and $x_3 = x_3'$, proving the sought injectivity. □

In the discrete setting, we can prove the following result which gives the possibility to displace the elements of a finite set B in order to satisfy condition (6.4.5) with respect to a fixed finite set A.

Proposition 6.4.3. *Assume that $\dim \mathbb{H} \geqslant 2$ and $A \subset \mathbb{H}$ is a finite set. For every finite set of distinct points $B = \{b_n\}_{n=1}^N \subset \mathbb{H}$ there exists a finite set $B' := \{b_n'\}_{n=1}^N$ of distinct points with $|b_n' - b_n| < 1$ such that, setting*

$$b_n(s) := (1-s)b_n + sb_n', \quad B(s) := \{b_n(s)\}_{n=1}^N, \tag{6.4.7}$$

we have that $\#B(s) = N$ for all $s \in [0,1]$ and

$$(B(s) - B(s)) \cap \operatorname{dir}(A) = \{0\} \quad \text{for every } s \in (0,1]. \tag{6.4.8}$$

In particular, for every $t \in (0,1)$ the restriction of the map x^t to $A \times B(s)$ is injective for every $s \in (0,1]$.

Proof. We split the proof of the Proposition in two steps.

Claim 1: *there exists a finite set of distinct points* $B'' := \{b''_n\}_{n=1}^N$ *with* $|b''_n - b_n| < 1$ *satisfying*

$$(B'' - B'') \cap \mathrm{dir}(A) = \{0\}. \tag{6.4.9}$$

We can argue by induction with respect to the cardinality N of the set B. The statement is obvious in the case $N = 1$ (it is sufficient to choose $b''_1 := b_1$).

Let us assume that the property holds for all the sets of cardinality $N - 1 \geqslant 1$. We can thus find a finite set of distinct points $B''_{N-1} = \{b''_n\}_{n=1}^{N-1}$ satisfying $(B''_{N-1} - B''_{N-1}) \cap \mathrm{dir}(A) = \{0\}$. We look for a point $b''_N \in U \setminus B''_{N-1}$, where $U := \{x \in \mathbb{H} : |x - b_N| < 1\}$, such that $B''_N := B''_{N-1} \cup \{b''_N\}$ satisfies (6.4.9). b''_N should therefore satisfy

$$b''_N \in U, \quad b''_N - b''_n \notin \mathrm{dir}(A) \quad \text{for every } n \in \{1, \cdots, N-1\}.$$

Such a point surely exists, since $\mathrm{dir}(A)$ is a closed set with empty interior (here we use the fact that the dimension of $\mathbb{H}$ is at least 2) and the union $\bigcup_{n=1}^{N-1} \big(b''_n + \mathrm{dir}(A)\big)$ has empty interior as well, so that it cannot contain the open set U.

Claim 2: *If* B'' *satisfies the properties of the previous claim, then there exists* $\delta \in (0, 1]$ *such that setting*

$$b'_n := (1 - \delta) b_n + \delta b''_n, \tag{6.4.10}$$

the set $B' = \{b'_n\}_{n=1}^N$ *satisfies the thesis.*

We denote by a the cardinality $\#A$ of A and we first make a simple remark: for every $z, z'' \in \mathbb{H}$

$$\#\{s \in [0, 1] : z(s) := (1 - s)z + sz'' \in \mathrm{dir}(A)\} > a^2 \quad \Rightarrow \quad z, z'' \in \mathrm{dir}(A). \tag{6.4.11}$$

Indeed, the set $A - A$ contains at most a^2 distinct elements, so that if the left hand side of (6.4.11) is true, then there are at least two distinct values $s_1, s_2 \in [0, 1]$, $r_1, r_2 \in \mathbb{R}$ and a vector $w \in A - A$ such that $(1 - s_1)z + s_1 z'' = r_1 w$, $(1 - s_2)z + s_2 z'' = r_2 w$. We then get

$$z(s) = z(s_1) + \frac{s - s_1}{s_2 - s_1}(z(s_2) - z(s_1)) = r_1 w + \frac{(s - s_1)(r_2 - r_1)}{s_2 - s_1} w \in \mathrm{dir}(A)$$

for every $s \in [0, 1]$,hence (6.4.11). As a particular consequence of (6.4.11) we get that if z'' does not belong to $\mathrm{dir}(A)$, then the set $\{s \in (0, 1] : z(s) := (1 - s)z + sz'' \in \mathrm{dir}(A)\}$ is finite, so that

$$\forall\, z, z'' \in \mathbb{H} :\; z'' \notin \mathrm{dir}(A) \quad \Rightarrow \quad \exists\, \delta > 0 :\; (1 - s)z + sz'' \notin \mathrm{dir}(A) \quad \forall\, s \in (0, \delta]. \tag{6.4.12}$$

Let us now apply property (6.4.12) to all the pairs (z, z'') of the form $z = b_n - b_m$, $z'' = b''_n - b''_m$, $n, m \in \{1, \cdots, N\}$, with $n \neq m$. Since $b''_n - b''_m \notin \mathrm{dir}(A)$ we deduce that there exists $\delta_{n,m} > 0$ such that

$$(1 - s)(b_n - b_m) + s(b''_n - b''_m) \notin \mathrm{dir}(A) \quad \text{for every } s \in (0, \delta_{n,m}]. \tag{6.4.13}$$

Setting

$$\tilde{\delta} := \min\{|b_n - b_m| \,:\, n, m \in \{1, \dots, N\},\, n \neq m\} > 0$$

and choosing $\delta := \min_{n,m}\{\delta_{n,m}, \tilde{\delta}/3\} > 0$, then it is not difficult to check that B' satisfies the thesis, with b'_n as in (6.4.10). Indeed, $|b_n - b'_n| = \delta|b_n - b''_n| < 1$, and for every $s \in [0,1]$ and n we get

$$b_n(s) := (1-s)b_n + sb'_n = (1-s)b_n + s(1-\delta)b_n + s\delta b''_n = (1-\delta s)b_n + \delta s b''_n$$

so that

$$b_n(s) - b_m(s) = (1-\delta s)(b_n - b_m) + \delta s(b''_n - b''_m) \notin \mathrm{dir}(A)$$

thanks to (6.4.13) and the fact that $s\delta \leqslant \delta_{n,m}$. □

6.5 LOCAL OPTIMALITY OF DISCRETE COUPLINGS

We want to study the behaviour of the Wasserstein distance along couplings between discrete measures. The main quantitative information is contained in the following lemma.

Lemma 6.5.1. *Let $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{H})$, $\boldsymbol{\gamma} \in \Gamma(\mu_0, \mu_1)$. If μ_0 has finite support $S = \{\bar{x}_1, \cdots, \bar{x}_M\}$ with $\delta := \min\left\{|\bar{x}_i - \bar{x}_j| : i, j \in \{1, \cdots, M\},\, i \neq j\right\} > 0$ and*

$$\sup\left\{|y - x| : (x,y) \in \mathrm{supp}\,\boldsymbol{\gamma}\right\} \leqslant \delta/2 \tag{6.5.1}$$

then $\boldsymbol{\gamma} \in \Gamma_o(\mu_0, \mu_1)$ and $W_2^2(\mu_0, \mu_1) = \int |y - x|^2\, d\boldsymbol{\gamma}$.

Proof. It is sufficient to prove that the support of γ satisfies the cyclical monotonicity condition (2.4.9).

If $\{(x_n, y_n)\}_{n=1}^N$ are points in $\mathrm{supp}\,\boldsymbol{\gamma}$ with $x_0 := x_N$ and $x_n \neq x_{n-1}$ then

$$\begin{aligned}
\langle y_n, x_n - x_{n-1}\rangle &= \langle y_n - x_n, x_n - x_{n-1}\rangle + \langle x_n, x_n - x_{n-1}\rangle \\
&\geqslant -\frac{\delta}{2}|x_n - x_{n-1}| + \frac{1}{2}|x_n - x_{n-1}|^2 + \frac{1}{2}|x_n|^2 - \frac{1}{2}|x_{n-1}|^2 \\
&\geqslant \frac{1}{2}|x_n|^2 - \frac{1}{2}|x_{n-1}|^2
\end{aligned}$$

since $|y_n - x_n| \leqslant \delta/2$ and $|x_n - x_{n-1}| \geqslant \delta$. If $x_n = x_{n-1}$ we trivially have $\langle y_n, x_n - x_{n-1}\rangle = \frac{1}{2}|x_n|^2 - \frac{1}{2}|x_{n-1}|^2$, so that

$$\sum_{n=1}^N \langle y_n, x_n - x_{n-1}\rangle \geqslant \sum_{n=1}^N \frac{1}{2}|x_n|^2 - \frac{1}{2}|x_{n-1}|^2 = \frac{1}{2}|x_N|^2 - \frac{1}{2}|x_0|^2 = 0.$$

□

As a consequence we obtain the following result.

Theorem 6.5.2. *Let $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{H})$ be two measures with finite support, $\boldsymbol{\gamma} \in \Gamma(\mu_0, \mu_1)$ and $\mu_t := (x^t)_\sharp \boldsymbol{\gamma}$, $t \in [0,1]$. Then the following properties hold.*

1. *For every* $s \in [0,1]$ *there exists* $\delta > 0$ *such that for every* $t \in [0,1]$ *with* $|t-s| \leqslant \delta$ $\gamma_{s,t} := (x^s, x^t)_\sharp \gamma$ *is an optimal plan between* μ_s *and* μ_t*, so that*

$$W_2^2(\mu_s,\mu_t) = \int |y-x|^2 \, d\gamma_{s,t} = |t-s|^2 \int |y-x|^2 \, d\gamma(x,y). \tag{6.5.2}$$

2. *There exist a finite number of points* $t_0 = 0 < t_1 < t_2 < \cdots < t_K = 1$ *such that for every* $k = 1, \cdots, K$, $\mu|_{[t_{k-1},t_k]}$ *is a minimal constant speed geodesic and*

$$W_2^2(\mu_{t'},\mu_{t''}) = |t''-t'|^2 \int |y-x|^2 \, d\gamma(x,y) \quad \textit{for every } t',t'' \in [t_{k-1},t_k]. \tag{6.5.3}$$

3. *The length of the curve* $(\mu_t)_{t\in[0,1]}$ *coincides with* $\left(\int |y-x|^2 \, d\gamma\right)^{1/2}$.

Proof. The first statement follows by Lemma 6.5.1, since every measure μ_s has finite support and for every $t \in [0,1]$

$$\begin{aligned} \sup\left\{|y-x| : (x,y) \in \operatorname{supp} \gamma_{s,t}\right\} &= |t-s| \sup\left\{|y-x| : (x,y) \in \operatorname{supp} \gamma\right\} \\ &\leqslant |t-s| \max\{|y-x| : x \in \operatorname{supp} \mu_0,\ y \in \operatorname{supp} \mu_1\}. \end{aligned}$$

In order to prove the second claim, we define an increasing sequence $(t_n)_{n=0}^\infty \subset [0,1]$ by induction as follows:

- $t_0 := 0$;
- if $t_n < 1$ then $t_{n+1} := \sup\left\{t \in (t_n,1] : W_2^2(\mu_{t_n},\mu_t) = |t-t_n|^2 \int |y-x|^2 \, d\gamma\right\}$;
- if $t_n = 1$ then $t_{n+1} = 1$.

The sequence is well defined thanks to the first claim. It is easy to see that there exists $K \in \mathbb{N}$ such that $t_K = 1$. If not, t_n would be strictly increasing with limit $t_\infty \leqslant 1$ as $n \to \infty$. By the first claim, there exists $r > 0$ such that the restriction of μ to $[t_\infty - r, t_\infty]$ is a minimal geodesic, so that whenever $t_n \geqslant t_\infty - r$ we should get $t_{n+1} = t_\infty$, a contradiction.

Claim (3) follows immediately by (2). □

7

DISSIPATIVITY IN WASSERSTEIN SPACES

In this chapter we present the notion of dissipativity in $\mathcal{P}_2(\mathbb{H})$ and we study dissipative operators on the Wasserstein space. In particular Section 7.1 introduces the pseudo scalar products/duality pairings between probability measures; Section 7.2 presents a few results concerning the differentiability of the Wasserstein distance along absolutely continuous curves; in Sections 7.3 and 7.4 we study the properties of duality pairings and how they interact with geodesics in $\mathcal{P}_2(\mathbb{H})$; in Sections 7.5, 7.6 and 7.7 we introduce the notion of dissipative operator on $\mathcal{P}_2(\mathbb{H})$, called Multivalued Probability Vector Field, we study its behaviour along geodesics and notions of extensions; finally Section 7.8 deals with a refined notion of dissipativity related to discrete measures.
In this whole Chapter, $\mathbb{H}$ is a fixed, possibly infinite dimensional, separable Hilbert space with $\dim(\mathbb{H}) \geqslant 2$.

This Chapter is the result of a collaboration with Giulia Cavagnari and Giuseppe Savaré and, except for Section 7.8, it appeared in [34].

7.1 DIRECTIONAL DERIVATIVES OF THE WASSERSTEIN DISTANCE AND DUALITY PAIRINGS

We start from a concavity property of

$$f(s,t) := \frac{1}{2} W_2^2(\mathsf{exp}^s_\sharp \Phi_0, \mathsf{exp}^t_\sharp \Phi_1), \quad s,t \in \mathbb{R}, \tag{7.1.1}$$

with $\Phi_0, \Phi_1 \in \mathcal{P}_2(\mathsf{T}\mathbb{H})$ and $\mathsf{exp}^t : \mathsf{T}\mathbb{H} \to \mathbb{H}$ is defined as

$$\mathsf{exp}^t(x,v) := x + tv, \quad (x,v) \in \mathsf{T}\mathbb{H}. \tag{7.1.2}$$

We also use the notation

$$|\Phi|_2 := \int_{\mathsf{T}\mathbb{H}} |v|^2 \, d\Phi(x,v) \quad \Phi \in \mathcal{P}(\mathsf{T}\mathbb{H}). \tag{7.1.3}$$

Moreover, recalling Theorem 2.1.1 and Remark 2.1.2, we give the following definition.

Definition 7.1.1. Given $\Phi \in \mathcal{P}_2(\mathsf{T}\mathbb{H}|\mu)$, the *barycenter of* Φ is the function $\mathbf{b}_\Phi \in L^2_\mu(\mathbb{H};\mathbb{H})$ defined by

$$\mathbf{b}_\Phi(x) := \int_{\mathbb{H}} v \, d\Phi_x(v) \quad \text{for } \mu\text{-a.e. } x \in \mathbb{H},$$

where $\{\Phi_x\}_{x\in\mathbb{H}} \subset \mathcal{P}_2(\mathbb{H})$ is the disintegration of Φ w.r.t. μ.

The proof of the following result comes from [5, Proposition 7.3.1].

Lemma 7.1.2. *Let $\Phi_0, \Phi_1 \in \mathcal{P}_2(\mathbb{TH})$, $s, t \in \mathbb{R}$, and let $\vartheta^{s,t} \in \Gamma(\exp^s_\sharp \Phi_0, \exp^t_\sharp \Phi_1)$. Then there exists $\boldsymbol{\Theta}^{s,t} \in \Gamma(\Phi_0, \Phi_1)$ such that $(\exp^s, \exp^t)_\sharp \boldsymbol{\Theta}^{s,t} = \vartheta^{s,t}$.*

Proof. Define, for every $r, s, t \in \mathbb{R}$,

$$\Sigma^r : \mathbb{TH} \to \mathbb{TH}, \quad \Sigma^r(x, v) := (\exp^r(x, v), v),$$
$$\Lambda^{s,t} : \mathbb{TH} \times \mathbb{TH} \to \mathbb{TH} \times \mathbb{TH}, \quad \Lambda^{s,t} := (\Sigma^s, \Sigma^t).$$

Consider the probabilities $(\Sigma^s)_\sharp \Phi_0, (\Sigma^t)_\sharp \Phi_1$ and $\vartheta^{s,t}$. They are constructed in such a way that there exists $\boldsymbol{\Psi}^{s,t} \in \mathcal{P}(\mathbb{TH} \times \mathbb{TH})$ s.t.

$$(x^0, v^0)_\sharp \boldsymbol{\Psi}^{s,t} = (\Sigma^s)_\sharp \Phi_0, \quad (x^1, v^1)_\sharp \boldsymbol{\Psi}^{s,t} = (\Sigma^t)_\sharp \Phi_1, \quad (x^0, x^1)_\sharp \boldsymbol{\Psi}^{s,t} = \vartheta^{s,t},$$

where we adopted the notation $x^i(x_0, v_0, x_1, v_1) := x_i$ and $v^i(x_0, v_0, x_1, v_1) := v_i$, $i = 0, 1$. We conclude by taking $\boldsymbol{\Theta}^{s,t} := (\Lambda^{-s,-t})_\sharp \boldsymbol{\Psi}^{s,t}$. □

Proposition 7.1.3. *Let $\Phi_0, \Phi_1 \in \mathcal{P}_2(\mathbb{TH})$ with $\mu_1 = x_\sharp \Phi_1$ and $\varphi^2 := |\Phi_0|_2^2 + |\Phi_1|_2^2$, let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function defined by (7.1.1) and let $h, g : \mathbb{R} \to \mathbb{R}$ be defined by*

$$\begin{aligned} h(s) &:= f(s, s) = \frac{1}{2} W_2^2(\exp^s_\sharp \Phi_0, \exp^s_\sharp \Phi_1), \\ g(s) &:= f(s, 0) = \frac{1}{2} W_2^2(\exp^s_\sharp \Phi_0, \mu_1). \end{aligned} \tag{7.1.4}$$

1. *The function $(s, t) \mapsto f(s, t) - \frac{1}{2}\varphi^2(s^2 + t^2)$ is concave, i.e. it holds*

$$\begin{aligned} f((1-\alpha)s_0 + \alpha s_1, (1-\alpha)t_0 + \alpha t_1) \geqslant (1-\alpha) f(s_0, t_0) + \alpha f(s_1, t_1) \\ - \frac{1}{2}\alpha(1-\alpha)\Big[(s_1 - s_0)^2 + (t_1 - t_0)^2\Big]\varphi^2 \end{aligned} \tag{7.1.5}$$

   for every $s_0, s_1, t_0, t_1 \in \mathbb{R}$ and every $\alpha \in [0, 1]$.

2. *The function $s \mapsto h(s) - \varphi^2 s^2$ is concave.*

3. *the function $s \mapsto g(s) - \frac{1}{2} s^2 |\Phi_0|_2^2$ is concave.*

Proof. Let us first prove (7.1.5). We set $s := (1-\alpha)s_0 + \alpha s_1$, $t := (1-\alpha)t_0 + \alpha t_1$ and we apply Lemma 7.1.2 to find $\boldsymbol{\Theta} \in \Gamma(\Phi_0, \Phi_1)$ such that $(\exp^s, \exp^t)_\sharp \boldsymbol{\Theta} \in \Gamma_o(\exp^s_\sharp \Phi_0, \exp^t_\sharp \Phi_1)$. Then, recalling the Hilbertian identity

$$|(1-\alpha)a + \alpha b|^2 = (1-\alpha)|a|^2 + \alpha|b|^2 - \alpha(1-\alpha)|a - b|^2, \quad a, b \in \mathbb{H},$$

we have

$$
\begin{aligned}
&W_2^2(\exp_\sharp^s \Phi_0, \exp_\sharp^t \Phi_1) \\
&= \int |x_0 + sv_0 - (x_1 + tv_1)|^2 \, d\boldsymbol{\Theta} \\
&= \int |(1-\alpha)(x_0 + s_0 v_0) + \alpha(x_0 + s_1 v_0) - \\
&\qquad (1-\alpha)(x_1 + t_0 v_1) - \alpha(x_1 + t_1 v_1)|^2 \, d\boldsymbol{\Theta} \\
&= (1-\alpha) \int |x_0 + s_0 v_0 - (x_1 + t_0 v_1)|^2 \, d\boldsymbol{\Theta} \\
&\quad + \alpha \int |x_0 + s_1 v_0 - (x_1 + t_1 v_1)|^2 \, d\boldsymbol{\Theta} \\
&\quad - \alpha(1-\alpha) \int |(s_1 - s_0) v_0 + (t_1 - t_0) v_1|^2 \, d\boldsymbol{\Theta} \\
&\geqslant (1-\alpha) W_2^2(\exp_\sharp^{s_0} \Phi_0, \exp_\sharp^{t_0} \Phi_1) + \alpha W_2^2(\exp_\sharp^{s_1} \Phi_0, \exp_\sharp^{t_1} \Phi_1) \\
&\quad - \alpha(1-\alpha)\Big((s_1 - s_0)^2 + (t_1 - t_0)^2\Big)\Big(\int |v_0|^2 \, d\Phi_0 + \int |v_1|^2 \, d\Phi_1\Big).
\end{aligned}
$$

which is the thesis. Claims (2) and (3) follow as particular cases when $t = s$ or $t = 0$. □

Semi-concavity guarantees the existence of right and left derivatives at $(0,0)$: given $\alpha, \beta \in \mathbb{R}$, we have (see e.g. [63, Ch. VI, Prop. 1.1.2]) that

$$
\begin{aligned}
f'_r(\alpha, \beta) &= \sup_{\rho > 0} \frac{f(\alpha\rho, \beta\rho) - f(0,0)}{\rho} - \frac{\rho\varphi^2}{2}(\alpha^2 + \beta^2), \\
f'_l(\alpha, \beta) &= \inf_{\rho > 0} \frac{f(0,0) - f(-\alpha\rho, -\beta\rho)}{\rho} + \frac{\rho\varphi^2}{2}(\alpha^2 + \beta^2).
\end{aligned}
$$

f'_r (resp. f'_l) is a concave (resp. convex) and positively 1-homogeneous function. We have

$$f'_r(-\alpha, -\beta) = -f'_l(\alpha, \beta) \quad \text{for every } \alpha, \beta \in \mathbb{R}, \tag{7.1.6}$$

$$f'_l(\alpha, \beta) \geqslant f'_r(\alpha, \beta) \quad \text{for every } \alpha, \beta \in \mathbb{R}, \tag{7.1.7}$$

$$f'_r(\alpha, \beta) \geqslant \alpha f'_r(1,0) + \beta f'_r(0,1) \quad \text{for every } \alpha, \beta \geqslant 0, \tag{7.1.8}$$

$$f(s,t) \leqslant f(0,0) + f'_r(s,t) - \frac{\varphi^2}{2}(s^2 + t^2) \quad \text{for every } s, t \in \mathbb{R}.$$

In addition

$$f'_r(1,0) = g'_r(0) = \lim_{\rho \downarrow 0} \frac{g(\rho) - g(0)}{\rho}$$

where g is as in (7.1.4); an analogous result holds for $f'_l(1,0)$. We will use the following notation for f'_r, f'_l, g'_r and g'_l, setting also

$$\mathcal{P}_2(\mathbb{TH}|\mu) := \left\{\Phi \in \mathcal{P}_2(\mathbb{TH}) \mid x_\sharp \Phi = \mu\right\}, \quad \mu \in \mathcal{P}_2(\mathbb{H}). \tag{7.1.9}$$

Definition 7.1.4. Let $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{H})$, $\Phi_0 \in \mathcal{P}_2(\mathbb{TH}|\mu_0)$ and $\Phi_1 \in \mathcal{P}_2(\mathbb{TH}|\mu_1)$. We define

$$[\Phi_0, \mu_1]_r := \lim_{s\downarrow 0} \frac{W_2^2(\exp_\sharp^s \Phi_0, \mu_1) - W_2^2(\mu_0, \mu_1)}{2s},$$
$$[\Phi_0, \mu_1]_l := \lim_{s\downarrow 0} \frac{W_2^2(\mu_0, \mu_1) - W_2^2(\exp_\sharp^{-s} \Phi_0, \mu_1)}{2s},$$

and analogously

$$[\Phi_0, \Phi_1]_r := \lim_{t\downarrow 0} \frac{W_2^2(\exp_\sharp^t \Phi_0, \exp_\sharp^t \Phi_1) - W_2^2(\mu_0, \mu_1)}{2t},$$
$$[\Phi_0, \Phi_1]_l := \lim_{t\downarrow 0} \frac{W_2^2(\mu_0, \mu_1) - W_2^2(\exp_\sharp^{-t} \Phi_0, \exp_\sharp^{-t} \Phi_1)}{2t}.$$

Recalling the definitions of f and g given by (7.1.1) and (7.1.4), with Φ_0 and Φ_1 as above, we notice that

$$[\Phi_0, \mu_1]_r = g_r'(0) = f_r'(1,0),$$
$$[\Phi_0, \mu_1]_l = g_l'(0) = f_l'(1,0),$$
$$[\Phi_0, \Phi_1]_r = f_r'(1,1),$$
$$[\Phi_0, \Phi_1]_l = f_l'(1,1).$$

Remark 7.1.5. Notice that $[\Phi_0, \mu_1]_r = [\Phi_0, \Phi_{\mu_1}]_r$ and $[\Phi_0, \mu_1]_l = [\Phi_0, \Phi_{\mu_1}]_l$, where

$$\Phi_{\mu_1} = (\boldsymbol{i}_{\mathbb{H}}, 0)_\sharp \mu_1 \in \mathcal{P}_2(\mathbb{TH}).$$

Moreover, given $\Phi \in \mathcal{P}(\mathbb{TH})$ and using the notation

$$-\Phi := J_\sharp \Phi, \quad \text{with } J(x,v) := (x, -v), \tag{7.1.10}$$

we have

$$[-\Phi_0, -\Phi_1]_r = -[\Phi_0, \Phi_1]_l, \quad \text{and} \quad [-\Phi_0, \mu_1]_r = -[\Phi_0, \mu_1]_l.$$

In particular, the properties of $[\cdot,\cdot]_l$ (in $\mathcal{P}_2(\mathbb{TH}) \times \mathcal{P}_2(\mathbb{TH})$ or $\mathcal{P}_2(\mathbb{TH}) \times \mathcal{P}_2(\mathbb{H})$) and the ones of $[\cdot,\cdot]_r$ in $\mathcal{P}_2(\mathbb{TH}) \times \mathcal{P}_2(\mathbb{H})$ can be easily derived by the corresponding ones of $[\cdot,\cdot]_r$ in $\mathcal{P}_2(\mathbb{TH}) \times \mathcal{P}_2(\mathbb{TH})$.

Recalling (7.1.8) and (7.1.6) we obtain the following result.

Corollary 7.1.6. *For every $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{H})$ and for every $\Phi_0 \in \mathcal{P}_2(\mathbb{TH}|\mu_0)$, $\Phi_1 \in \mathcal{P}_2(\mathbb{TH}|\mu_1)$, it holds*

$$[\Phi_0, \mu_1]_r + [\Phi_1, \mu_0]_r \leqslant [\Phi_0, \Phi_1]_r \quad \textit{and} \quad [\Phi_0, \mu_1]_l + [\Phi_1, \mu_0]_l \geqslant [\Phi_0, \Phi_1]_l.$$

Our aim is now to show a useful characterization of the above pairings. We denote by $x^0, v^0, x^1 : \mathbb{TH} \times \mathbb{H} \to \mathbb{H}$ the projection maps of a point (x_0, v_0, x_1) in $\mathbb{TH} \times \mathbb{H}$ (and similarly for $\mathbb{TH} \times \mathbb{TH}$ with x^0, v^0, x^1, v^1).

Let us introduce the following sets.

Definition 7.1.7. For every $\Phi_0 \in \mathcal{P}(\mathbb{TH})$ with $\mu_0 = x_\sharp \Phi_0$ and $\mu_1 \in \mathcal{P}_2(\mathbb{H})$ we set

$$\Lambda(\Phi_0, \mu_1) := \left\{ \boldsymbol{\sigma} \in \Gamma(\Phi_0, \mu_1) \mid (x^0, x^1)_\sharp \boldsymbol{\sigma} \in \Gamma_o(\mu_0, \mu_1) \right\}.$$

Analogously, for every $\Phi_0, \Phi_1 \in \mathcal{P}(\mathbb{TH})$ with $\mu_0 = x_\sharp \Phi_0$ and $\mu_1 = x_\sharp \Phi_1$ in $\mathcal{P}_2(\mathbb{H})$ we set

$$\Lambda(\Phi_0, \Phi_1) := \left\{ \boldsymbol{\Theta} \in \Gamma(\Phi_0, \Phi_1) \mid (x^0, x^1)_\sharp \boldsymbol{\Theta} \in \Gamma_o(\mu_0, \mu_1) \right\}.$$

Similar results to what follows (with analogous proofs) can be found also in [53, Theorem 4.2] and [44, Corollary 3.18] where $\mathbb{H}$ is a smooth compact Riemannian manifold.

Theorem 7.1.8. *For every $\Phi_0, \Phi_1 \in \mathcal{P}_2(\mathbb{TH})$ and $\mu_1 \in \mathcal{P}_2(\mathbb{H})$ we have*

$$[\Phi_0, \mu_1]_r = \min \left\{ \int_{\mathbb{TH} \times \mathbb{H}} \langle x_0 - x_1, v_0 \rangle \, d\boldsymbol{\sigma} \mid \boldsymbol{\sigma} \in \Lambda(\Phi_0, \mu_1) \right\}, \tag{7.1.11}$$

$$[\Phi_0, \Phi_1]_r = \min \left\{ \int_{\mathbb{TH} \times \mathbb{TH}} \langle x_0 - x_1, v_0 - v_1 \rangle \, d\boldsymbol{\Theta} \mid \boldsymbol{\Theta} \in \Lambda(\Phi_0, \Phi_1) \right\}. \tag{7.1.12}$$

We denote by $\Lambda_o(\Phi_0, \mu_1)$ (resp. $\Lambda_o(\Phi_0, \Phi_1)$) the subset of $\Lambda(\Phi_0, \mu_1)$ (resp. $\Lambda(\Phi_0, \Phi_1)$) where the minimum in (7.1.11) *(resp.* (7.1.12)*) is attained.*

Proof. First, we recall that the minima in the right hand side are attained since $\Lambda(\Phi_0, \mu_1)$ and $\Lambda(\Phi_0, \Phi_1)$ are compact subsets of $\mathcal{P}_2(\mathbb{TH} \times \mathbb{H})$ and $\mathcal{P}_2(\mathbb{TH} \times \mathbb{TH})$ respectively by Lemma 2.4.1 and the integrands are continuous functions with quadratic growth. Thanks to Remark 7.1.5, we only need to prove the equality (7.1.12). For every $\boldsymbol{\Theta} \in \Lambda(\Phi_0, \Phi_1)$ and setting $\mu_0 = x_\sharp \Phi_0$, $\mu_1 = x_\sharp \Phi_1$, we have

$$\begin{aligned}
&W_2^2(\mathsf{exp}_\sharp^s(\Phi_0), \mathsf{exp}_\sharp^s(\Phi_1)) \\
&\leqslant \int_{\mathbb{TH} \times \mathbb{TH}} |(x_0 - x_1) + s(v_0 - v_1)|^2 \, d\boldsymbol{\Theta} \\
&= \int_{\mathbb{H}^2} |x_0 - x_1|^2 \, d(x^0, x^1)_\sharp \boldsymbol{\Theta} + 2s \int_{\mathbb{TH} \times \mathbb{TH}} \langle x_0 - x_1, v_0 - v_1 \rangle \, d\boldsymbol{\Theta} \\
&\quad + s^2 \int_{\mathbb{H}^2} |v_0 - v_1|^2 \, d\boldsymbol{\Theta} \\
&= W_2^2(\mu_0, \mu_1) + 2s \int_{\mathbb{TH} \times \mathbb{TH}} \langle x_0 - x_1, v_0 - v_1 \rangle \, d\boldsymbol{\Theta} + s^2 \int_{\mathbb{H}^2} |v_0 - v_1|^2 \, d\boldsymbol{\Theta}
\end{aligned}$$

and this immediately implies

$$[\Phi_0, \Phi_1]_r \leqslant \min \left\{ \int_{\mathbb{TH} \times \mathbb{TH}} \langle x_0 - x_1, v_0 - v_1 \rangle \, d\boldsymbol{\Theta} \mid \boldsymbol{\Theta} \in \Lambda(\Phi_0, \Phi_1) \right\}.$$

In order to prove the converse inequality, thanks to Lemma 7.1.2, for every $s > 0$ we can find $\boldsymbol{\Theta}_s \in \Gamma(\Phi_0, \Phi_1)$ s.t.

$$(\mathsf{exp}^s, \mathsf{exp}^s)_\sharp \boldsymbol{\Theta}_s \in \Gamma_o(\mathsf{exp}_\sharp^s \Phi_0, \mathsf{exp}_\sharp^s \Phi_1).$$

Then

$$\frac{W_2^2(\exp_\sharp^s \Phi_0, \exp_\sharp^s \Phi_1) - W_2^2(\mu_0,\mu_1)}{2s} \geqslant \frac{1}{2s}\int_{\mathbb{TH}\times\mathbb{TH}} |(x_0-x_1)+s(v_0-v_1)|^2 \,d\boldsymbol{\Theta}_s - \frac{1}{2s}\int_{\mathbb{TH}\times\mathbb{TH}} |x_0-x_1|^2\,d\boldsymbol{\Theta}_s \geqslant \int_{\mathbb{TH}\times\mathbb{TH}} \langle x_0-x_1, v_0-v_1\rangle\,d\boldsymbol{\Theta}_s. \tag{7.1.13}$$

Since $\Gamma(\Phi_0,\Phi_1)$ is compact in $\mathcal{P}_2(\mathbb{TH}\times\mathbb{TH})$, there exists a vanishing sequence $k\mapsto s_k$ and $\boldsymbol{\Theta}\in\Gamma(\Phi_0,\Phi_1)$ s.t. $\boldsymbol{\Theta}_{s_k}\to\boldsymbol{\Theta}$ in $\mathcal{P}_2(\mathbb{TH}\times\mathbb{TH})$. Moreover it holds $(\exp^{s_k},\exp^{s_k})_\sharp\boldsymbol{\Theta}_{s_k}\to(x^0,x^1)_\sharp\boldsymbol{\Theta}$ in $\mathcal{P}(\mathbb{TH}\times\mathbb{TH})$ so that $(x^0,x^1)_\sharp\boldsymbol{\Theta}\in\Gamma_o(\mu_0,\mu_1)$, and therefore $\boldsymbol{\Theta}\in\Lambda(\Phi_0,\Phi_1)$. The convergence in $\mathcal{P}_2(\mathbb{TH}\times\mathbb{TH})$ yields

$$\lim_k \int_{\mathbb{TH}\times\mathbb{TH}} \langle x_0-x_1, v_0-v_1\rangle\,d\boldsymbol{\Theta}_{s_k} = \int_{\mathbb{TH}\times\mathbb{TH}} \langle x_0-x_1, v_0-v_1\rangle\,d\boldsymbol{\Theta},$$

so that, passing to the limit in (7.1.13) along the sequence s_k, we obtain

$$[\Phi_0,\Phi_1]_r \geqslant \int_{\mathbb{TH}\times\mathbb{TH}} \langle x_0-x_1, v_0-v_1\rangle\,d\boldsymbol{\Theta}$$

for some $\boldsymbol{\Theta}\in\Lambda(\Phi_0,\Phi_1)$. □

Corollary 7.1.9. *Let $\Phi_0,\Phi_1\in\mathcal{P}_2(\mathbb{TH})$ and $\mu_1\in\mathcal{P}_2(\mathbb{H})$, then*

$$[\Phi,\mu_1]_l = \max\left\{\int_{\mathbb{TH}\times\mathbb{H}} \langle x_0-x_1, v_0\rangle\,d\boldsymbol{\sigma} \mid \boldsymbol{\sigma}\in\Lambda(\Phi_0,\mu_1)\right\}, \tag{7.1.14}$$

$$[\Phi_0,\Phi_1]_l = \max\left\{\int_{\mathbb{TH}\times\mathbb{TH}} \langle x_0-x_1, v_0-v_1\rangle\,d\boldsymbol{\Theta} \mid \boldsymbol{\Theta}\in\Lambda(\Phi_0,\Phi_1)\right\}.$$

7.2 DIRECTIONAL DERIVATIVES OF THE WASSERSTEIN DISTANCE ALONG A.C. CURVES

We present now the differentiability properties of the map $\mathcal{I}\ni t\mapsto \frac{1}{2}W_2^2(\mu_t,\nu)$ along a locally absolutely continuous curve $\mu:\mathcal{I}\to\mathcal{P}_2(\mathbb{H})$, with $\mathcal{I}$ an open interval of $\mathbb{R}$ and $\nu\in\mathcal{P}_2(\mathbb{H})$.

Theorem 7.2.1. *Let $\mu:\mathcal{I}\to\mathcal{P}_2(\mathbb{H})$ be a locally absolutely continuous curve and let $\boldsymbol{v}:\mathcal{I}\times\mathbb{H}\to\mathbb{H}$ and $A(\mu)$ be as in Theorem 2.4.6. Then, for every $\nu\in\mathcal{P}_2(\mathbb{H})$ and every $t\in A(\mu)$, it holds*

$$\lim_{h\downarrow 0}\frac{W_2^2(\mu_{t+h},\nu)-W_2^2(\mu_t,\nu)}{2h} = \left[(\boldsymbol{i}_{\mathbb{H}},\boldsymbol{v}_t)_\sharp\mu_t,\nu\right]_r, \tag{7.2.1}$$

$$\lim_{h\uparrow 0}\frac{W_2^2(\mu_{t+h},\nu)-W_2^2(\mu_t,\nu)}{2h} = \left[(\boldsymbol{i}_{\mathbb{H}},\boldsymbol{v}_t)_\sharp\mu_t,\nu\right]_l,$$

so that the map $s\mapsto W_2^2(\mu_s,\nu)$ is left and right differentiable at every $t\in A(\mu)$. In particular,

1. *if* $t \in A(\mu)$ *and* $\nu \in \mathcal{P}_2(\mathbb{H})$ *are s.t. there exists a unique optimal transport plan between* μ_t *and* ν*, then the map* $s \mapsto W_2^2(\mu_s, \nu)$ *is differentiable at* t*;*

2. *there exists a subset* $A(\mu, \nu) \subset A(\mu)$ *of full Lebesgue measure such that* $s \mapsto W_2^2(\mu_s, \nu)$ *is differentiable in* $A(\mu, \nu)$ *and*

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}W_2^2(\mu_t, \nu) &= \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t, \nu\right]_r = \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t, \nu\right]_l \\ &= \int \langle \boldsymbol{v}_t(x_1), x_1 - x_2\rangle \, d\boldsymbol{\mu}(x_1, x_2)\end{aligned}$$

for every $\boldsymbol{\mu} \in \Gamma_o(\mu_t, \nu)$, $t \in A(\mu, \nu)$.

Proof. Let $\nu \in \mathcal{P}_2(\mathbb{H})$ and for every $t \in \mathcal{I}$ we set $\Phi_t := (\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t \in \mathcal{P}_2(\mathsf{T}\mathbb{H})$. By Theorem 7.1.8, we have

$$\begin{aligned}\lim_{h\downarrow 0} \frac{W_2^2(\exp_\sharp^h \Phi_t, \nu) - W_2^2(\mu_t, \nu)}{2h} &= \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t, \nu\right]_r, \\ \lim_{h\uparrow 0} \frac{W_2^2(\exp_\sharp^h \Phi_t, \nu) - W_2^2(\mu_t, \nu)}{2h} &= \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t, \nu\right]_l.\end{aligned}$$

Since $\exp_\sharp^h \Phi_t = (\boldsymbol{i}_{\mathbb{H}} + h\boldsymbol{v}_t)_\sharp \mu_t$, then thanks to Theorem 2.4.6 we have that the above limits coincide respectively with the limits in the statement, for all $t \in A(\mu)$.

Claim (1) comes by the characterizations given in Theorem 7.1.8 and Corollary 7.1.9. Indeed, if there exists a unique optimal transport plan between μ_t and ν, then $\left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t, \nu\right]_r = \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t, \nu\right]_l$.

Claim (2) is a simple consequence of the fact that $s \mapsto W_2^2(\mu_s, \nu)$ is differentiable a.e. in $\mathcal{I}$. □

Remark 7.2.2. In Theorem 7.2.1 we can actually replace $\boldsymbol{v}$ with any Borel velocity field $\boldsymbol{w}$ solving the continuity equation for μ and s.t. $\|\boldsymbol{w}_t\|_{L^2_{\mu_t}} \in L^1_{loc}(\mathcal{I})$. Indeed, we notice that by [5, Lemma 5.3.2],

$$\begin{aligned}\Lambda((\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t, \nu) &= \{(x^0, \boldsymbol{v}_t \circ x^0, x^1)_\sharp \boldsymbol{\gamma} \mid \boldsymbol{\gamma} \in \Gamma_o(\mu_t, \nu)\}, \\ \Lambda((\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{w}_t)_\sharp \mu_t, \nu) &= \{(x^0, \boldsymbol{w}_t \circ x^0, x^1)_\sharp \boldsymbol{\gamma} \mid \boldsymbol{\gamma} \in \Gamma_o(\mu_t, \nu)\},\end{aligned}$$

so that, by [5, Proposition 8.5.4], we get

$$\begin{aligned}\left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t, \nu\right]_r &= \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{w}_t)_\sharp \mu_t, \nu\right]_r, \\ \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t, \nu\right]_l &= \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{w}_t)_\sharp \mu_t, \nu\right]_l.\end{aligned}$$

Theorem 7.2.3. *Let* $\mu^1, \mu^2 : \mathcal{I} \to \mathcal{P}_2(\mathbb{H})$ *be locally absolutely continuous curves and let* $\boldsymbol{v}^1, \boldsymbol{v}^2 : \mathcal{I} \times \mathbb{H} \to \mathbb{H}$ *be the corresponding Wasserstein velocity fields satisfying* (2.4.8) *in* $A(\mu^1)$ *and* $A(\mu^2)$ *respectively. Then, for every* $t \in A(\mu^1) \cap A(\mu^2)$*, it holds*

$$\begin{aligned}\lim_{h\downarrow 0} \frac{W_2^2(\mu^1_{t+h}, \mu^2_{t+h}) - W_2^2(\mu^1_t, \mu^2_t)}{2h} &= \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}^1_t)_\sharp \mu^1_t, (\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}^2_t)_\sharp \mu^2_t\right]_r, \\ \lim_{h\uparrow 0} \frac{W_2^2(\mu^1_{t+h}, \mu^2_{t+h}) - W_2^2(\mu^1_t, \mu^2_t)}{2h} &= \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}^1_t)_\sharp \mu^1_t, (\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}^2_t)_\sharp \mu^2_t\right]_l.\end{aligned}$$

In particular, there exists a subset $A \subset A(\mu^1) \cap A(\mu^2)$ *of full Lebesgue measure such that* $s \mapsto W_2^2(\mu_s^1, \mu_s^2)$ *is differentiable in* A *and*

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}W_2^2(\mu_t^1,\mu_t^2) &= \left[(\boldsymbol{i}_{\mathbb{H}},\boldsymbol{v}_t^1)_\sharp\mu_t^1,(\boldsymbol{i}_{\mathbb{H}},\boldsymbol{v}_t^2)_\sharp\mu_t^2\right]_r = \left[(\boldsymbol{i}_{\mathbb{H}},\boldsymbol{v}_t^1)_\sharp\mu_t^1,(\boldsymbol{i}_{\mathbb{H}},\boldsymbol{v}_t^2)_\sharp\mu_t^2\right]_l \\ &= \int \langle \boldsymbol{v}_t^1-\boldsymbol{v}_t^2, x_1-x_2\rangle \, d\boldsymbol{\mu}(x_1,x_2)\end{aligned} \tag{7.2.2}$$

for every $\boldsymbol{\mu} \in \Gamma_o(\mu_t^1, \mu_t^2)$, $t \in A$.

The proof of Theorem 7.2.3 follows by the same argument of the proof of Theorem 7.2.1.

Remark 7.2.4. In general, if $\mu : \mathcal{I} \to \mathcal{P}_2(\mathbb{H})$ is a locally absolutely continuous curve and $\nu \in \mathcal{P}_2(\mathbb{H})$, then the map $\mathcal{I} \ni s \mapsto W_2^2(\mu_s, \nu)$ is locally absolutely continuous and thus differentiable in a set of full measure $A(\mu, \nu) \subset \mathcal{I}$ which, in principle, depends both on μ and ν. What Theorem 7.2.1 shows is that, independently of ν, there is a full measure set $A(\mu)$, depending only on μ, where this map is left and right differentiable. If moreover ν and $t \in A(\mu)$ are such that there is a unique optimal transport plan between them, we can actually conclude that the above map is differentiable at t. The next Example 7.2.5 shows how the result in Theorem 7.2.1 is optimal, proving the existence of a locally absolutely continuous curve $\mu : [0, +\infty) \to \mathcal{P}_2(\mathbb{R}^2)$ s.t. the full measure set of differentiability points of the map $[0, +\infty) \ni s \mapsto W_2^2(\mu_s, \nu)$ depends also on $\nu \in \mathcal{P}_2(\mathbb{R}^2)$.

Example 7.2.5. It is enough to show that

$$\begin{aligned}&\text{for every } t_0 \in A(\mu) \text{ there exist } \nu_0 \in \mathcal{P}_2(\mathbb{R}^2)\\ &\text{and } \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \in \Gamma_o(\mu_{t_0}, \nu_0) \text{ s.t. } L(\boldsymbol{\gamma}_1) \neq L(\boldsymbol{\gamma}_2),\end{aligned}$$

where $A(\mu)$ is as in Theorem 2.4.6 and, for $\boldsymbol{\gamma} \in \mathcal{P}_2(\mathbb{R}^2 \times \mathbb{R}^2)$ s.t. $x^0_\sharp \boldsymbol{\gamma} = \mu_t$, we define

$$L(\boldsymbol{\gamma}) := \int_{\mathbb{H}^2} \langle \boldsymbol{v}_t(x), x - y\rangle \, d\boldsymbol{\gamma}(x, y).$$

Indeed this will imply that $\left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_{t_0})_\sharp \mu_{t_0}, \nu_0\right]_r \neq \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_{t_0})_\sharp \mu_{t_0}, \nu_0\right]_l$, hence the non differentiability at t_0.

Let us consider two regular functions $u : [0, +\infty) \to \mathbb{R}^2$ and $r : [0, +\infty) \to \mathbb{R}$ s.t. $|u_t| = 1$ for every $t \geqslant 0$. Let $\omega : [0, +\infty) \to \mathbb{R}^2$ be defined as the orthogonal direction to u_t:

$$\omega_t := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} u_t, \qquad t \geqslant 0.$$

Being the norm of u constant in time, there exists some regular $\lambda : (0, +\infty) \to \mathbb{R}$ s.t. $\dot{u}_t = \lambda_t \omega_t$ for every $t > 0$. Finally we define

$$\begin{aligned} x_1 &: [0, +\infty) \to \mathbb{R}^2, && x_1(t) := r_t u_t, \\ x_2 &: [0, +\infty) \to \mathbb{R}^2, && x_2(t) := -r_t u_t, \\ \mu &: [0, +\infty) \to \mathcal{P}_2(\mathbb{R}^2), && \mu_t := \frac{1}{2}\left(\delta_{x_1(t)} + \delta_{x_2(t)}\right). \end{aligned}$$

Observe that $\dot{x}_1(t) = \dot{r}_t u_t + r_t \dot{u}_t = -\dot{x}_2(t)$ for every $t > 0$. Moreover, for every $\varphi \in C_c^\infty(\mathbb{R}^2)$ and $t > 0$, we have

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^2} \varphi \, d\mu_t &= \frac{d}{dt}\left(\frac{1}{2}\varphi(x_1(t)) + \frac{1}{2}\varphi(x_2(t))\right)\\ &= \frac{1}{2}\nabla\varphi(x_1(t))\,\dot{x}_1(t) + \frac{1}{2}\nabla\varphi(x_2(t))\,\dot{x}_2(t)\\ &= \int_{\mathbb{R}^2}\langle v_t(x), \nabla\varphi(x)\rangle \, d\mu_t,\end{aligned}$$

where

$$\boldsymbol{v}_t(x) := \begin{cases}\dot{x}_1(t) & \text{if } x = x_1(t),\\ \dot{x}_2(t) & \text{if } x = x_2(t),\end{cases} \qquad t > 0.$$

Hence, the above defined vector field $\boldsymbol{v}_t$ solves the continuity equation with μ_t. Let $t_0 \in A(\mu)$ and let us define $\omega_0 := \omega(t_0)$, $\nu_0 := \frac{1}{2}\delta_{\omega_0} + \frac{1}{2}\delta_{-\omega_0}$ and the plans $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \in \Gamma_o(\mu_{t_0}, \nu_0)$ by

$$\begin{aligned}\boldsymbol{\gamma}_1 &:= \frac{1}{2}\delta_{x_1(t_0)} \otimes \delta_{\omega_0} + \frac{1}{2}\delta_{x_2(t_0)} \otimes \delta_{-\omega_0},\\ \boldsymbol{\gamma}_2 &:= \frac{1}{2}\delta_{x_2(t_0)} \otimes \delta_{\omega_0} + \frac{1}{2}\delta_{x_1(t_0)} \otimes \delta_{-\omega_0}.\end{aligned}$$

Notice that they are optimal since any plan in $\Gamma(\mu_{t_0}, \nu_0)$ has the same cost, being the points $\omega_0, x_1(t_0), x_2(t_0), -\omega_0$ the vertexes of a rhombus. Finally, we compute $L(\boldsymbol{\gamma}_1)$ and $L(\boldsymbol{\gamma}_2)$:

$$\begin{aligned}L(\boldsymbol{\gamma}_1) &= \int_{\mathbb{R}^2\times\mathbb{R}^2}\langle x - y, \boldsymbol{v}_t(x)\rangle \, d\boldsymbol{\gamma}_1(x,y)\\ &= \frac{1}{2}\langle \dot{x}_1(t_0), x_1(t_0) - \omega_0\rangle + \frac{1}{2}\langle \dot{x}_2(t_0), x_2(t_0) + \omega_0\rangle\\ &= \langle \dot{x}_1(t_0), x_1(t_0) - \omega_0\rangle = \langle \dot{r}_{t_0}u_{t_0} + r_{t_0}\dot{u}_{t_0}, r_{t_0}u_{t_0} - \omega_0\rangle\\ &= r_{t_0}\dot{r}_{t_0} - r_{t_0}\lambda_{t_0},\\ L(\boldsymbol{\gamma}_2) &= \int_{\mathbb{R}^2\times\mathbb{R}^2}\langle x - y, \boldsymbol{v}_t(x)\rangle \, d\boldsymbol{\gamma}_2(x,y)\\ &= \frac{1}{2}\langle \dot{x}_2(t_0), x_2(t_0) - \omega_0\rangle + \frac{1}{2}\langle \dot{x}_1(t_0), x_1(t_0) + \omega_0\rangle\\ &= \langle \dot{x}_1(t_0), x_1(t_0) + \omega_0\rangle\\ &= \langle \dot{r}_{t_0}u_{t_0} + r_{t_0}\dot{u}_{t_0}, r_{t_0}u_{t_0} + \omega_0\rangle\\ &= r_{t_0}\dot{r}_{t_0} + r_{t_0}\lambda_{t_0}.\end{aligned}$$

In this way, if $r_{t_0} \neq 0$ and $\lambda_{t_0} \neq 0$ we have $L(\boldsymbol{\gamma}_1) \neq L(\boldsymbol{\gamma}_2)$. A possible choice for u and r satisfying the assumptions is

$$u_t := (\cos(t), \sin(t)), \qquad r_t = 1, \qquad t \geqslant 0,$$

so that $\lambda_t = 1$ for every $t > 0$.

We conclude this section with the following property of the upper derivative of a distance (we state it for the Wasserstein distance, but a general distance could be considered).

Lemma 7.2.6. *Let* $\mu : \mathcal{I} \to \mathcal{P}_2(X)$, $\nu \in \mathcal{P}_2(X)$, $t \in \mathcal{I}$, $\boldsymbol{\sigma}_t \in \Gamma_o(\mu_t, \nu)$, *and consider the constant speed geodesic* $\nu^t : [0,1] \to \mathcal{P}_2(X)$ *defined by* $\nu^t_s := (x^s)_\sharp \boldsymbol{\sigma}_t$ *for every* $s \in [0,1]$. *The upper right and left Dini derivatives* $b^\pm : (0,1] \to \mathbb{R}$ *defined by*

$$b^+(s) := \frac{1}{2s} \limsup_{h \downarrow 0} \frac{W_2^2(\mu_{t+h}, \nu^t_s) - W_2^2(\mu_t, \nu^t_s)}{h},$$
$$b^-(s) := \frac{1}{2s} \limsup_{h \downarrow 0} \frac{W_2^2(\mu_t, \nu^t_s) - W_2^2(\mu_{t-h}, \nu^t_s)}{h}$$

are respectively decreasing and increasing in $(0,1]$.

Proof. Take $0 \leqslant s' < s \leqslant 1$. Since $\nu^t : [0,1] \to \mathcal{P}_2(X)$ is a constant speed geodesic from μ_t to ν, we have

$$W_2(\mu_t, \nu^t_s) = W_2(\mu_t, \nu^t_{s'}) + W_2(\nu^t_{s'}, \nu^t_s),$$

then, by triangular inequality

$$\begin{aligned} W_2(\mu_{t+h}, \nu^t_s) - W_2(\mu_t, \nu^t_s) &\leqslant W_2(\mu_{t+h}, \nu^t_{s'}) + W_2(\nu^t_{s'}, \nu^t_s) - W_2(\mu_t, \nu^t_s) \\ &= W_2(\mu_{t+h}, \nu^t_{s'}) - W_2(\mu_t, \nu^t_{s'}). \end{aligned}$$

Dividing by $h > 0$ and passing to the limit as $h \downarrow 0$ we obtain that the function $a : [0,1] \to \mathbb{R}$ defined by

$$a^+(s) := \limsup_{h \downarrow 0} \frac{W_2(\mu_{t+h}, \nu^t_s) - W_2(\mu_t, \nu^t_s)}{h}$$

is decreasing. It is then sufficient to observe that for $s > 0$

$$b^+(s) = a^+(s) \frac{W_2(\mu_t, \nu^t_s)}{s} = a^+(s)\, W_2(\mu_t, \nu).$$

The monotonicity property of b^- follows by the same argument. □

7.3 CONVEXITY AND SEMICONTINUITY OF DUALITY PAIRINGS

We present a few results about the semicontinuity and convexity properties of $[\cdot,\cdot]_r$ and $[\cdot,\cdot]_l$.

We use the notation for $\mu \in \mathcal{P}(\mathbb{H})$

$$m_2^2(\mu) := \int_{\mathbb{H}} |x|^2 \, d\mu(x), \quad \mu \in \mathcal{P}(\mathbb{H}). \tag{7.3.1}$$

Lemma 7.3.1. *Let* $(\Phi_n)_{n \in \mathbb{N}} \subset \mathcal{P}_2(\mathsf{T}\mathbb{H})$ *be converging to* Φ *in* $\mathcal{P}_2^{sw}(\mathsf{T}\mathbb{H})$, *and let* $(\nu_n)_{n \in \mathbb{N}} \subset \mathcal{P}_2(\mathbb{H})$ *be converging to* ν *in* $\mathcal{P}_2(\mathbb{H})$. *Then*

$$\liminf_n \, [\Phi_n, \nu_n]_r \geqslant [\Phi, \nu]_r \quad \textit{and} \quad \limsup_n \, [\Phi_n, \nu_n]_l \leqslant [\Phi, \nu]_l \,. \tag{7.3.2}$$

Finally, if $(\Phi^i_n)_{n \in \mathbb{N}}$, $i = 0, 1$, *are sequences converging to* Φ^i *in* $\mathcal{P}_2^{sw}(\mathsf{T}\mathbb{H})$ *then*

$$\liminf_{n \to \infty} \left[\Phi^0_n, \Phi^1_n\right]_r \geqslant \left[\Phi^0, \Phi^1\right]_r, \qquad \limsup_{n \to \infty} \left[\Phi^0_n, \Phi^1_n\right]_l \geqslant \left[\Phi^0, \Phi^1\right]_l \,. \tag{7.3.3}$$

Proof. We just consider the proof of the first inequality (7.3.2); the other statements follow by similar arguments and by Remark 7.1.5.

We can extract a subsequence of $(\Phi_n)_{n\in\mathbb{N}}$ (not relabeled) s.t. the lim inf is achieved as a limit. We have to prove that

$$\lim_n [\Phi_n, \nu_n]_r \geqslant [\Phi, \nu]_r \,. \tag{7.3.4}$$

For every $n \in \mathbb{N}$ take $\boldsymbol{\sigma}_n \in \Lambda_o(\Phi_n, \nu_n)$ and $\bar{\vartheta}_n = (x^0, x^1)_\sharp \boldsymbol{\sigma}_n$. Since the marginals of $\bar{\vartheta}_n$ are converging w.r.t. W_2, the family $(\bar{\vartheta}_n)_{n\in\mathbb{N}}$ is relatively compact in $\mathcal{P}_2(\mathbb{H}^2)$. Hence, $(\boldsymbol{\sigma}_n)_{n\in\mathbb{N}}$ is relatively compact in $\mathcal{P}_2^{sws}(\mathbb{TH} \times \mathbb{H})$ by Proposition 6.3.3, since the moments $\int |v_0|^2 \, d\boldsymbol{\sigma}_n(x_0, v_0, x_1) = |\Phi_n|_2^2$ are uniformly bounded by assumption. Thus, possibly passing to a further subsequence, we have that $(\boldsymbol{\sigma}_n)_{n\in\mathbb{N}}$ converges to some $\boldsymbol{\sigma}$ in $\mathcal{P}_2^{sws}(\mathbb{TH} \times \mathbb{H})$. In particular $\boldsymbol{\sigma} \in \Lambda(\Phi, \nu)$ since optimality of the $\mathbb{H}^2$ marginals is preserved by narrow convergence. Indeed, it sufficies to use [5, Proposition 7.1.3] noting that

$$\int |x_0 - x_1|^2 \, d\boldsymbol{\sigma}_n \leqslant 2\, m_2^2(x_\sharp \Phi_n) + 2\, m_2^2(\nu_n) \leqslant K,$$

for some $K \geqslant 0$.

The relation in (6.3.3) then yields

$$\lim_{n\to\infty} [\Phi_n, \nu_n]_r = \lim_{n\to\infty} \int \langle v_0, x_0 - x_1 \rangle \, d\boldsymbol{\sigma}_n = \int \langle v_0, x_0 - x_1 \rangle \, d\boldsymbol{\sigma}$$

which yields (7.3.4) since the r.h.s. is larger than $[\Phi, \nu]_r$ by Theorem 7.1.8. □

Remark 7.3.2. Notice that in the special case in which $\Lambda(\Phi, \nu)$ is a singleton, then the limit exists and it holds

$$\lim_{n\to\infty} [\Phi_n, \nu_n]_r = [\Phi, \nu]_r \,, \qquad \lim_{n\to\infty} [\Phi_n, \nu_n]_l = [\Phi, \nu]_l \,.$$

Lemma 7.3.3. *For every $\mu, \nu \in \mathcal{P}_2(\mathbb{H})$ the maps $\Phi \mapsto [\Phi, \nu]_r$ and $(\Phi, \Psi) \mapsto [\Phi, \Psi]_r$ (resp. $\Phi \mapsto [\Phi, \nu]_l$ and $(\Phi, \Psi) \mapsto [\Phi, \Psi]_l$) are convex (resp. concave) in $\mathcal{P}_2(\mathbb{TH}|\mu)$ and $\mathcal{P}_2(\mathbb{TH}|\mu) \times \mathcal{P}_2(\mathbb{TH}|\nu)$.*

Proof. We prove the convexity of $(\Phi, \Psi) \mapsto [\Phi, \Psi]_r$ in $\mathcal{P}_2(\mathbb{TH}|\mu) \times \mathcal{P}_2(\mathbb{TH}|\nu)$; the argument of the proofs of the other statements are completely analogous.

Let $\Phi_k \in \mathcal{P}_2(\mathbb{TH}|\mu)$, $\Psi_k \in \mathcal{P}_2(\mathbb{TH}|\nu)$, and let $\beta_k \geqslant 0$, with $\sum_k \beta_k = 1$, $k = 1, \cdots, K$. We set $\Phi = \sum_{k=1}^K \beta_k \Phi_k$, $\Psi = \sum_{k=1}^K \beta_k \Psi_k$. For every k let us select $\boldsymbol{\Theta}_k \in \Lambda(\Phi_k, \Psi_k)$ such that

$$[\Phi_k, \Psi_k]_r = \int \langle v_1 - v_0, x_1 - x_0 \rangle \, d\boldsymbol{\Theta}_k.$$

It is not difficult to check that $\boldsymbol{\Theta} := \sum_k \beta_k \boldsymbol{\Theta}_k \in \Lambda(\Phi, \Psi)$ so that

$$\begin{aligned} [\Phi, \Psi]_r &\leqslant \int \langle v_1 - v_0, x_1 - x_0 \rangle \, d\boldsymbol{\Theta} \\ &= \sum_k \beta_k \int \langle v_1 - v_0, x_1 - x_0 \rangle \, d\boldsymbol{\Theta}_k \\ &= \sum_k \beta_k \, [\Phi_k, \Psi_k]_r \,. \end{aligned}$$

□

7.4 BEHAVIOUR OF DUALITY PAIRINGS ALONG GEODESICS

We know that the quantities $[\cdot,\cdot]_r$ and $[\cdot,\cdot]_l$ may be different when $\Gamma_o(\mu_0,\mu_1)$ is not a singleton. We thus expect a nice behaviour along geodesics. In the following definition, we use the notation

$$x^t(x_0,x_1) := (1-t)x_0 + tx_1, \quad v^0(x_0,v_0,x_1) := v_0$$

for every $(x_0,v_0,x_1) \in \mathsf{T}\mathbb{H} \times \mathbb{H}$, $t \in [0,1]$.

Definition 7.4.1. For $\boldsymbol{\vartheta} \in \mathcal{P}_2(\mathbb{H}\times\mathbb{H})$, $t \in [0,1]$, $\vartheta_t = x^t_\sharp\boldsymbol{\vartheta}$ and $\Phi_t \in \mathcal{P}_2(\mathsf{T}\mathbb{H}|\vartheta_t)$, we set

$$\Gamma_t(\Phi_t,\boldsymbol{\vartheta}) := \left\{\boldsymbol{\sigma} \in \mathcal{P}_2(\mathsf{T}\mathbb{H}\times\mathbb{H}) \mid (x^0,x^1)_\sharp\boldsymbol{\sigma} = \boldsymbol{\vartheta} \text{ and } (x^t\circ(x^0,x^1),v^0)_\sharp\boldsymbol{\sigma} = \Phi_t\right\}, \tag{7.4.1}$$

which is not empty since $\vartheta_t = x^t_\sharp\boldsymbol{\vartheta} = x_\sharp\Phi_t$. We set

$$[\Phi_t,\boldsymbol{\vartheta}]_{b,t} := \int \Big\langle x_0 - x_1, \mathbf{b}_{\Phi_t}(x^t(x_0,x_1))\Big\rangle\, d\boldsymbol{\vartheta}(x_0,x_1),$$
$$[\Phi_t,\boldsymbol{\vartheta}]_{r,t} := \min\left\{\int\langle x_0-x_1,v_0\rangle\, d\boldsymbol{\sigma}(x_0,v_0,x_1) \mid \boldsymbol{\sigma}\in\Gamma_t(\Phi_t,\boldsymbol{\vartheta})\right\},$$
$$[\Phi_t,\boldsymbol{\vartheta}]_{l,t} := \max\left\{\int\langle x_0-x_1,v_0\rangle\, d\boldsymbol{\sigma}(x_0,v_0,x_1) \mid \boldsymbol{\sigma}\in\Gamma_t(\Phi_t,\boldsymbol{\vartheta})\right\}.$$

If moreover $\Phi_0 \in \mathcal{P}_2(\mathsf{T}\mathbb{H}|\vartheta_0)$, $\Phi_1 \in \mathcal{P}_2(\mathsf{T}\mathbb{H}|\vartheta_1)$, $\boldsymbol{\vartheta} \in \Gamma(\vartheta_0,\vartheta_1)$, we define

$$[\Phi_0,\Phi_1]_{r,\boldsymbol{\vartheta}} := [\Phi_0,\boldsymbol{\vartheta}]_{r,0} - [\Phi_1,\boldsymbol{\vartheta}]_{l,1},$$
$$[\Phi_0,\Phi_1]_{l,\boldsymbol{\vartheta}} := [\Phi_0,\boldsymbol{\vartheta}]_{l,0} - [\Phi_1,\boldsymbol{\vartheta}]_{r,1}.$$

If $(\Phi_t)_x$ is the disintegration of Φ_t with respect to $\vartheta_t = x_\sharp\Phi_t$, we can consider the barycentric coupling $\boldsymbol{\sigma}_t := \int_{\mathbb{H}\times\mathbb{H}}(\Phi_t)_{x^t}\, d\boldsymbol{\vartheta} \in \Gamma_t(\Phi_t,\boldsymbol{\vartheta})$, i.e.

$$\int \psi(x_0,v_0,x_1)\, d\boldsymbol{\sigma}_t = \int\left[\int \psi(x_0,v_0,x_1)\, d(\Phi_t)_{(1-t)x_0+tx_1}(v_0)\right] d\boldsymbol{\vartheta}(x_0,x_1)$$

so that $[\Phi_t,\boldsymbol{\vartheta}]_{b,t} = \int\langle v_0, x_0-x_1\rangle\, d\boldsymbol{\sigma}_t$ and

$$[\Phi_t,\boldsymbol{\vartheta}]_{r,t} \leqslant [\Phi_t,\boldsymbol{\vartheta}]_{b,t} \leqslant [\Phi_t,\boldsymbol{\vartheta}]_{l,t}.$$

If we set

$$s : \mathbb{H}^2 \to \mathbb{H}^2, \quad s(x_0,x_1) := (x_1,x_0), \tag{7.4.2}$$

with an analogous definition in $\mathsf{T}\mathbb{H}\times\mathbb{H}$, given by $s(x_0,v_0,x_1) := (x_1,v_0,x_0)$, it is easy to check that

$$\boldsymbol{\sigma} \in \Gamma_t(\Phi_t,\boldsymbol{\vartheta}) \quad \Leftrightarrow \quad s_\sharp\boldsymbol{\sigma} \in \Gamma_{1-t}(\Phi_t, s_\sharp\boldsymbol{\vartheta})$$

so that

$$[\Phi_t,\boldsymbol{\vartheta}]_{r,t} = -[\Phi_t,s_\sharp\boldsymbol{\vartheta}]_{l,1-t}, \quad [\Phi_t,\boldsymbol{\vartheta}]_{l,t} = -[\Phi_t,s_\sharp\boldsymbol{\vartheta}]_{r,1-t}. \tag{7.4.3}$$

(7.1.11) and (7.1.14) can be simplified in some situations, as the following remark clarifies.

Remark 7.4.2 (Particular cases). Suppose that $\vartheta \in \mathcal{P}_2(\mathbb{H}^2)$, $t \in [0,1]$, $\vartheta_t = x^t_\sharp\vartheta$, $\Phi_t \in \mathcal{P}_2(\mathsf{T}\mathbb{H}|\vartheta_t)$ and $x^t : \mathbb{H}^2 \to \mathbb{H}$ is ϑ-essentially injective so that ϑ is concentrated on a Borel map

$$(X^0_t, X^1_t) : \mathbb{H} \to \mathbb{H} \times \mathbb{H}, \text{ i.e. } \vartheta = (X^0_t, X^1_t)_\sharp\vartheta_t.$$

In this case $\Gamma_t(\Phi_t, \vartheta)$ contains a unique element given by $(X^0_t \circ x, v, X^1_t \circ x)_\sharp\Phi_t$ and

$$\begin{aligned}[\Phi_t, \vartheta]_{r,t} &= [\Phi_t, \vartheta]_{l,t} = [\Phi_t, \vartheta]_{b,t} \\ &= \int \langle v, X^0_t(x) - X^1_t(x)\rangle \, d\Phi_t(x, v) = \int \langle \mathbf{b}_{\Phi_t}, X^0_t - X^1_t\rangle \, d\vartheta_t.\end{aligned} \tag{7.4.4}$$

When $t = 0$ and ϑ is the unique element of $\Gamma_o(\vartheta_0, \vartheta_1)$ then $X^0_t(x) = x$ and we obtain

$$\begin{aligned}[\Phi_t, \vartheta_1]_r &= [\Phi_t, \vartheta_1]_l = [\Phi_t, \vartheta]_{r,0} = [\Phi_t, \vartheta]_{l,0} \\ &= \int \langle v, x - X^1_t(x)\rangle \, d\Phi_t(x, v) = \int \langle \mathbf{b}_{\Phi_t}, x - X^1_t(x)\rangle \, d\vartheta_0(x).\end{aligned}$$

Another simple case is when

$$\Phi_t = (\mathbf{i}_\mathbb{H}, \boldsymbol{w})_\sharp\vartheta_t$$

for some vector field $\boldsymbol{w} \in L^2_{\vartheta_t}(\mathbb{H}; \mathbb{H})$ (i.e. its disintegration Φ_x w.r.t. ϑ_t takes the form $\delta_{\boldsymbol{w}(x)}$ and $\boldsymbol{w} = \mathbf{b}_{\Phi_t}$). We have

$$[\Phi_t, \vartheta]_{r,t} = [\Phi_t, \vartheta]_{l,t} = \int \Big\langle \boldsymbol{w}((1-t)x_0 + tx_1), x_0 - x_1 \Big\rangle \, d\vartheta(x_0, x_1).$$

In particular we get

$$[\Phi_t, \vartheta_1]_r = \min\left\{ \int \langle \boldsymbol{w}(x), x_0 - x_1\rangle \, d\vartheta(x_0, x_1) \mid \vartheta \in \Gamma_o(\vartheta_0, \vartheta_1) \right\}.$$

The following is a simple property of the brackets under restriction.

Lemma 7.4.3. *For every $\vartheta \in \mathcal{P}_2(\mathbb{H}^2)$, every $0 \leqslant s < t \leqslant 1$ and every $\Phi \in \mathcal{P}_2(\mathsf{T}\mathbb{H}|x^s_\sharp\vartheta)$, $\Psi \in \mathcal{P}_2(\mathsf{T}\mathbb{H}|x^t_\sharp\vartheta)$ we have*

$$[\Phi, \vartheta]_{r,s} = \frac{1}{t-s}[\Phi, (x^s, x^t)_\sharp\vartheta]_{r,0}, \quad [\Psi, \vartheta]_{l,t} = \frac{1}{t-s}[\Psi, (x^s, x^t)_\sharp\vartheta]_{l,1}. \tag{7.4.5}$$

Proof. If we define $T : \mathsf{T}\mathbb{H} \times \mathbb{H} \to \mathsf{T}\mathbb{H} \times \mathbb{H}$ and $\mathcal{L} : \mathcal{P}_2(\mathsf{T}\mathbb{H} \times \mathbb{H}) \to \mathbb{R}$ as

$$\begin{aligned} T(x_0, v_0, x_1) &:= (x^s(x_0, x_1), v_0, x^t(x_0, x_1)), \\ \mathcal{L}(\boldsymbol{\sigma}) &:= \int_{\mathsf{T}\mathbb{H} \times \mathbb{H}} \langle v_0, x_0 - x_1\rangle \, d\boldsymbol{\sigma}(x_0, v_0, x_1), \end{aligned}$$

it is clear that

$$\begin{aligned} [\Phi, \boldsymbol{\mu}]_{r,s} &= \inf\{\mathcal{L}(\boldsymbol{\sigma}) \mid \boldsymbol{\sigma} \in \Gamma_s(\Phi, \boldsymbol{\mu})\}, \\ [\Phi, (x^s, x^t)_\sharp\boldsymbol{\mu}]_{r,0} &= \inf\left\{\mathcal{L}(\boldsymbol{\sigma}) \mid \boldsymbol{\sigma} \in \Gamma_0(\Phi, (x^s, x^t)_\sharp\boldsymbol{\mu})\right\}. \end{aligned}$$

The first equality in the statement follows by $T_\sharp(\Gamma_s(\Phi, \boldsymbol{\mu})) = \Gamma_0(\Phi, (x^s, x^t)_\sharp\boldsymbol{\mu})$ and $\mathcal{L}(T_\sharp\boldsymbol{\sigma}) = (t-s)\mathcal{L}(\boldsymbol{\sigma})$ for every $\boldsymbol{\sigma} \in \mathcal{P}_2(\mathsf{T}\mathbb{H} \times \mathbb{H})$. The second equality follows from the first one and (7.4.2). □

The above Remark 7.4.2 applies to the case of geodesics in $\mathcal{P}_2(\mathbb{H})$.

Lemma 7.4.4. *Let $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{H})$, $\mu : [0,1] \to \mathcal{P}_2(\mathbb{H})$ be a constant speed geodesic induced by an optimal plan $\boldsymbol{\mu} \in \Gamma_o(\mu_0, \mu_1)$ by the relation*

$$\mu_t = \mathsf{x}^t_\sharp \boldsymbol{\mu}, \quad t \in [0,1], \quad \textit{where} \quad \mathsf{x}^t(x_0, x_1) = (1-t)x_0 + tx_1.$$

If $t \in (0,1)$, $\Phi_t \in \mathcal{P}_2(\mathsf{T}\mathbb{H}|\mu_t)$, $\hat{\boldsymbol{\mu}} = \mathsf{s}_\sharp \boldsymbol{\mu} \in \Gamma_o(\mu_1, \mu_0)$, with s the reversion map in (7.4.2), then

$$\begin{aligned} \frac{1}{1-t}[\Phi_t, \mu_1]_r &= \frac{1}{1-t}[\Phi_t, \mu_1]_l \\ &= [\Phi_t, \boldsymbol{\mu}]_{r,t} \\ &= [\Phi_t, \boldsymbol{\mu}]_{l,t} \\ &= -\frac{1}{t}[\Phi_t, \mu_0]_r \\ &= -\frac{1}{t}[\Phi_t, \mu_0]_l \\ &= -[\Phi_t, \hat{\boldsymbol{\mu}}]_{r,1-t} \\ &= -[\Phi_t, \hat{\boldsymbol{\mu}}]_{l,1-t}. \end{aligned} \tag{7.4.6}$$

Proof. The crucial fact is that $\mathsf{x}^t : \mathbb{H}^2 \to \mathbb{H}$ is injective on $\text{supp}(\boldsymbol{\mu})$ and thus a bijection on its image $\text{supp}(\mu_t)$. Indeed, take $(x_0, x_1), (x_0', x_1') \in \text{supp}(\boldsymbol{\mu})$, then

$$\begin{aligned} \left|\mathsf{x}^t(x_0,x_1) - \mathsf{x}^t(x_0',x_1')\right|^2 &= (1-t)^2|x_0-x_0'|^2 + t^2|x_1-x_1'|^2 \\ &\quad + 2t(1-t)\langle x_0-x_0', x_1-x_1'\rangle \\ &\geqslant (1-t)^2|x_0-x_0'|^2 + t^2|x_1-x_1'|^2 \end{aligned}$$

thanks to the cyclical monotonicity of $\text{supp}(\boldsymbol{\mu})$.

Then, for every $x \in \text{supp}(\mu_t)$, there exists a unique couple

$$(x_0, x_1) = (X^0_t(x), X^1_t(x)) \in \text{supp}(\boldsymbol{\mu})$$

s.t. $x = (1-t)x_0 + tx_1$, where we refer to Remark 7.4.2 for the definitions of X^0_t, X^1_t (cf. also [107, Theorem 5.29]). Hence, in Figure 1 all maps are bijections,

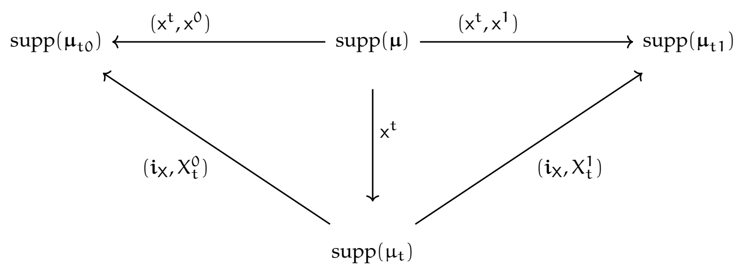

Figure 1: Interpolation of plans and maps.

where $\boldsymbol{\mu}_{t1} = (\mathsf{x}^t, \mathsf{x}^1)_\sharp \boldsymbol{\mu} = (\boldsymbol{i}_\mathbb{H}, X^1_t)_\sharp \mu_t$ is the unique element of $\Gamma_o(\mu_t, \mu_1)$ and

$\mu_{t0} = (x^t, x^0)_\sharp \mu = (\mathbf{i}_{\mathbb{H}}, X_t^0)_\sharp \mu_t = (x^{1-t}, x^1)_\sharp \hat{\mu}$ is the unique element of $\Gamma_o(\mu_t, \mu_0)$ (see Theorem 2.4.4). Since

$$\frac{x - X_t^1(x)}{1-t} = \frac{x - x_1}{1-t} = x_0 - x_1 = -\frac{x - x_0}{t} = -\frac{x - X_t^0(x)}{t},$$

and $\Lambda(\Phi_t, \mu_1) = \{(\mathbf{i}_{\mathsf{T}\mathbb{H}}, X_t^1 \circ x)_\sharp \Phi_t\}$ thanks to Theorem 2.4.4, by Theorem 7.1.8 and Corollary 7.1.9 we have

$$[\Phi_t, \mu_1]_r = [\Phi_t, \mu_1]_l = \int_{\mathsf{T}\mathbb{H}} \langle v, x - X_t^1(x)\rangle \, d\Phi_t(x, v).$$

Analogously, $\Lambda(\Phi_t, \mu_0) = \{(\mathbf{i}_{\mathsf{T}\mathbb{H}}, X_t^0 \circ x)_\sharp \Phi_t\}$. Hence

$$[\Phi_t, \mu_0]_r = [\Phi_t, \mu_0]_l = \int_{\mathsf{T}\mathbb{H}} \langle v, x - X_t^0(x)\rangle \, d\Phi_t(x, v).$$

Also recalling (7.4.3) and (7.4.4) we conclude. □

7.5 MULTIVALUED PROBABILITY VECTOR FIELDS AND DISSIPATIVITY

Definition 7.5.1. A *multivalued probability vector field* $\mathbf{F}$ is a nonempty subset of $\mathcal{P}_2(\mathsf{T}\mathbb{H})$ with domain $D(\mathbf{F}) := x_\sharp(\mathbf{F}) = \{x_\sharp \Phi : \Phi \in \mathbf{F}\}$. Given $\mu \in \mathcal{P}_2(\mathbb{H})$, we define the *section* $\mathbf{F}[\mu]$ of $\mathbf{F}$ as

$$\mathbf{F}[\mu] := (x_\sharp)^{-1}(\mu) \cap \mathbf{F} = \left\{\Phi \in \mathbf{F} \mid x_\sharp \Phi = \mu\right\}.$$

A *selection* $\mathbf{F}'$ of $\mathbf{F}$ is a subset of $\mathbf{F}$ such that $D(\mathbf{F}') = D(\mathbf{F})$. We call $\mathbf{F}$ a *probability vector field* (PVF) if $x_\sharp$ is injective in $\mathbf{F}$, i.e. $\mathbf{F}[\mu]$ contains a unique element for every $\mu \in D(\mathbf{F})$. A MPVF $\mathbf{F}$ is a vector field if for every $\mu \in D(\mathbf{F})$, the section $\mathbf{F}[\mu]$ contains a unique element Φ concentrated on a map, i.e. $\Phi = (\mathbf{i}_{\mathbb{H}}, \mathbf{b}_\Phi)_\sharp \mu$.

Remark 7.5.2. We can equivalently formulate Definition 7.5.1 by considering $\mathbf{F}$ as a multifunction, as in the case, e.g., of the Wasserstein subdifferential $\boldsymbol{\partial}\mathcal{F}$ of a function $\mathcal{F} : \mathcal{P}_2(\mathbb{H}) \to (-\infty, +\infty]$, see [5, Ch. 10] and Section 9.5.1. According to this viewpoint, a MPVF is a set-valued map $\mathbf{F} : \mathcal{P}_2(\mathbb{H}) \supset D(\mathbf{F}) \rightrightarrows \mathcal{P}_2(\mathsf{T}\mathbb{H})$ such that $x_\sharp \Phi = \mu$ for all $\Phi \in \mathbf{F}[\mu]$. In this way, each section $\mathbf{F}[\mu]$ is nothing but the image of $\mu \in D(\mathbf{F})$ through $\mathbf{F}$. In this case, *probability vector fields* correspond to single valued maps: this notion has been used in [94] with the aim of describing a sort of velocity field on $\mathcal{P}(\mathbb{H})$, and later in [93] dealing with Multivalued Probability Vector Fields (called Probability Multifunctions).

Definition 7.5.3 (Metrically λ-dissipative MPVF)**.** A MPVF $\mathbf{F} \subset \mathcal{P}_2(\mathsf{T}\mathbb{H})$ is (metrically) λ-*dissipative*, with $\lambda \in \mathbb{R}$, if

$$[\Phi_0, \Phi_1]_r \leqslant \lambda W_2^2(\mu_0, \mu_1) \quad \text{for every } \Phi_0, \Phi_1 \in \mathbf{F},\ \mu_i = x_\sharp \Phi_i. \tag{7.5.1}$$

We say that $\mathbf{F}$ is (metrically) λ-*accretive* if $-\mathbf{F} = \{-\Phi : \Phi \in \mathbf{F}\}$ (recall (7.1.10)) is $-\lambda$-dissipative, i.e.

$$[\Phi_0, \Phi_1]_l \geqslant \lambda W_2^2(\mu_0, \mu_1) \quad \text{for every } \Phi_0, \Phi_1 \in \mathbf{F},\ \mu_i = x_\sharp \Phi_i.$$

In Section 9.5 we present a few examples of λ-dissipative MPVFs.

Remark 7.5.4. Notice that (7.5.1) is equivalent to asking for the existence of a coupling $\boldsymbol{\Theta} \in \Lambda(\Phi_0, \Phi_1)$ (thus $(x^0, x^1)_\sharp \boldsymbol{\Theta}$ is optimal between $\mu_0 = x_\sharp \Phi_0$ and $\mu_1 = x_\sharp \Phi_1$) such that

$$\int \langle v_1 - v_0, x_1 - x_0 \rangle \, d\boldsymbol{\Theta} \leqslant \lambda W_2^2(\mu_0, \mu_1) = \lambda \int |x_1 - x_0|^2 \, d\boldsymbol{\Theta}.$$

The λ-dissipativity condition (7.5.1) has a natural metric interpretation: if $\Phi_0, \Phi_1 \in \mathbf{F}$ with $\mu_0 = x_\sharp \Phi_0$, $\mu_1 = x_\sharp \Phi_1$, performing a first order Taylor expansion of the map

$$t \mapsto \frac{1}{2} W_2^2(\exp^t \Phi_0, \exp^t \Phi_1)$$

at $t = 0$, recalling Definition 7.1.4, we have

$$W_2^2(\exp^t \Phi_0, \exp^t \Phi_1) \leqslant (1 + 2\lambda t) W_2^2(\mu_0, \mu_1) + o(t) \quad \text{as } t \downarrow 0.$$

Remark 7.5.5. Thanks to Corollary 7.1.6, (7.5.1) implies the weaker condition

$$[\Phi_0, \mu_1]_r + [\Phi_1, \mu_0]_r \leqslant \lambda W_2^2(\mu_0, \mu_1) \quad \text{for every } \Phi_0, \Phi_1 \in \mathbf{F},\ \mu_i = x_\sharp \Phi_i. \tag{7.5.2}$$

It is clear that the inequality of (7.5.2) implies the inequality of (7.5.1) whenever $\Gamma_o(\mu_0, \mu_1)$ contains only one element. More generally, we will see in Corollary 7.6.6 that (7.5.2) is in fact equivalent to (7.5.1) when $D(\mathbf{F})$ is geodesically convex (according to Definition 2.4.2).

Analogously to the Hilbertian setting, λ-dissipativity can be reduced to dissipativity (meaning 0-dissipativity) as shown in Lemma 7.5.6. Let us introduce the map

$$L^\lambda : \mathsf{TH} \to \mathsf{TH}, \quad L^\lambda(x, v) := (x, v - \lambda x).$$

Lemma 7.5.6. $\mathbf{F}$ *is a λ-dissipative* MPVF *(resp. satisfies* (7.5.2)*) if and only if* $\mathbf{F}^\lambda := L^\lambda_\sharp(\mathbf{F}) = \{L^\lambda_\sharp \Phi \mid \Phi \in \mathbf{F}\}$ *is dissipative (resp. satisfies* (7.5.2) *with* $\lambda = 0$*).*

Proof. Let us first check the case of (7.5.2). If $\boldsymbol{\sigma} \in \mathcal{P}_2(\mathsf{TH} \times \mathbb{H})$ with $(x^i)_\sharp \boldsymbol{\sigma} = \mu_i$, $i = 0, 1$, the transformed plan $\boldsymbol{\sigma}^\lambda := (L^\lambda, \boldsymbol{i}_{\mathbb{H}})_\sharp \boldsymbol{\sigma}$ satisfies

$$\int \langle v_0, x_0 - x_1 \rangle \, d\boldsymbol{\sigma}^\lambda = \int \langle v_0 - \lambda x_0, x_0 - x_1 \rangle \, d\boldsymbol{\sigma}$$
$$= \int \langle v_0, x_0 - x_1 \rangle \, d\boldsymbol{\sigma} - \frac{\lambda}{2} \int |x_0 - x_1|^2 \, d\boldsymbol{\sigma} \tag{7.5.3}$$
$$+ \frac{\lambda}{2} \Big(m_2^2(\mu_1) - m_2^2(\mu_0) \Big). \tag{7.5.4}$$

Since $\boldsymbol{\sigma} \in \Lambda_o(\Phi_0, \mu_1)$ if and only if $\boldsymbol{\sigma}^\lambda \in \Lambda_o(L^\lambda_\sharp \Phi_0, \mu_1)$, (7.5.3) yields

$$\int \langle v_0, x_0 - x_1 \rangle \, d\boldsymbol{\sigma}^\lambda = \int \langle v_0, x_0 - x_1 \rangle \, d\boldsymbol{\sigma}$$
$$- \frac{\lambda}{2} \Big(m_2^2(\mu_0) - m_2^2(\mu_1) + W_2^2(\mu_0, \mu_1) \Big)$$

and therefore

$$[L^\lambda_\sharp \Phi_0, \mu_1]_r = [\Phi_0, \mu_1]_r - \frac{\lambda}{2}\Big(m_2^2(\mu_0) - m_2^2(\mu_1) + W_2^2(\mu_0, \mu_1)\Big). \tag{7.5.5}$$

Using the corresponding identity for $\left[L^\lambda_\sharp \Phi_1, \mu_0\right]_r$ we obtain that $\mathbf{F}^\lambda$ is dissipative.

Similarly, if $\boldsymbol{\Theta} \in \mathcal{P}_2(\mathsf{T}\mathbb{H} \times \mathsf{T}\mathbb{H})$ with $x^i_\sharp \boldsymbol{\Theta} = \mu_i$, the plan $\boldsymbol{\Theta}^\lambda := (L^\lambda, L^\lambda)_\sharp \boldsymbol{\Theta}$ satisfies

$$\begin{aligned}\int \langle v_0 - v_1, x_0 - x_1\rangle \, d\boldsymbol{\Theta}^\lambda &= \int \langle v_0 - v_1 - \lambda(x_0 - x_1), x_0 - x_1\rangle \, d\boldsymbol{\Theta} \\ &= \int \langle v_0 - v_1, x_0 - x_1\rangle \, d\boldsymbol{\Theta} - \lambda \int |x_0 - x_1|^2 \, d\boldsymbol{\Theta}.\end{aligned} \tag{7.5.6}$$

Reasoning with a similar argument as for the case of assumption (7.5.2), using the identity (7.5.6), we get the equivalence between the λ-dissipativity of $\mathbf{F}$ and the dissipativity of $\mathbf{F}^\lambda$. □

We conclude the section showing that a Lipschitz like condition similar to the one considered in [94] (see Section 9.5.5) implies λ-dissipativity.

Lemma 7.5.7. *Suppose that the MPVF $\mathbf{F}$ satisfies*

$$\mathcal{W}_2(\mathbf{F}[\nu], \mathbf{F}[\nu']) \leqslant L W_2(\nu, \nu') \quad \textit{for every } \nu, \nu' \in D(\mathbf{F}),$$

where $\mathcal{W}_2 : \mathcal{P}_2(\mathsf{T}\mathbb{H}) \times \mathcal{P}_2(\mathsf{T}\mathbb{H}) \to [0, +\infty)$ is defined by

$$\mathcal{W}_2^2(\Phi_0, \Phi_1) = \inf\left\{\int_{\mathsf{T}\mathbb{H}\times\mathsf{T}\mathbb{H}} |v_0 - v_1|^2 \, d\boldsymbol{\Theta}(x_0, v_0, x_1, v_1) : \boldsymbol{\Theta} \in \Lambda(\Phi_0, \Phi_1)\right\},$$

with $\Lambda(\cdot,\cdot)$ as in Definition 7.1.7. Then $\mathbf{F}$ is λ-dissipative according to (7.5.1), for $\lambda := \frac{1}{2}(1 + L^2)$

Proof. Let $\nu', \nu'' \in D(\mathbf{F})$, then by Theorem 7.1.8 and Young's inequality, we have

$$\begin{aligned}\left[\mathbf{F}[\nu'], \mathbf{F}[\nu'']\right]_r &= \min\left\{\int_{\mathsf{T}\mathbb{H}\times\mathsf{T}\mathbb{H}} \langle x' - x'', v' - v''\rangle \, d\boldsymbol{\Theta} \;:\; \boldsymbol{\Theta} \in \Lambda(\mathbf{F}[\nu'], \mathbf{F}[\nu''])\right\} \\ &\leqslant \frac{1}{2}\left(W_2^2(\nu', \nu'') + \mathcal{W}_2^2(\mathbf{F}[\nu'], \mathbf{F}[\nu''])\right) \\ &\leqslant \frac{L^2 + 1}{2} W_2^2(\nu', \nu'').\end{aligned}$$

□

7.6 BEHAVIOUR OF DISSIPATIVE MPVFS ALONG GEODESICS

We want to study the interaction between MPVFs and geodesic. In the classical Hilbert setting, given a map $F : H \to H$ in a Hilbert space H, it is easy to see that the function

$$f(t) := \langle F(x_t), x_0 - x_1\rangle, \quad x_t = (1-t)x_0 + t x_1, \quad t \in [0, 1] \tag{7.6.1}$$

is monotone increasing.

Let $\mathbf{F} \subset \mathcal{P}_2(\mathsf{TH})$, $\mu_0, \mu_1 \in \overline{D(\mathbf{F})}$, $\boldsymbol{\mu} \in \Gamma_o(\mu_0, \mu_1)$. In order to compute the measure-theoretic analogue of the scalar product in (7.6.1), we need to define the set

$$I(\boldsymbol{\mu}|\mathbf{F}) := \left\{ t \in [0,1] : x^t_\sharp \boldsymbol{\mu} \in D(\mathbf{F}) \right\}, \tag{7.6.2}$$

since we can evaluate the MPVF $\mathbf{F}$ along geodesics only for time instants $t \in [0,1]$ at which they lie inside the domain.

Definition 7.6.1. Let $\mathbf{F} \subset \mathcal{P}_2(\mathsf{TH})$ be a MPVF. Let $\mu_0, \mu_1 \in \overline{D(\mathbf{F})}$, $\boldsymbol{\mu} \in \Gamma_o(\mu_0, \mu_1)$ and let $\mu_t := x^t_\sharp \boldsymbol{\mu}$, $t \in [0,1]$. For every $t \in I(\boldsymbol{\mu}|\mathbf{F})$ we define

$$[\mathbf{F}, \boldsymbol{\mu}]_{r,t} := \sup \left\{ [\Phi_t, \boldsymbol{\mu}]_{r,t} \mid \Phi_t \in \mathbf{F}[\mu_t] \right\},$$
$$[\mathbf{F}, \boldsymbol{\mu}]_{l,t} := \inf \left\{ [\Phi_t, \boldsymbol{\mu}]_{l,t} \mid \Phi_t \in \mathbf{F}[\mu_t] \right\}.$$

Theorem 7.6.2. *Let us suppose that the MPVF $\mathbf{F}$ satisfies (7.5.2), let $\mu_0, \mu_1 \in \overline{D(\mathbf{F})}$, and let $\boldsymbol{\mu} \in \Gamma_o(\mu_0, \mu_1)$. Then the following properties hold*

1. $[\mathbf{F}, \boldsymbol{\mu}]_{l,t} \leqslant [\mathbf{F}, \boldsymbol{\mu}]_{r,t}$ *for every* $t \in (0,1) \cap I(\boldsymbol{\mu}|\mathbf{F})$;

2. $[\mathbf{F}, \boldsymbol{\mu}]_{r,s} \leqslant [\mathbf{F}, \boldsymbol{\mu}]_{l,t} + \lambda(t-s) W_2^2(\mu_0, \mu_1)$ *for every* $s, t \in I(\boldsymbol{\mu}|\mathbf{F})$, $s < t$;

3. *the maps*

   $$t \mapsto [\mathbf{F}, \boldsymbol{\mu}]_{r,t} + \lambda t\, W_2^2(\mu_0, \mu_1) \quad \textit{and} \quad t \mapsto [\mathbf{F}, \boldsymbol{\mu}]_{l,t} + \lambda t\, W_2^2(\mu_0, \mu_1)$$

   are increasing respectively in $I(\boldsymbol{\mu}|\mathbf{F}) \setminus \{1\}$ *and in* $I(\boldsymbol{\mu}|\mathbf{F}) \setminus \{0\}$;

4. *if* t_0 *is a right accumulation point of* $I(\boldsymbol{\mu}|\mathbf{F})$, *then*

   $$\lim_{t \downarrow t_0} [\mathbf{F}, \boldsymbol{\mu}]_{r,t} = \lim_{t \downarrow t_0} [\mathbf{F}, \boldsymbol{\mu}]_{l,t} \tag{7.6.3}$$

   and these right limits exist. If, instead, t_0 *is a left accumulation point of* $I(\boldsymbol{\mu}|\mathbf{F})$, *the same holds with the right limits in (7.6.3) replaced by the left limits at* t_0;

(5) $[\mathbf{F}, \boldsymbol{\mu}]_{l,t} = [\mathbf{F}, \boldsymbol{\mu}]_{r,t}$ *at every interior point* t *of* $I(\boldsymbol{\mu}|\mathbf{F})$ *where one of them is continuous.*

Proof. Throughout all the proof we set

$$f_r(t) := [\mathbf{F}, \boldsymbol{\mu}]_{r,t} \quad \text{and} \quad f_l(t) := [\mathbf{F}, \boldsymbol{\mu}]_{l,t}. \tag{7.6.4}$$

Thanks to Lemma 7.5.6 and in particular to (7.5.5), it is easy to check that it is sufficient to consider the dissipative case $\lambda = 0$.

1. It is a direct consequence of Lemma 7.4.4 and the definitions of f_r and f_l.

2. We prove that for every $\Phi_s \in \mathbf{F}[\mu_s]$ and $\Phi'_t \in \mathbf{F}[\mu_t]$ it holds

$$[\Phi_s, \boldsymbol{\mu}]_{r,s} \leqslant [\Phi'_t, \boldsymbol{\mu}]_{l,t}. \tag{7.6.5}$$

The thesis will follow immediately passing to the sup over $\Phi_s \in \mathbf{F}[\mu_s]$ in the l.h.s. and to the inf over $\Phi'_t \in \mathbf{F}[\mu_t]$ in the r.h.s.. It is enough to prove (7.6.5) in case at least one between s, t belongs to $(0, 1)$. Let us define the map $L : \mathcal{P}_2(\mathsf{T}\mathbb{H} \times \mathbb{H}) \to \mathbb{R}$ as

$$L(\gamma) := \int_{\mathsf{T}\mathbb{H}\times\mathbb{H}} \langle v_0, x_0 - x_1\rangle \, d\gamma(x_0, v_0, x_1) \quad \gamma \in \mathcal{P}_2(\mathsf{T}\mathbb{H} \times \mathbb{H}).$$

Observe that, since it never happens that $s = 0$ and $t = 1$ at the same time, the map $T_{s,t} : \Gamma_s(\Phi_s, \boldsymbol{\mu}) \to \Lambda(\Phi_s, \mu_t)$, with $\Gamma_s(\cdot, \cdot)$ as in (7.4.1) and $\Lambda(\cdot, \cdot)$ as in Definition 7.1.7, defined as

$$T_{s,t}(\boldsymbol{\sigma}) := \left(x^s \circ (x^0, x^1), v^0, x^t \circ (x^0, x^1)\right)_\sharp \boldsymbol{\sigma}$$

is a bijection s.t. $(t - s)L(\boldsymbol{\sigma}) = L(T_{s,t}(\boldsymbol{\sigma}))$ for every $\boldsymbol{\sigma} \in \Gamma_s(\Phi_s, \boldsymbol{\mu})$. This immediately gives that

$$(t - s)[\Phi_s, \boldsymbol{\mu}]_{r,s} = [\Phi_s, \mu_t]_r \, .$$

In the same way we can deduce that

$$(s - t)[\Phi'_t, \boldsymbol{\mu}]_{l,t} = \left[\Phi'_t, \mu_s\right]_r \, .$$

Thanks to the dissipativity assumption (7.5.2) of $\mathbf{F}$, we get

$$(t - s)[\Phi_s, \boldsymbol{\mu}]_{r,s} - (t - s)[\Phi'_t, \boldsymbol{\mu}]_{l,t} = [\Phi_s, \mu_t]_r + \left[\Phi'_t, \mu_s\right]_r \leqslant 0.$$

3. Combining (1) and (2) we have that for every $s, t \in I(\boldsymbol{\mu}|\mathbf{F})$ with $0 < s < t < 1$ it holds

$$f_l(s) \leqslant f_r(s) \leqslant f_l(t) \leqslant f_r(t), \tag{7.6.6}$$

with f_r, f_l as in (7.6.4). This implies that both f_l and f_r are increasing in $I(\boldsymbol{\mu}|\mathbf{F}) \cap (0, 1)$. Observe that, again combining (1) and (2), it also holds

$$f_r(0) \leqslant f_l(t) \leqslant f_r(t),$$
$$f_l(t) \leqslant f_r(t) \leqslant f_l(1)$$

for every $t \in I(\boldsymbol{\mu}|\mathbf{F}) \setminus \{0, 1\}$, and then f_r is increasing in $I(\boldsymbol{\mu}|\mathbf{F}) \setminus \{1\}$ and f_l is increasing in $I(\boldsymbol{\mu}|\mathbf{F}) \setminus \{0\}$.

4. It is an immediate consequence of (7.6.6).

5. It is a straightforward consequence of (4). □

Thanks to Theorem 7.6.2(4), we have

$$\lim_{t\downarrow 0}[\mathbf{F},\boldsymbol{\mu}]_{r,t} = \lim_{t\downarrow 0}[\mathbf{F},\boldsymbol{\mu}]_{l,t},$$
$$\lim_{t\uparrow 1}[\mathbf{F},\boldsymbol{\mu}]_{r,t} = \lim_{t\uparrow 1}[\mathbf{F},\boldsymbol{\mu}]_{l,t},$$

and those limits exist whenever the starting time $t_0 = 0$ and the final time $t_1 = 1$ are accumulation points of $I(\boldsymbol{\mu}|\mathbf{F})$, respectively. Due to the importance played by these objects in Section 8.1, we give the following definitions. These are intended to weaken the requirement for the operator's domain $D(\mathbf{F})$ to be open or geodesically convex.

Definition 7.6.3. Let $\mathbf{F} \subset \mathcal{P}_2(\mathsf{TH})$, $\mu_0, \mu_1 \in \overline{D(\mathbf{F})}$, $\boldsymbol{\mu} \in \Gamma_o(\mu_0,\mu_1)$. We define the sets

$$\Gamma_o^i(\mu_0,\mu_1|\mathbf{F}) := \left\{\boldsymbol{\mu} \in \Gamma_o(\mu_0,\mu_1) : i \text{ is an acc. point of } I(\boldsymbol{\mu}|\mathbf{F})\right\}, i = 0,1 \tag{7.6.7}$$

$$\Gamma_o^{01}(\mu_0,\mu_1|\mathbf{F}) := \Gamma_o^0(\mu_0,\mu_1|\mathbf{F}) \cap \Gamma_o^1(\mu_0,\mu_1|\mathbf{F}), \tag{7.6.8}$$

$$\Gamma(\mu_0,\mu_1|\mathbf{F}) := \left\{\boldsymbol{\mu} \in \Gamma(\mu_0,\mu_1) \mid \mathsf{x}^t_\sharp\boldsymbol{\mu} \in D(\mathbf{F}) \text{ for every } t \in [0,1]\right\} \tag{7.6.9}$$

Notice that these sets depend on $\mathbf{F}$ just through $D(\mathbf{F})$. In particular, if $\mu_0, \mu_1 \in D(\mathbf{F})$ and $D(\mathbf{F})$ is open or geodesically convex according to Definition 2.4.2 then $\Gamma_o^{01}(\mu_0,\mu_1|\mathbf{F}) \neq \emptyset$.

By the previous discussion, the next definition is well posed.

Definition 7.6.4. Let us suppose that the MPVF $\mathbf{F}$ satisfies (7.5.2), let $\mu_0, \mu_1 \in \overline{D(\mathbf{F})}$.

$$\text{If } \boldsymbol{\mu} \in \Gamma_o^0(\mu_0,\mu_1|\mathbf{F}) \text{ we set } \quad [\mathbf{F},\boldsymbol{\mu}]_{0+} := \lim_{t\downarrow 0}[\mathbf{F},\boldsymbol{\mu}]_{r,t} = \lim_{t\downarrow 0}[\mathbf{F},\boldsymbol{\mu}]_{l,t}.$$
$$\text{If } \boldsymbol{\mu} \in \Gamma_o^1(\mu_0,\mu_1|\mathbf{F}) \text{ we set } \quad [\mathbf{F},\boldsymbol{\mu}]_{1-} := \lim_{t\uparrow 1}[\mathbf{F},\boldsymbol{\mu}]_{r,t} = \lim_{t\uparrow 1}[\mathbf{F},\boldsymbol{\mu}]_{l,t}.$$

In the following statements, we make use of the objects introduced in Definition 7.6.3 in order to get refined dissipativity conditions involving the limiting pseudo-scalar products of Definition 7.6.4. These results will be useful in the sequel: in Proposition 7.7.3 they allow to get a dissipativity property of a suitable notion of extension $\hat{\mathbf{F}}$ of $\mathbf{F}$; in Section 8.1 (see in particular Lemma 8.1.3) they are relevant to study the properties of so-called λ-EVI solutions for a λ-dissipative MPVF $\mathbf{F}$.

Corollary 7.6.5. *Let us keep the same notation of Theorem 7.6.2 and let $s \in I(\boldsymbol{\mu}|\mathbf{F}) \cap (0,1)$ with $\Phi \in \mathbf{F}[\mu_s]$.*

1. *If $\boldsymbol{\mu} \in \Gamma_o^0(\mu_0,\mu_1|\mathbf{F})$, we have that*

$$[\mathbf{F},\boldsymbol{\mu}]_{0+} \leqslant [\Phi,\boldsymbol{\mu}]_{l,s} + \lambda s W^2 = [\Phi,\boldsymbol{\mu}]_{r,s} + \lambda s W^2; \tag{7.6.10}$$

if moreover $\Phi_0 \in \mathbf{F}[\mu_0]$ then

$$[\Phi_0,\mu_1]_r \leqslant [\Phi_0,\boldsymbol{\mu}]_{r,0} \leqslant [\mathbf{F},\boldsymbol{\mu}]_{0+}. \tag{7.6.11}$$

2. *If* $\mu \in \Gamma^1_o(\mu_0, \mu_1|\mathbf{F})$, *we have that*

$$[\Phi, \mu]_{l,s} - \lambda(1-s)W^2 = [\Phi, \mu]_{r,s} - \lambda(1-s)W^2 \leqslant [\mathbf{F}, \mu]_{1-};$$

if moreover $\Phi_1 \in \mathbf{F}[\mu_1]$ *then*

$$[\mathbf{F}, \mu]_{1-} \leqslant [\Phi_1, \mu]_{l,1} \leqslant -[\Phi_1, \mu_0]_r \tag{7.6.12}$$

3. *In particular, for every* $\Phi_0 \in \mathbf{F}[\mu_0]$, $\Phi_1 \in \mathbf{F}[\mu_1]$ *and* $\mu \in \Gamma^{01}_o(\mu_0, \mu_1|\mathbf{F})$ *we obtain*

$$[\Phi_0, \Phi_1]_{r,\mu} \leqslant [\mathbf{F}, \mu]_{0+} - [\mathbf{F}, \mu]_{1-} \leqslant \lambda W_2^2(\mu_0, \mu_1). \tag{7.6.13}$$

(7.6.13) immediately yields the following property.

Corollary 7.6.6. *Suppose that a MPVF* $\mathbf{F}$ *satisfies*

$$\textit{for every } \mu_0, \mu_1 \in D(\mathbf{F}) \textit{ the set } \Gamma^{01}_o(\mu_0, \mu_1|\mathbf{F}) \textit{ of (7.6.8) is not empty} \tag{7.6.14}$$

(e.g. if $D(\mathbf{F})$ *is open or geodesically convex), then* $\mathbf{F}$ *is* λ*-dissipative according to* (7.5.1) *if and only if it satisfies* (7.5.2).

Proposition 7.6.7. *Let* $\mathbf{F} \subset \mathcal{P}_2(\mathsf{TH})$ *be a MPVF satisfying* (7.5.2), *let* $\mu_0 \in \overline{D(\mathbf{F})}$ *and let* $\Phi \in \mathcal{P}_2(\mathsf{TH}|\mu_0)$. *Consider the following statements*

(P1) $[\Phi, \mu]_r + [\Psi, \mu_0]_r \leqslant \lambda W_2^2(\mu_0, \mu)$ *for every* $\Psi \in \mathbf{F}$ *with* $\mu = x_\sharp \Psi$;

(P2) *for every* $\mu \in D(\mathbf{F})$ *there exists* $\Psi \in \mathbf{F}[\mu]$ *s.t.* $[\Phi, \mu]_r + [\Psi, \mu_0]_r \leqslant \lambda W_2^2(\mu_0, \mu)$;

(P3) $[\Phi, \mu]_{r,0} \leqslant [\mathbf{F}, \mu]_{0+}$ *for every* $\mu_1 \in \overline{D(\mathbf{F})}$, $\mu \in \Gamma^0_o(\mu_0, \mu_1|\mathbf{F})$;

(P4) $[\Phi, \mu]_{r,0} \leqslant [\mathbf{F}, \mu]_{0+}$ *for every* $\mu_1 \in D(\mathbf{F})$, $\mu \in \Gamma^0_o(\mu_0, \mu_1|\mathbf{F})$;

(P5) $[\Phi, \mu]_{r,0} \leqslant \lambda W_2^2(\mu_0, \mu_1) + [\mathbf{F}, \mu]_{1-}$ *for every* $\mu_1 \in \overline{D(\mathbf{F})}$, $\mu \in \Gamma^1_o(\mu_0, \mu_1|\mathbf{F})$;

(P6) $[\Phi, \mu]_{r,0} \leqslant \lambda W_2^2(\mu_0, \mu_1) + [\mathbf{F}, \mu]_{1-}$ *for every* $\mu_1 \in D(\mathbf{F})$, $\mu \in \Gamma^1_o(\mu_0, \mu_1|\mathbf{F})$.

Then the following hold

1. (P1) $\Rightarrow$ (P2) $\Rightarrow$ (P3) $\Rightarrow$ (P4);
2. (P1) $\Rightarrow$ (P2) $\Rightarrow$ (P5) $\Rightarrow$ (P6);
3. *if for every* $\mu_1 \in D(\mathbf{F})$ $\Gamma^0_o(\mu_0, \mu_1|\mathbf{F}) \neq \emptyset$, *then* (P4) $\Rightarrow$ (P1) *(in particular,* (P1), (P2), (P3), (P4) *are equivalent);*
4. *if for every* $\mu_1 \in D(\mathbf{F})$ $\Gamma^1_o(\mu_0, \mu_1|\mathbf{F}) \neq \emptyset$, *then* (P6) $\Rightarrow$ (P1) *(in particular,* (P1), (P2), (P5), (P6) *are equivalent).*

Proof. We first prove that (P2) ⇒ (P3),(P5). Let us choose an arbitrary $\mu_1 \in \overline{D(\mathbf{F})}$; by the definition of $[\mathbf{F}, \boldsymbol{\mu}]_{r,t}$ and arguing as in the proof of Theorem 7.6.2(2), for all $\boldsymbol{\mu} \in \Gamma_o(\mu_0, \mu_1)$ and $t \in I(\boldsymbol{\mu}|\mathbf{F})$ there exists $\Psi_t \in \mathbf{F}[\mu_t]$ such that

$$\begin{aligned} [\Phi, \boldsymbol{\mu}]_{r,0} &= \frac{1}{t}\, [\Phi, \mu_t]_r \\ &\leqslant -\frac{1}{t}\, [\Psi_t, \mu_0]_r + t\lambda W_2^2(\mu_0, \mu_1) \\ &= [\Psi_t, \boldsymbol{\mu}]_{r,t} + t\lambda W_2^2(\mu_0, \mu_1) \\ &\leqslant [\mathbf{F}, \boldsymbol{\mu}]_{r,t} + t\lambda W_2^2(\mu_0, \mu_1), \end{aligned}$$

where we also used (7.4.6). If $\boldsymbol{\mu} \in \Gamma_o^0(\mu_0, \mu_1|\mathbf{F})$, by passing to the limit as $t \downarrow 0$ we get (P3).

In the second case, assuming that $\boldsymbol{\mu} \in \Gamma_o^1(\mu_0, \mu_1|\mathbf{F})$, we can pass to the limit as $t \uparrow 1$ and we get (P5).

We now prove item (3). Let $\mu_1 \in D(\mathbf{F})$, $\Psi \in \mathbf{F}[\mu_1]$, $\boldsymbol{\mu} \in \Gamma_o^0(\mu_0, \mu_1|\mathbf{F})$, $s \in I(\boldsymbol{\mu}|\mathbf{F}) \cap (0,1)$, $\Phi_s \in \mathbf{F}[\mu_s]$, with $\mu_s = \mathsf{x}^s_\sharp \boldsymbol{\mu}$. Assuming (P4) and using (7.6.11), (7.6.10), (7.4.6) and (7.5.2), we have

$$\begin{aligned} [\Phi, \mu_1]_r \leqslant [\Phi, \boldsymbol{\mu}]_{r,0} \leqslant [\mathbf{F}, \boldsymbol{\mu}]_{0+} &\leqslant [\Phi_s, \boldsymbol{\mu}]_{r,s} + \lambda s W_2^2(\mu_0, \mu_1) \\ &= \frac{1}{1-s}\, [\Phi_s, \mu_1]_r + \lambda s W_2^2(\mu_0, \mu_1) \\ &\leqslant -\frac{1}{1-s}\, [\Psi, \mu_s]_r + \lambda(1+s) W_2^2(\mu_0, \mu_1). \end{aligned}$$

By Lemma 7.3.1, letting $s \downarrow 0$ we get (P1). Item (4) follows by (7.6.11), (7.6.12). □

7.7 EXTENSIONS OF DISSIPATIVE MPVFS

Let us describe a few properties of extensions of λ-dissipative MPVFs. The first one is about the sequential closure in $\mathcal{P}_2^{sw}(\mathsf{T}\mathbb{H})$: given $A \subset \mathcal{P}_2(\mathsf{T}\mathbb{H})$, we denote by $\mathrm{cl}(A)$ its sequential closure defined by

$$\mathrm{cl}(A) := \Big\{ \Phi \in \mathcal{P}_2(\mathsf{T}\mathbb{H}) : \exists\, (\Phi_n)_{n\in\mathbb{N}} \subset A : \Phi_n \to \Phi \text{ in } \mathcal{P}_2^{sw}(\mathsf{T}\mathbb{H}) \Big\}.$$

Proposition 7.7.1. *If* $\mathbf{F}$ *is a* λ*-dissipative MPVF according to* (7.5.1)*, then its sequential closure* $\mathrm{cl}(\mathbf{F})$ *is* λ*-dissipative as well according to* (7.5.1)*.*

Proof. If Φ^i, $i = 0, 1$, belong to $\mathrm{cl}(\mathbf{F})$, we can find sequences $(\Phi^i_n)_{n\in\mathbb{N}} \subset \mathbf{F}$ such that $\Phi^i_n \to \Phi^i$ in $\mathcal{P}_2^{sw}(\mathsf{T}\mathbb{H})$ as $n \to \infty$, $i = 0, 1$. It is then sufficient to pass to the limit in the inequality

$$\left[\Phi^0_n, \Phi^1_n\right]_r \leqslant \lambda W_2^2(\mu^0_n, \mu^1_n), \quad \mu^i_n = \mathsf{x}_\sharp \Phi^i_n$$

using the lower semicontinuity property (7.3.3) and the fact that convergence in $\mathcal{P}_2^{sw}(\mathsf{T}\mathbb{H})$ yields $\mu^i_n \to \mathsf{x}_\sharp \Phi^i$ in $\mathcal{P}_2(\mathbb{H})$ as $n \to \infty$. □

A second result concerns the convexification of the sections of $\mathbf{F}$. For every $\mu \in D(\mathbf{F})$ we set

$$\begin{aligned} \operatorname{co}(\mathbf{F})[\mu] &:= \text{the convex hull of } \mathbf{F}[\mu] \\ &= \left\{ \sum_k \alpha_k \Phi_k : \Phi_k \in \mathbf{F}[\mu], \alpha_k \geqslant 0, \sum_k \alpha_k = 1 \right\}, \end{aligned} \tag{7.7.1}$$

$$\overline{\operatorname{co}}(\mathbf{F})[\mu] := \operatorname{cl}(\operatorname{co}(\mathbf{F})[\mu]). \tag{7.7.2}$$

Notice that if $\mathbf{F}[\mu]$ is bounded in $\mathcal{P}_2(\mathsf{T\!H})$ then $\overline{\operatorname{co}}(\mathbf{F})[\mu]$ coincides with the closed convex hull of $\mathbf{F}[\mu]$.

Proposition 7.7.2. *If $\mathbf{F}$ is λ-dissipative according to (7.5.1), then $\operatorname{co}(\mathbf{F})$ and $\overline{\operatorname{co}}(\mathbf{F})$ are λ-dissipative as well according to (7.5.1).*

Proof. By Proposition 7.7.1 and noting that $\overline{\operatorname{co}}(\mathbf{F}) \subset \operatorname{cl}(\operatorname{co}(\mathbf{F}))$, it is sufficient to prove that $\operatorname{co}(\mathbf{F})$ is λ-dissipative. By Lemma 7.5.6 it is not restrictive to assume $\lambda = 0$. Let $\Phi^i \in \operatorname{co}(\mathbf{F})[\mu_i]$, $i = 0, 1$; there exist positive coefficients α^i_k, $k = 1, \cdots, K$, with $\sum_k \alpha^i_k = 1$, and elements $\Phi^i_k \in \mathbf{F}[\mu^i]$, $i = 0, 1$, such that $\Phi^i = \sum_{k=1}^K \alpha^i_k \Phi^i_k$. Setting $\beta_{h,k} := \alpha^0_h \alpha^1_k$, we can apply Lemma 7.3.3 and we obtain

$$\left[\Phi^0, \Phi^1\right]_r = \left[\sum_{h,k} \beta_{h,k} \Phi^0_h, \sum_{h,k} \beta_{h,k} \Phi^1_k\right]_r \leqslant \sum_{h,k} \beta_{h,k} \left[\Phi^0_h, \Phi^1_k\right]_r \leqslant 0. \qquad \square$$

We recall that in the Hilbertian case (cf. e.g. [26]), a fundamental role is played by the notion of maximality for a dissipative operator $F \subset H \times H$. Indeed, this notion enables to extablish the existence and uniqueness of solutions of the corresponding evolution equation and to get crucial properties of the resolvent operator. Moreover, if F is maximal, in order to prove that an element $(x, v) \in H \times H$ belongs to F it is enough to verify that it satisfies the dissipativity inequality

$$\langle v - w, x - y \rangle \leqslant 0 \quad \text{for every } (y, w) \in F. \tag{7.7.3}$$

For these reasons, if F is not maximal it is important to study its maximal extension, whose elements (x, v) must satisfy (7.7.3).

By analogy with the Hilbertian framework, it is interesting to study the properties of the *extended MPVF* defined by

$$\hat{\mathbf{F}} := \left\{ \Phi \in \mathcal{P}_2(\mathsf{T\!H}) : \begin{array}{l} \mu = x_\sharp \Phi \in \overline{D(\mathbf{F})}, \\[1ex] [\Phi, \nu]_r + [\Psi, \mu]_r \leqslant \lambda W_2^2(\mu, \nu) \quad \forall \Psi \in \mathbf{F},\ \nu = x_\sharp \Psi \end{array} \right\}. \tag{7.7.4}$$

This notion of extension $\hat{\mathbf{F}}$ of a MPVF $\mathbf{F}$ will be involved later in Section 8.1 dealing with differential inclusions in Wasserstein spaces, in particular in Theorem 8.1.4 and in Subsection 8.3.

It is obvious that $\mathbf{F} \subset \hat{\mathbf{F}}$; if the domain of $\mathbf{F}$ satisfies the geometric condition (7.7.6), the following result shows that $\hat{\mathbf{F}}$ provides the maximal λ-dissipative extension of $\mathbf{F}$.

Proposition 7.7.3. *Let $\mathbf{F}$ be a λ-dissipative MPVF according to (7.5.1).*

(a) *If* $\mathbf{F}' \supset \mathbf{F}$ *is* λ*-dissipative according to* (7.5.1)*, with* $\mathrm{D}(\mathbf{F}') \subset \overline{\mathrm{D}(\mathbf{F})}$*, then* $\mathbf{F}' \subset \hat{\mathbf{F}}$*. In particular* $\overline{\mathrm{co}}(\mathrm{cl}(\mathbf{F})) \subset \hat{\mathbf{F}}$*.*

(b) $\widehat{\mathrm{cl}(\mathbf{F})} = \hat{\mathbf{F}}$ *and* $\widehat{\mathrm{co}(\mathbf{F})} = \hat{\mathbf{F}}$*.*

(c) $\hat{\mathbf{F}}$ *is sequentially closed and* $\hat{\mathbf{F}}[\mu]$ *is convex for every* $\mu \in \mathrm{D}(\hat{\mathbf{F}})$*.*

(d) *If* $\mathrm{D}(\mathbf{F})$ *satisfies* (7.6.14)*, then the restriction of* $\hat{\mathbf{F}}$ *to* $\mathrm{D}(\mathbf{F})$ *is* λ*-dissipative according to* (7.5.1) *and for every* $\mu_0, \mu_1 \in \mathrm{D}(\mathbf{F})$ *it holds*

$$[\mathbf{F}, \boldsymbol{\mu}]_{0+} = [\hat{\mathbf{F}}, \boldsymbol{\mu}]_{0+}, \quad [\mathbf{F}, \boldsymbol{\mu}]_{1-} = [\hat{\mathbf{F}}, \boldsymbol{\mu}]_{1-} \quad \textit{for every } \boldsymbol{\mu} \in \Gamma_o^{01}(\mu_0, \mu_1|\mathbf{F}). \tag{7.7.5}$$

(e) *If* $\mu_0 \in \overline{\mathrm{D}(\mathbf{F})}$*,* $\mu_1 \in \mathrm{D}(\mathbf{F})$ *and* $\Gamma_o^1(\mu_0, \mu_1|\mathbf{F}) \neq \emptyset$ *then*

$$\Phi_i \in \hat{\mathbf{F}}[\mu_i] \quad \Rightarrow \quad [\Phi_0, \Phi_1]_r \leqslant \lambda W_2^2(\mu_0, \mu_1).$$

(f) *If*

$$\textit{for every } \mu_0, \mu_1 \in \overline{\mathrm{D}(\mathbf{F})} \textit{ the set } \Gamma_o^{01}(\mu_0, \mu_1|\mathbf{F}) \textit{ is not empty,} \tag{7.7.6}$$

then $\hat{\mathbf{F}}$ *is* λ*-dissipative as well according to* (7.5.1) *and for every* $\mu_0, \mu_1 \in \overline{\mathrm{D}(\mathbf{F})}$ (7.7.5) *holds.*

Proof. Claim (a) is obvious since every λ-dissipative extension $\mathbf{F}'$ of $\mathbf{F}$ in $\overline{\mathrm{D}(\mathbf{F})}$ satisfies $\mathbf{F}' \subset \hat{\mathbf{F}}$.

(b) Let us prove that if $\Phi \in \hat{\mathbf{F}}$ then $\Phi \in \widehat{\mathrm{cl}(\mathbf{F})}$. If $\Psi \in \mathrm{cl}(\mathbf{F})$ we can find a sequence $(\Psi_n)_{n\in\mathbb{N}} \subset \mathbf{F}$ converging to Ψ in $\mathcal{P}_2^{sw}(\mathsf{TH})$ as $n \to \infty$. We can then pass to the limit in the inequalities

$$[\Phi, \nu_n]_r + [\Phi_n, \mu]_r \leqslant \lambda W_2^2(\mu, \nu_n), \quad \mu = x_\sharp\Phi, \ \nu_n = x_\sharp\Psi_n,$$

using the lower semicontinuity results of Lemma 7.3.1. We conclude since $\overline{\mathrm{D}(\mathbf{F})} = \overline{\mathrm{D}(\mathrm{cl}(\mathbf{F}))}$.

In order to prove that $\Phi \in \hat{\mathbf{F}} \ \Rightarrow \ \Phi \in \widehat{\mathrm{co}(\mathbf{F})}$ we take $\Psi = \sum \alpha_k \Psi_k \in \mathrm{co}(\mathbf{F})$; for some $\Psi_k \in \mathbf{F}[\nu]$, $\nu = x_\sharp\Psi \in \mathrm{D}(\mathbf{F})$, and positive coefficients α_k, $k = 1, \cdots, K$, with $\sum_k \alpha_k = 1$. Taking a convex combination of the inequalities

$$[\Phi, \nu]_r + [\Psi_k, \mu]_r \leqslant \lambda W_2^2(\mu, \nu), \quad \text{for every } k = 1, \cdots, K,$$

and using Lemma 7.3.3 we obtain

$$[\Phi, \nu]_r + [\Psi, \mu]_r \leqslant \sum_k \alpha_k \Big([\Phi, \nu]_r + [\Psi_k, \mu]_r \Big) \leqslant \lambda W_2^2(\mu, \nu).$$

The proof of claim (c) follows by a similar argument.

(d) Let $\mu_i \in \mathrm{D}(\mathbf{F})$, $\Phi_i \in \hat{\mathbf{F}}[\mu_i]$, $i = 0, 1$, and $\boldsymbol{\mu} \in \Gamma_o^{01}(\mu_0, \mu_1|\mathbf{F})$. The implication (P1)⇒(P4) of Proposition 7.6.7 applied to $\boldsymbol{\mu}$ and to $s_\sharp\boldsymbol{\mu}$, with s the reversion map in (7.4.2), yields

$$[\Phi_0, \boldsymbol{\mu}]_{r,0} \leqslant [\mathbf{F}, \boldsymbol{\mu}]_{0+}, \quad [\Phi_1, s_\sharp\boldsymbol{\mu}]_{r,0} \leqslant [\mathbf{F}, s_\sharp\boldsymbol{\mu}]_{0+} = -[\mathbf{F}, \boldsymbol{\mu}]_{1-}$$

so that (7.6.13) yields

$$[\Phi_0, \Phi_1]_r \leqslant [\Phi_0, \boldsymbol{\mu}]_{r,0} + [\Phi_1, s_\sharp \boldsymbol{\mu}]_{r,0} \leqslant [\mathbf{F}, \boldsymbol{\mu}]_{0+} - [\mathbf{F}, \boldsymbol{\mu}]_{1-} \leqslant \lambda W_2^2(\mu_0, \mu_1).$$

In order to prove (7.7.5) we observe that $\mathbf{F} \subset \hat{\mathbf{F}}$ so that, for every $\boldsymbol{\mu} \in \Gamma_o^{01}(\mu_0, \mu_1|\mathbf{F})$ and every $t \in I(\boldsymbol{\mu}|\mathbf{F})$, we have $[\mathbf{F}, \boldsymbol{\mu}]_{r,t} \leqslant [\hat{\mathbf{F}}, \boldsymbol{\mu}]_{r,t}$ and $[\mathbf{F}, \boldsymbol{\mu}]_{l,t} \geqslant [\hat{\mathbf{F}}, \boldsymbol{\mu}]_{l,t}$, hence (7.7.5) is a consequence of Definition 7.6.4 and Theorem 7.6.2.

The proof of claim (f) follows by the same argument.

In the case of claim (e), we use the implication (P1)⇒(P6) of Proposition 7.6.7 applied to $\boldsymbol{\mu}$ and the implication (P1)⇒(P3) applied to $s_\sharp \boldsymbol{\mu}$, obtaining

$$[\Phi_0, \boldsymbol{\mu}]_{r,0} \leqslant \lambda W_2^2(\mu_0, \mu_1) + [\mathbf{F}, \boldsymbol{\mu}]_{1-}, \quad [\Phi_1, s_\sharp \boldsymbol{\mu}]_{r,0} \leqslant [\mathbf{F}, s_\sharp \boldsymbol{\mu}]_{0+} = -[\mathbf{F}, \boldsymbol{\mu}]_{1-}$$

and then

$$[\Phi_0, \Phi_1]_r \leqslant [\Phi_0, \boldsymbol{\mu}]_{r,0} + [\Phi_1, s_\sharp \boldsymbol{\mu}]_{r,0} \leqslant \lambda W_2^2(\mu_0, \mu_1). \qquad \square$$

7.8 DISSIPATIVITY IN THE DISCRETE SETTING

We want to show that, in case we restrict our attention to discrete measures, the properties shown in Theorem 7.6.2 holds also for non necessarily optimal couplings. To this aim, let us consider the space of probability measures with finite and with compact support

$$\begin{aligned} \mathcal{P}_f(\mathbb{H}) &:= \left\{ \mu \in \mathcal{P}(\mathbb{H}) : \text{supp}(\mu) \text{ is finite} \right\}, \\ \mathcal{P}_c(\mathbb{H}) &:= \left\{ \mu \in \mathcal{P}(\mathbb{H}) : \text{supp}(\mu) \text{ is compact} \right\} \end{aligned} \tag{7.8.1}$$

and the set

$$\mathcal{P}_N(\mathbb{H}) := \left\{ \mu \in \mathcal{P}_f(\mathbb{H}) : N\mu(A) \in \mathbb{N}\ \forall A \subset \mathbb{H} \right\}, \tag{7.8.2}$$

with $N \in \mathbb{N}$.

If $\mathbf{F}$ is a MPVF, we correspondingly set

$$D_f(\mathbf{F}) := D(\mathbf{F}) \cap \mathcal{P}_f(\mathbb{H}). \tag{7.8.3}$$

For every $\mu_0, \mu_1 \in \mathcal{P}_c(\mathbb{H})$ we define the L^∞-Wasserstein distance by

$$W_\infty(\mu_0, \mu_1) := \min \left\{ \|x^0 - x^1\|_{L^\infty(\mathbb{H}\times\mathbb{H}, \boldsymbol{\mu})} : \boldsymbol{\mu} \in \Gamma(\mu_0, \mu_1) \right\}. \tag{7.8.4}$$

In the following, we investigate the results recalled in Theorem 7.6.2 in the case of marginals μ_0, μ_1 with finite support, but removing the optimality requirement over the coupling $\boldsymbol{\mu}$.

Lemma 7.8.1. *Let $\mathbf{F}$ be a MPVF satisfying (7.5.2). If $\mu_0, \mu_1 \in D_f(\mathbf{F})$ and $\boldsymbol{\mu} \in \Gamma(\mu_0, \mu_1|\mathbf{F})$ is such that x^t is $\boldsymbol{\mu}$-essentially injective for every $t \in (0, 1)$, then*

$$[\mathbf{F}, \boldsymbol{\mu}]_{r,s} - [\mathbf{F}, \boldsymbol{\mu}]_{l,t} \leqslant \lambda(t-s)W^2, \quad W^2 := \int |x_0 - x_1|^2 \, d\boldsymbol{\mu}, \quad \textit{for every } 0 \leqslant s < t \leqslant 1. \tag{7.8.5}$$

In particular, $t \mapsto [\mathbf{F}, \boldsymbol{\mu}]_{r,t} + \lambda W^2 t$ and $t \mapsto [\mathbf{F}, \boldsymbol{\mu}]_{l,t} + \lambda W^2 t$ are increasing respectively in $[0, 1)$ and in $(0, 1]$, $[\mathbf{F}, \boldsymbol{\mu}]_{l,t} = [\mathbf{F}, \boldsymbol{\mu}]_{r,t}$ at every $t \in (0, 1)$ where one of them is continuous, hence they coincide outside a countable set of discontinuities.

Proof. By Theorem 7.6.2 it is not restrictive to assume $\lambda = 0$; we can also assume $s = 0$ and $t = 1$ thanks to (7.4.5). We set $\mu_t := x^t_\sharp \mu$ and we select arbitrarily $\Phi_t \in \mathbf{F}[\mu_t]$.

Applying Theorem 6.5.2, we can find points $t_0 = 0 < t_1 < \cdots < t_K = 1$ such that

$$\mu^k := (x^{t_{k-1}}, x^{t_k})_\sharp \mu \in \Gamma(\mu_{t_{k-1}}, \mu_{t_k} | \mathbf{F}) \cap \Gamma_o(\mu_{t_{k-1}}, \mu_{t_k}) \quad \text{for every } k = 1, \cdots, K.$$

In particular, from (7.4.5) and Theorem 7.6.2(2), we get

$$[\Phi_{t_{k-1}}, \mu]_{r,t_{k-1}} = \frac{1}{t_k - t_{k-1}} [\Phi_{t_{k-1}}, \mu^k]_{r,0} \leqslant \frac{1}{t_k - t_{k-1}} [\Phi_{t_k}, \mu^k]_{l,1} = [\Phi_{t_k}, \mu]_{l,t_k}. \tag{7.8.6}$$

Since x^t is μ-essentially injective, Remark 7.4.2 yields $[\Phi_{t_k}, \mu]_{l,t_k} = [\Phi_{t_k}, \mu]_{r,t_k}$ so that

$$[\Phi_0, \mu]_{r,0} \leqslant [\Phi_1, \mu]_{l,1}.$$

Taking the supremum w.r.t. $\Phi_0 \in \mathbf{F}[\mu_0]$ and the infimum w.r.t. $\Phi_1 \in \mathbf{F}[\mu_1]$ we obtain (7.8.5). The last part of the statement follows as in the proof of Theorem 7.6.2. □

Lemma 7.8.2. *Let* $\mathbf{F}$ *be a MPVF satisfying* (7.5.2), *let* $\mu_0, \mu_1 \in D_f(\mathbf{F})$, $\boldsymbol{\mu} \in \Gamma(\mu_0, \mu_1 | \mathbf{F})$ *(see Definition 7.6.3) and let* $\mu_t = x^t_\sharp \boldsymbol{\mu}$, $t \in [0, 1]$. *Assume that one of the following conditions is satisfied:*

1. *for every* $t \in (0, 1)$ μ_t *belongs to the interior of* $D_f(\mathbf{F})$ *in* $\mathcal{P}_f(\mathbb{H})$ *with respect to the* W_∞*-topology;*
2. *there exists* $N \in \mathbb{N}$ *such that* $\boldsymbol{\mu} \in \mathcal{P}_N(\mathbb{H} \times \mathbb{H})$ *and for every* $t \in (0, 1)$ μ_t *belongs to the interior of* $D(\mathbf{F}) \cap \mathcal{P}_N(\mathbb{H})$ *in* $\mathcal{P}_N(\mathbb{H})$ *with respect to the* W_∞*-topology.*

Then

$$[\mathbf{F}, \boldsymbol{\mu}]_{r,s} - [\mathbf{F}, \boldsymbol{\mu}]_{l,t} \leqslant \lambda(t-s) W^2, \quad W^2 := \int |x_0 - x_1|^2 \, d\boldsymbol{\mu}, \quad \textit{for every } 0 \leqslant s < t \leqslant 1. \tag{7.8.7}$$

Proof. We prove the Lemma only in case 1., being the proof in case 2. analogous. By Theorem 7.6.2 it is not restrictive to assume $\lambda = 0$; we can also assume $s = 0$ and $t = 1$ thanks to (7.4.5). By Theorem 6.5.2 we can find $0 < \delta < 1/2$ and $\tau \in (\delta, 1-\delta)$ s.t. x^δ, x^τ and $x^{1-\delta}$ are $\boldsymbol{\mu}$-essentially injective and $(x^0, x^\delta)_\sharp \boldsymbol{\mu}$, $(x^{1-\delta}, x^1)_\sharp \boldsymbol{\mu}$ are optimal. In this way, since by Theorem 7.6.2 the relation (7.8.7) is true both for the case $s = 0, t = \delta$ and $s = 1-\delta, t = 1$, we only need to prove it for $s = \delta$ and $t = 1 - \delta$.
We set $A = \text{supp}(\mu_\delta) \cup \text{supp}(\mu_{1-\delta})$ and $B = \text{supp}(\mu_\tau)$. By compactness, we can find $\varepsilon > 0$ such that every measure with finite support in the W_∞-neighborhood of radious $\varepsilon > 0$ around μ_t is contained in $D(\mathbf{F})$ for every $\delta \leqslant t \leqslant 1-\delta$.
Applying Proposition 6.4.3 we can find a map $\mathbf{b} : B \to \mathbb{H}$ with values in the open ball of radious ε centered at 0 such that setting $\mathbf{b}^s(x) := x + s\mathbf{b}(x)$ for every

$s \in [0,1]$ and $x \in B$, the set $B^s := \mathbf{b}^s(B)$ satisfies $(B^s - B^s) \cap \mathrm{dir}(A) = \{0\}$ and $\# B^s = \# \mathrm{supp}(\mu_\tau)$ for every $s \in (0,1]$. Considering the measures $\nu_s := \mathbf{b}^s_\sharp(\mu_\tau)$ we can pick $\Psi_s \in \mathbf{F}[\nu_s]$ with barycenter $\boldsymbol{\nu}_s : B^s \to \mathbb{H}$.
Now for every $(x_0, x_1) \in \mathrm{supp}((x^\delta, x^{1-\delta})_\sharp \boldsymbol{\mu})$ we set

$$x_a := x^a(x_0, x_1), \quad \mathbf{b}^{s,\tau} := \mathbf{b}^s(x_a), \quad \boldsymbol{\nu}^{s,\tau} := \boldsymbol{\nu}^s(\mathbf{b}^{s,\tau}),$$

where $a = \frac{\tau - \delta}{1-2\delta}$. Let us consider $\Phi_\delta \in \mathbf{F}[\mu_\delta]$, $\Phi_{1-\delta} \in \mathbf{F}[\mu_{1-\delta}]$ and $\boldsymbol{\sigma} \in \mathcal{P}(\mathsf{T}\mathbb{H} \times \mathsf{T}\mathbb{H})$ s.t. $(x^0, x^1)_\sharp \boldsymbol{\sigma} = (x^\delta, x^{1-\delta})_\sharp \boldsymbol{\mu}$, $(x^0, v^0)_\sharp \boldsymbol{\sigma} = \Phi_\delta$ and $(x^1, v^1)_\sharp \boldsymbol{\sigma} = \Phi_{1-\delta}$. For every $(x_0, v_0, x_1, v_1) \in \mathrm{supp}(\boldsymbol{\sigma})$ we have

$$\begin{aligned}
\langle v_0 - v_1, x_0 - x_1 \rangle &= \langle v_0 - \boldsymbol{\nu}^{s,\tau}, x_0 - x_1 \rangle + \langle v_1 - \boldsymbol{\nu}^{s,\tau}, x_1 - x_0 \rangle \\
&= \frac{1}{a} \langle v_0 - \boldsymbol{\nu}^{s,\tau}, x_0 - x_a \rangle + \frac{1}{1-a} \langle v_1 - \boldsymbol{\nu}^{s,\tau}, x_1 - x_a \rangle \\
&= \frac{1}{a} \langle v_0 - \boldsymbol{\nu}^{s,\tau}, x_0 - \mathbf{b}^{s,\tau} \rangle + \frac{1}{1-a} \langle v_1 - \boldsymbol{\nu}^{s,\tau}, x_1 - \mathbf{b}^{s,\tau} \rangle \\
&\quad + \frac{1}{a} \langle v_0 - \boldsymbol{\nu}^{s,\tau}, \mathbf{b}^{s,\tau} - x_a \rangle + \frac{1}{1-a} \langle v_1 - \boldsymbol{\nu}^{s,\tau}, \mathbf{b}^{s,\tau} - x_a \rangle \\
&= \frac{1}{a} \langle v_0 - \boldsymbol{\nu}^{s,\tau}, x_0 - \mathbf{b}^{s,\tau} \rangle + \frac{1}{1-a} \langle v_1 - \boldsymbol{\nu}^{s,\tau}, x_1 - \mathbf{b}^{s,\tau} \rangle \\
&\quad + \frac{1}{a(1-a)} \langle \boldsymbol{\nu}^{1,\tau} - \boldsymbol{\nu}^{s,\tau}, \mathbf{b}^{s,\tau} - x_a \rangle \\
&\quad + \frac{1}{a(1-a)} \langle (1-a) v_0 + a v_1 - \boldsymbol{\nu}^{1,\tau}, \mathbf{b}^{s,\tau} - x_a \rangle \\
&= \frac{1}{a} \langle v_0 - \boldsymbol{\nu}^{s,\tau}, x_0 - \mathbf{b}^{s,\tau} \rangle + \frac{1}{1-a} \langle v_1 - \boldsymbol{\nu}^{s,\tau}, x_1 - \mathbf{b}^{s,\tau} \rangle \\
&\quad + \frac{s}{(1-s) a (1-a)} \langle \boldsymbol{\nu}^{1,\tau} - \boldsymbol{\nu}^{s,\tau}, \mathbf{b}^{1,\tau} - \mathbf{b}^{s,\tau} \rangle \\
&\quad + \frac{s}{a(1-a)} \langle (1-a) v_0 + a v_1 - \boldsymbol{\nu}^{1,\tau}, \mathbf{b}^{1,\tau} - x_a \rangle.
\end{aligned}$$

We have that

$$\begin{aligned}
\int \langle v_0 - \boldsymbol{\nu}^{s,\tau}(x_0, x_1), x_0 - \mathbf{b}^{s,\tau}(x_0, x_1) \rangle \, d\boldsymbol{\sigma} &= [\Phi_\delta, \boldsymbol{\mu}^{s,\tau}]_{r,0} - [\Psi_s, \boldsymbol{\mu}^{s,\tau}]_{l,1}, \\
\int \langle v_1 - \boldsymbol{\nu}^{s,\tau}(x_0, x_1), x_1 - \mathbf{b}^{s,\tau}(x_0, x_1) \rangle \, d\boldsymbol{\sigma} &= [\Phi_{1-\delta}, \tilde{\boldsymbol{\mu}}^{s,\tau}]_{r,0} - [\Psi_s, \tilde{\boldsymbol{\mu}}^{s,\tau}]_{l,1}
\end{aligned}$$

and

$$\begin{aligned}
\int \langle \boldsymbol{\nu}^{1,\tau}(x_0, x_1) - \boldsymbol{\nu}^{s,\tau}(x_0, x_1), \mathbf{b}^{1,\tau}(x_0, x_1) - \mathbf{b}^{s,\tau}(x_0, x_1) \rangle \, d\boldsymbol{\sigma} \\
= [\Psi_1, \boldsymbol{\vartheta}^{s,\tau}]_{r,0} - [\Psi_s, \boldsymbol{\vartheta}^{s,\tau}]_{l,1},
\end{aligned}$$

where $\boldsymbol{\mu}^{s,\tau} = (x^0, \mathbf{b}^{s,\tau})_\sharp \boldsymbol{\sigma}$, $\tilde{\boldsymbol{\mu}}^{s,\tau} = (x^1, \mathbf{b}^{s,\tau})_\sharp \boldsymbol{\sigma}$, $\boldsymbol{\vartheta}^{s,\tau} = (\mathbf{b}^{1,\tau}, \mathbf{b}^{s,\tau}) \sharp \boldsymbol{\sigma}$ and the equalities with the pseudo scalar products come from the fact that all those plans are concentrated on a map w.r.t. their first marginal (here we are using the $\boldsymbol{\mu}$-essential injectivity of x^δ, x^τ, $x^{1-\delta}$ and the fact that the cardinality of B^s is constant w.r.t. s). By construction, these plans satisfy the hypotheses of Lemma 7.8.1 so that we end up with

$$\int \langle v_0 - v_1, x_0 - x_1 \rangle \, d\boldsymbol{\sigma} \leqslant \frac{s}{a(1-a)} \int \langle (1-a)v_0 + av_1 - \boldsymbol{v}^{1,\tau}, \boldsymbol{b}^{1,\tau} - x_a \rangle \, d\boldsymbol{\sigma}.$$

Passing to the limit as $s \downarrow 0$ we obtain

$$\int \langle v_0 - v_1, x_0 - x_1 \rangle \, d\boldsymbol{\sigma} \leqslant 0.$$

Passing to the supremum w.r.t. $\Phi_\delta \in \mathbf{F}[\mu_\delta]$ and to the infimum w.r.t. $\Phi_{1-\delta} \in \mathbf{F}[\mu_{1-\delta}]$, we get

$$[\mathbf{F}, (x^\delta, x^{1-\delta})_\sharp \boldsymbol{\mu}]_{r,0} - [\mathbf{F}, (x^\delta, x^{1-\delta})_\sharp \boldsymbol{\mu}]_{l,1} \leqslant 0,$$

which is (7.8.7) with $s = \delta$ and $t = 1 - \delta$ thanks to (7.4.5). □

8 MEASURE DIFFERENTIAL INCLUSIONS AND EVI

In this chapter we treat the notion of EVI evolution, its properties and its relation to a weaker notion of solution, called barycentric property. In particular Section 8.1 is devoted to the definition of EVI solution and its properties; in Section 8.2 we present a few consequences of the definition in terms of properties of solutions to the evolution problem; finally in Section 8.3 we compare the notion of EVI solution with the barycentric one.

This Chapter is the result of a collaboration with Giulia Cavagnari and Giuseppe Savaré and it appeared in [34].

For the whole chapter, $\mathbb{H}$ denotes a separable Hilbert space.

8.1 METRIC CHARACTERIZATION

In this section we study a suitable notion of solution to the (formal) problem

$$\dot{\mu}_t \in \mathbf{F}[\mu_t], \qquad t \in \mathcal{I}, \tag{8.1.1}$$

where $\mathbf{F}$ is a MPVF as in Definition 7.5.1 and $\mathcal{I}$ is a connected subset of $\mathbb{R}$.

Reasoning in analogy with the theory of gradient flows in $\mathcal{P}_2(\mathbb{H})$, the naive way to interpret (8.1.1) is to ask for a locally absolutely continuous curve $\mu : \mathcal{I} \to \mathcal{P}_2(\mathbb{H})$ to satisfy

$$(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t \in \mathbf{F}[\mu_t] \quad \text{for a.e. } t \in \mathcal{I}, \tag{8.1.2}$$

where $\boldsymbol{v}$ is the velocity vector of μ (see Theorem 2.4.6).

However, there is no reason why a given $\mathbf{F}[\mu_t]$ should contain vectors of the tangent space $\mathrm{Tan}_{\mu_t} \mathcal{P}_2(\mathbb{H})$. We thus introduce a different notion of solution to (8.1.2), inspired by the EVI formulation for gradient flows, and we will eventually obtain the inclusion (8.1.2) for the extended MPVF $\hat{\mathbf{F}}$ introduced in (7.7.4).

It is not difficult to see that, if $\mathbf{F}$ is λ-dissipative according to (7.5.1), also using Theorem 7.2.1 and Remark 7.5.5, every locally absolutely continuous curve satisfying (8.1.2) also satisfies the Evolution Variational Inequality (λ-EVI)

$$\frac{1}{2}\frac{d}{dt}W_2^2(\mu_t, \nu) \leqslant \lambda W_2^2(\mu_t, \nu) - [\Phi, \mu_t]_r \quad \text{in } \mathscr{D}'\big(\operatorname{int}(\mathcal{I})\big), \tag{λ-EVI}$$

for every $\nu \in D(\mathbf{F})$ and every $\Phi \in \mathbf{F}[\nu]$, where $[\cdot, \cdot]_r$ is the functional pairing in Definition 7.1.4 and the writing $\mathscr{D}'\big(\operatorname{int}(\mathcal{I})\big)$ means that the expression has to be understood in the distributional sense over $\operatorname{int}(\mathcal{I})$. Starting from this heuristic remark, we give the following definition.

Definition 8.1.1 (λ-EVI solution)**.** Let $\mathbf{F}$ be a MPVF and let $\lambda \in \mathbb{R}$. We say that a continuous curve $\mu : \mathfrak{I} \to \overline{D(\mathbf{F})}$ is a λ-*EVI solution* to (8.1.1) for the MPVF $\mathbf{F}$ if (λ-EVI) holds for every $\nu \in D(\mathbf{F})$ and every $\Phi \in \mathbf{F}[\nu]$.
A λ-EVI solution μ is said to be a *strict solution* if $\mu_t \in D(\mathbf{F})$ for every $t \in \mathfrak{I}$, $t > \inf \mathfrak{I}$.
A λ-EVI solution μ is said to be a *global solution* if $\sup \mathfrak{I} = +\infty$.

See Example 9.5.5 for a justification of the mere continuity assumption on μ.

We recall that, given a function $\zeta : \mathfrak{I} \to \mathbb{R}$, the right upper and lower Dini derivatives of ζ at a point $t \in \mathfrak{I}$, $t < \sup \mathfrak{I}$ are defined as

$$\frac{d^+}{dt}\zeta(t) := \limsup_{h\downarrow 0} \frac{\zeta(t+h)-\zeta(t)}{h}, \qquad \frac{d}{dt_+}\zeta(t) := \liminf_{h\downarrow 0} \frac{\zeta(t+h)-\zeta(t)}{h}. \tag{8.1.3}$$

Remark 8.1.2. Arguing as in [85, Lemma A.1] and using the lower semicontinuity of the map $t \mapsto [\Phi, \mu_t]_r$, the distributional inequality of (λ-EVI) can be equivalently reformulated in terms of the right upper or lower Dini derivatives of the squared distance function and requiring the condition to hold for every $t \in \operatorname{int}(\mathfrak{I})$:

$$\frac{1}{2}\frac{d^+}{dt} W_2^2(\mu_t, x_\sharp\Phi) \leqslant \lambda W_2^2(\mu_t, x_\sharp\Phi) - [\Phi, \mu_t]_r \quad \text{for every } t \in \operatorname{int}(\mathfrak{I}),\ \Phi \in \mathbf{F} \tag{λ-EVI$_1$}$$

$$\frac{1}{2}\frac{d}{dt_+} W_2^2(\mu_t, x_\sharp\Phi) \leqslant \lambda W_2^2(\mu_t, x_\sharp\Phi) - [\Phi, \mu_t]_r \quad \text{for every } t \in \operatorname{int}(\mathfrak{I}),\ \Phi \in \mathbf{F}. \tag{λ-EVI$_2$}$$

A further equivalent formulation [85, Theorem 3.3] involves the difference quotients: for every $s, t \in \mathfrak{I}$, $s < t$

$$\begin{aligned} &e^{-2\lambda(t-s)} W_2^2(\mu_t, x_\sharp\Phi) - W_2^2(\mu_s, x_\sharp\Phi) \\ &\qquad \leqslant -2\int_s^t e^{-2\lambda(r-s)} [\Phi, \mu_r]_r \, dr \quad \text{for every } \Phi \in \mathbf{F}. \end{aligned} \tag{λ-EVI$_3$}$$

Finally, if μ is also locally absolutely continuous, then (λ-EVI$_1$) and (λ-EVI$_2$) are also equivalent to

$$\frac{1}{2}\frac{d}{dt} W_2^2(\mu_t, x_\sharp\Phi) \leqslant \lambda W_2^2(\mu_t, x_\sharp\Phi) - [\Phi, \mu_t]_r \qquad \text{for a.e. } t \in \mathfrak{I} \text{ and every } \Phi \in \mathbf{F}.$$

The following lemma discusses further properties of λ-EVI solutions. We refer respectively to (7.6.2), (7.6.7) and Definition 7.6.4 for the definitions of $I(\boldsymbol{\mu}|\mathbf{F})$, $\Gamma_o^i(\cdot, \cdot|\mathbf{F})$, with $i = 0, 1$, and for the definitions of $[\mathbf{F}, \boldsymbol{\mu}]_{0+}$ and $[\mathbf{F}, \boldsymbol{\mu}]_{1-}$.

Lemma 8.1.3. *Let $\mathbf{F}$ be a λ-dissipative MPVF according to (7.5.1) and let $\mu : \mathfrak{I} \to \overline{D(\mathbf{F})}$ be a continuous λ-EVI solution to (8.1.1). We have*

$$\begin{gathered} \frac{1}{2}\frac{d^+}{dt} W_2^2(\mu_t, \nu) \leqslant [\mathbf{F}, \boldsymbol{\mu}_t]_{0+} \\ \textit{for every } \nu \in \overline{D(\mathbf{F})},\ t \in \operatorname{int}(\mathfrak{I}),\ \boldsymbol{\mu}_t \in \Gamma_o^0(\mu_t, \nu|\mathbf{F}), \end{gathered} \tag{8.1.4a}$$

$$\begin{gathered} \frac{1}{2}\frac{d^+}{dt} W_2^2(\mu_t, \nu) \leqslant \lambda W_2^2(\mu_t, \nu) + [\mathbf{F}, \boldsymbol{\mu}_t]_{1-} \\ \textit{for every } \nu \in \overline{D(\mathbf{F})},\ t \in \operatorname{int}(\mathfrak{I}),\ \boldsymbol{\mu}_t \in \Gamma_o^1(\mu_t, \nu|\mathbf{F}). \end{gathered} \tag{8.1.4b}$$

If moreover μ is locally absolutely continuous with Wasserstein velocity field $\boldsymbol{\nu}$ satisfying (2.4.8) *for every* t *in the subset* $A(\mu) \subset \mathcal{I}$ *of full Lebesgue measure, then*

$$\begin{gathered}\left[(\mathbf{i}_{\mathbb{H}}, \boldsymbol{\nu}_t)_\sharp \mu_t, \nu\right]_r \leqslant \lambda W_2^2(\mu_t, \nu) - [\Phi, \mu_t]_r \\ \textit{for every } t \in A(\mu)\ \Phi \in \mathbf{F},\ \nu = x_\sharp \Phi,\end{gathered} \tag{8.1.5a}$$

$$\begin{gathered}[(\mathbf{i}_{\mathbb{H}}, \boldsymbol{\nu}_t)_\sharp \mu_t, \mu_t]_{r,0} \leqslant [\mathbf{F}, \mu_t]_{0+} \\ \textit{for every } t \in A(\mu),\ \nu \in \overline{D(\mathbf{F})},\ \boldsymbol{\mu}_t \in \Gamma_o^0(\mu_t, \nu|\mathbf{F}),\end{gathered} \tag{8.1.5b}$$

$$\begin{gathered}[(\mathbf{i}_{\mathbb{H}}, \boldsymbol{\nu}_t)_\sharp \mu_t, \mu_t]_{r,0} \leqslant \lambda W_2^2(\mu_t, \nu) + [\mathbf{F}, \mu_t]_{1-} \\ \textit{for every } t \in A(\mu),\ \nu \in \overline{D(\mathbf{F})},\ \boldsymbol{\mu}_t \in \Gamma_o^1(\mu_t, \nu|\mathbf{F}).\end{gathered} \tag{8.1.5c}$$

Proof. In order to check (8.1.5a) it is sufficient to combine (7.2.1) of Theorem 7.2.1 with (λ-EVI$_1$). (8.1.5b) and (8.1.5c) then follow applying Proposition 7.6.7. Let us now prove (8.1.4a): fix $\nu \in \overline{D(\mathbf{F})}$ and $t \in \operatorname{int}(\mathcal{I})$. Take $\boldsymbol{\mu}_t \in \Gamma_o(\mu_t, \nu)$ and define the constant speed geodesic $\nu^t : [0,1] \to \mathcal{P}_2(\mathbb{H})$ by $\nu^t_s := (x^s)_\sharp \boldsymbol{\mu}_t$, thus in particular $\nu^t_0 = \mu_t$ and $\nu^t_1 = \nu$. Then by Lemma 7.2.6, for every $s \in I(\boldsymbol{\mu}|\mathbf{F}) \cap (0,1)$ and $\Phi_s \in \mathbf{F}(\nu^t_s)$ we have

$$\begin{aligned}\frac{1}{2}\frac{d^+}{dt} W_2^2(\mu_t, \nu) &\leqslant \frac{1}{2s}\frac{d^+}{dt} W_2^2(\mu_t, \nu^t_s) \\ &\leqslant -\frac{1}{s}[\Phi_s, \boldsymbol{\mu}_t]_r + \frac{\lambda}{s} W_2^2(\mu_t, \nu^t_s) \\ &\leqslant [\mathbf{F}, \boldsymbol{\mu}_t]_{r,s} + \lambda s W_2^2(\mu_t, \nu),\end{aligned}$$

where the second inequality comes from (λ-EVI$_1$). Taking $\boldsymbol{\mu}_t \in \Gamma_o^0(\mu_t, \nu|\mathbf{F})$ and passing to the limit as $s \downarrow 0$ we get (8.1.4a). Analogously for (8.1.4b). □

The following result presents the relation between the notion of λ-EVI solution and of differential inclusion (8.1.2).

Theorem 8.1.4. *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1) *and let* $\mu : \mathcal{I} \to \overline{D(\mathbf{F})}$ *be a locally absolutely continuous curve.*

1. *If* μ *satisfies the differential inclusion* (8.1.2) *driven by any* λ*-dissipative extension of* $\mathbf{F}$ *in* $D(\mathbf{F})$*, then* μ *is also a* λ*-EVI solution to* (8.1.1) *for* $\mathbf{F}$.

2. μ *is a* λ*-EVI solution of* (8.1.1) *for* $\mathbf{F}$ *if and only if*

$$(\mathbf{i}_{\mathbb{H}}, \boldsymbol{\nu}_t)_\sharp \mu_t \in \hat{\mathbf{F}}[\mu_t] \quad \textit{for a.e. } t \in \mathcal{I}. \tag{8.1.6}$$

3. *If* $D(\mathbf{F})$ *satisfies* (7.6.14) *and* $\mu_t \in D(\mathbf{F})$ *for a.e.* $t \in \mathcal{I}$*, then the following properties are equivalent:*
   - μ *is a* λ*-EVI solution to* (8.1.1) *for* $\mathbf{F}$.
   - μ *satisfies* (8.1.5b).
   - μ *is a* λ*-EVI solution to* (8.1.1) *for the restriction of* $\hat{\mathbf{F}}$ *to* $D(\mathbf{F})$.

4. *If* $\mathbf{F}$ *satisfies* (7.7.6) *then* μ *is a* λ*-EVI solution to* (8.1.1) *for* $\mathbf{F}$ *if and only if it is a* λ*-EVI solution to* (8.1.1) *for* $\hat{\mathbf{F}}$.

Proof. (1) It is sufficient to apply Theorem 7.2.1 and the definition of λ-dissipativity.

The left-to-right implication $\Rightarrow$ of (2) follows by (8.1.5a) of Lemma 8.1.3 and the definition of $\hat{\mathbf{F}}$.

Conversely, if μ satisfies (8.1.6), $\nu \in D(\mathbf{F})$, $\Phi \in \mathbf{F}[\nu]$, then Theorem 7.2.1 and the definition of $\hat{\mathbf{F}}$ yield

$$\frac{1}{2}\frac{d}{dt}W_2^2(\mu_t,\nu) = \left[(\boldsymbol{i}_{\mathbb{H}},\boldsymbol{v}_t)_\sharp\mu_t,\nu\right]_r \leqslant \lambda W_2^2(\mu_t,\nu) - [\Phi,\mu_t]_r \quad \text{a.e. in } \mathcal{I}.$$

Claim (3) is an immediate consequence of Lemma 8.1.3, Proposition 7.7.3(d) and Proposition 7.6.7.

Claim (4) is a consequence of Proposition 7.7.3(f) and the λ-dissipativity of $\hat{\mathbf{F}}$. □

The result stated in Theorem 8.1.4 suggests a compatibility between the notion of EVI solution for a dissipative MPVF and the notion of gradient flow for a convex functional in $\mathcal{P}_2(\mathbb{H})$. This correspondence is analysed in Subsection 9.5.1, where we consider the particular case where the MPVF is the opposite of the Fréchet subdifferential of a proper, lower semicontinuous and convex functional $\mathcal{F} : \mathcal{P}_2(\mathbb{H}) \to (-\infty,+\infty]$ (see Proposition 9.5.2).

We derive a further useful a priori bound for λ-EVI solutions.

Proposition 8.1.5. *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1) *and let* $T \in (0,+\infty]$*. Every* λ*-EVI solution* $\mu : [0,T) \to \overline{D(\mathbf{F})}$ *with initial datum* $\mu_0 \in D(\mathbf{F})$ *satisfies the a priori bound*

$$W_2(\mu_t,\mu_0) \leqslant 2|\mathbf{F}|_2(\mu_0)\int_0^t e^{\lambda s}\,ds \tag{8.1.7}$$

for all $t \in [0,T)$*, where*

$$|\mathbf{F}|_2(\mu) := \inf\left\{|\Phi|_2 : \Phi \in \mathbf{F}[\mu]\right\}$$

for every $\mu \in D(\mathbf{F})$*.*

Proof. Let $\Phi \in \mathbf{F}(\mu_0)$. Then ($\lambda$-EVI) with $\nu := \mu_0$ yields

$$\frac{d^+}{dt}W_2^2(\mu_t,\mu_0) - 2\lambda W_2^2(\mu_t,\mu_0) \leqslant -2\,[\Phi,\mu_t]_r \leqslant 2|\Phi|_2\,W_2(\mu_t,\mu_0)$$

for every $t \in [0,T)$. We can then apply the estimate of Lemma [5, Lemma 4.1.8] to obtain

$$e^{-\lambda t}W_2(\mu_t,\mu_0) \leqslant 2|\Phi|_2\int_0^t e^{-\lambda s}\,ds$$

for all $t \in [0,T)$, which in turn yields (8.1.7). □

We conclude this section with a stability result w.r.t. uniform convergence.

Proposition 8.1.6. *If* $\mu_n : \mathcal{I} \to \overline{D(\mathbf{F})}$ *is a sequence of* λ*-EVI solutions locally uniformly converging to* μ *as* $n \to \infty$*, then* μ *is a* λ*-EVI solution.*

Proof. μ is a continuous curve defined in $\mathcal{I}$ with values in $\overline{D(\mathbf{F})}$. Using pointwise convergence, the lower semicontinuity of $\mu \mapsto [\Phi,\mu]_r$ of Lemma 7.3.1, and Fatou's Lemma, it is easy to pass to the limit in the equivalent characterization (λ-EVI$_3$) of λ-EVI solutions, written for μ_n. □

8.2 STABILITY AND UNIQUENESS

The following results provide stability properties of λ-EVI. In the first theorem we assume absolute continuity.

Theorem 8.2.1 (Stability for absolutely continuous solutions). *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1) *and let* $\mu^1, \mu^2 : [0,T) \to \overline{D(\mathbf{F})}$, $T \in (0,+\infty]$, *be* locally absolutely continuous λ*-EVI solutions to* (8.1.1)*. If* $\Gamma_o^0(\mu_t^1, \mu_t^2|\mathbf{F}) \neq \emptyset$ *for a.e.* $t \in (0,T)$*, then*

$$W_2(\mu_t^1, \mu_t^2) \leqslant W_2(\mu_0^1, \mu_0^2)\, e^{\lambda t} \quad \textit{for every } t \in [0,T). \tag{8.2.1}$$

In particular, if $\mu_0^1 = \mu_0^2$ *then* $\mu^1 \equiv \mu^2$ *in* $[0,T)$.

Proof. Since μ^1, μ^2 are locally absolutely continuous curves, we can apply Theorem 7.2.3 and find a subset $A \subset A(\mu^1) \cap A(\mu^2)$ of full Lebesgue measure such that (7.2.2) holds and $\Gamma_o^0(\mu_t^1, \mu_t^2|\mathbf{F}) \neq \emptyset$ for every $t \in A$. Selecting $\boldsymbol{\mu}_t \in \Gamma_o^0(\mu_t^1, \mu_t^2|\mathbf{F})$, we have

$$\frac{1}{2}\frac{d}{dt}W_2^2(\mu_t^1, \mu_t^2) = \int \langle \boldsymbol{v}_t^1(x_1), x_1 - x_2\rangle \, d\boldsymbol{\mu}_t(x_1,x_2) + \int \langle \boldsymbol{v}_t^2(x_2), x_2 - x_1\rangle \, d\boldsymbol{\mu}_t(x_1,x_2).$$

Note that

$$\begin{aligned} \Gamma_0\left((\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t^1)_\sharp \mu_t^1, \boldsymbol{\mu}_t\right) &= \Lambda\left((\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t^1)_\sharp \mu_t^1, \mu_t^2\right) = \left\{(x^0, \boldsymbol{v}_t^1 \circ x^0, x^1)_\sharp \boldsymbol{\mu}_t\right\}, \\ \Gamma_0\left((\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t^2)_\sharp \mu_t^2, s_\sharp \boldsymbol{\mu}_t\right) &= \Lambda\left((\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t^2)_\sharp \mu_t^2, \mu_t^1\right) = \left\{(x^1, \boldsymbol{v}_t^2 \circ x^1, x^0)_\sharp \boldsymbol{\mu}_t\right\} \end{aligned}$$

by [5, Lemma 5.3.2], where $\Gamma_0(\cdot,\cdot)$ is the set defined in (7.4.1) with $t = 0$ and $\Lambda(\cdot,\cdot)$ is defined in Definition 7.1.7. Hence, using (8.1.5b), (8.1.5c) and recalling the definition of reversion map s in (7.4.2), for every $t \in A$ we get

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}W_2^2(\mu_t^1, \mu_t^2) &= [(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t^1)_\sharp \mu_t^1, \boldsymbol{\mu}_t]_{r,0} + [(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t^2)_\sharp \mu_t^2, s_\sharp \boldsymbol{\mu}_t]_{r,0} \\ &\leqslant [\mathbf{F}, \boldsymbol{\mu}_t]_{0+} + \lambda W_2^2(\mu_t^1, \mu_t^2) + [\mathbf{F}, s_\sharp \boldsymbol{\mu}_t]_{1-} \\ &= \lambda W_2^2(\mu_t^1, \mu_t^2), \end{aligned}$$

where we also used the property

$$[\mathbf{F}, s_\sharp \boldsymbol{\mu}_t]_{1-} = -[\mathbf{F}, \boldsymbol{\mu}_t]_{0+}.$$ □

The next Theorem considers the situation when one curve is absolutely continuous and the other merely continuous. The argument presents some technicalities and comes by [89, Theorem 1.1]. Before stating and proving it, we present a simple lemma that allows us to pass from a differential inequality for the right upper Dini derivative to the corresponding distributional inequality (see also [85, Lemma A.1] and [51]).

Lemma 8.2.2. *Let* $(a,b) \subset \mathbb{R}$ *be an open interval (bounded or unbounded) and let* $\zeta, \eta : (a,b) \to \mathbb{R}$ *be s.t.* ζ *is continuous in* (a,b) *and* η *is measurable and locally bounded from above in* (a,b)*. If*

$$\frac{d^+}{dt}\zeta(t) \leqslant \eta(t)$$

for every $t \in (a, b)$*, then the above inequality holds also in the sense of distributions, meaning that*

$$-\int_a^b \zeta(t)\varphi'(t)\,dt \leqslant \int_a^b \eta(t)\varphi(t)\,dt$$

for every $\varphi \in C_c^\infty(a, b)$ *with* $\varphi \geqslant 0$*.*

Proof. Let $\varphi \in C_c^\infty(a, b)$ with $\varphi \geqslant 0$, then there exist $a < x < y < b$ s.t. the support of φ is contained in $[x, y]$; since η is locally bounded from above, there exists a positive constant $C > 0$ s.t. $\eta(t) \leqslant C$ for every $t \in [x, y]$. Then the function $t \mapsto \zeta(t) - Ct$ is such that

$$\frac{d^+}{dt}(\zeta(t) - Ct) \leqslant 0$$

for every $t \in [x, y]$, so that it is decreasing in $[x, y]$ and hence a function of bounded variation in $[x, y]$. Its distributional derivative is hence a non positive measure T on $[x, y]$ whose absolutely continuous part (w.r.t. the 1-dimensional Lebesgue measure on $[x, y]$) coincides a.e. with the right upper Dini derivative. Then we have

$$\begin{aligned} -\int_a^b (\zeta(t) - Ct)\varphi'(t)\,dt = T(\varphi) &= \int_a^b \frac{d^+}{dt}(\zeta(t) - Ct)\varphi(t)\,dt + T_s(\varphi) \\ &\leqslant \int_a^b (\eta - C)\varphi(t)\,dt, \end{aligned}$$

where T_s is the singular part of T. This immediately gives the thesis. □

Theorem 8.2.3 (Refined stability)**.** *Let* $T > 0$ *and* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1)*. Let*

(i) $\mu^1 : [0, T] \to \overline{D(\mathbf{F})}$ *be an absolutely continuous* λ*-EVI solution for* $\mathbf{F}$*, with* $\mu_0^1 \in D(\mathbf{F})$*;*

(ii) $\mu^2 : [0, T] \to \overline{D(\mathbf{F})}$ *be* λ*-EVI solution for* $\mathbf{F}$*.*

If at least one of the following properties hold:

1. $\Gamma_o^0(\mu_r^1, \mu_s^2 | \mathbf{F}) \neq \emptyset$ *for every* $s \in (0, T)$ *and* $r \in [0, T) \setminus N$ *with* $N \subset (0, T)$*,* $\mathcal{L}(N) = 0$*;*

2. μ^1 *satisfies* (8.1.2)*,*

then

$$W_2(\mu_t^1, \mu_t^2) \leqslant e^{\lambda t} W_2(\mu_0^1, \mu_0^2) \quad \textit{for every } t \in [0, T].$$

Proof. We extend μ^1 in $(-\infty, 0)$ with the constant value μ^1_0, denote by $\boldsymbol{\nu}$ the Wasserstein velocity field associated to μ^1 (and extended to 0 outside $A(\mu^1)$) and define the functions $w, f, h : (-\infty, T] \times [0, T] \to \mathbb{R}$ by

$$w(r,s) := W_2(\mu^1_r, \mu^2_s),$$
$$f(r,s) := \begin{cases} 2|\mathbf{F}|_2(\mu^1_0) w(0,s) & \text{if } r < 0, \\ 0 & \text{if } r \geqslant 0, \end{cases}$$
$$h(r,s) := \begin{cases} 0 & \text{if } r < 0, \\ 2\left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{\nu}_r)_\sharp \mu^1_r, \mu^2_s\right]_r & \text{if } r \geqslant 0. \end{cases}$$

Theorem 7.2.1 yields

$$\frac{\partial}{\partial r} w^2(r,s) = h(r,s) \quad \text{in } \mathscr{D}'(-\infty, T), \text{ for every } s \in [0,T]. \tag{8.2.2}$$

In case (1) holds, writing (8.1.4b) for μ^2 with $\nu = \mu^1_r$ and $r \in (-\infty, T] \setminus N$, then for every $\boldsymbol{\mu}_{rs} \in \Gamma^0_o(\mu^1_r, \mu^2_s | \mathbf{F})$ we obtain

$$\frac{d^+}{ds} w^2(r,s) \leqslant 2\lambda w^2(r,s) - 2[\mathbf{F}, \boldsymbol{\mu}_{rs}]_{0+} \quad \text{for } s \in (0,T) \text{ and } r \in (-\infty, T) \setminus N. \tag{8.2.3}$$

On the other hand (8.1.5b) yields

$$\begin{aligned} -2[\mathbf{F}, \boldsymbol{\mu}_{rs}]_{0+} &\leqslant -2[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{\nu}_r)_\sharp \mu^1_r, \boldsymbol{\mu}_{rs}]_{r,0} \leqslant -2\left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{\nu}_r)_\sharp \mu^1_r, \mu^2_s\right]_r \quad \text{for } r \in A(\mu^1) \setminus N, \\ -2[\mathbf{F}, \boldsymbol{\mu}_{rs}]_{0+} &\leqslant 2|\mathbf{F}|_2(\mu^1_0) w(0,s) = f(r,s) \quad \text{for every } r < 0. \end{aligned} \tag{8.2.4}$$

Combining (8.2.3) and (8.2.4) we obtain

$$\frac{d^+}{ds} w^2(r,s) \leqslant 2\lambda w^2(r,s) + f(r,s) - h(r,s) \quad \text{for } s \in (0,T),\ r \in (-\infty, 0] \cup A(\mu^1) \setminus N.$$

Since, recalling Theorem 2.4.6, we have $|h(r,s)| \leqslant 2|\dot{\mu}^1_r| \, w(r,s)$, then applying Lemma 8.2.2 we get

$$\frac{\partial}{\partial s} w^2(r,s) \leqslant 2\lambda w^2(r,s) + f(r,s) - h(r,s) \quad \text{in } \mathscr{D}'(0,T), \text{ for a.e. } r \in (-\infty, T]. \tag{8.2.5}$$

The expression in (8.2.5) can also be deduced in case (2) using (8.1.2).

By multiplying both inequalities (8.2.2) and (8.2.5) by $e^{-2\lambda s}$ we get

$$\begin{aligned} &\frac{\partial}{\partial r}\Big(e^{-2\lambda s} w^2(r,s)\Big) = e^{-2\lambda s} h(r,s) \\ &\quad \text{in } \mathscr{D}'(-\infty, T) \text{ and every } s \in [0,T], \\ &\frac{\partial}{\partial s}\Big(e^{-2\lambda s} w^2(r,s)\Big) \leqslant e^{-2\lambda s}\big(f(r,s) - h(r,s)\big) \\ &\quad \text{in } \mathscr{D}'(0,T) \text{ and a.e. } r \in (-\infty, T]. \end{aligned}$$

We fix $t \in [0,T]$ and $\varepsilon > 0$ and we apply the Divergence theorem in [89, Lemma 6.15] on the two-dimensional strip $Q^{\varepsilon}_{0,t}$ as in Figure 2,

$$Q^{\varepsilon}_{0,t} := \{(r,s) \in \mathbb{R}^2 \mid 0 \leqslant s \leqslant t\,,\; s-\varepsilon \leqslant r \leqslant s\}, \tag{8.2.6}$$

and we get

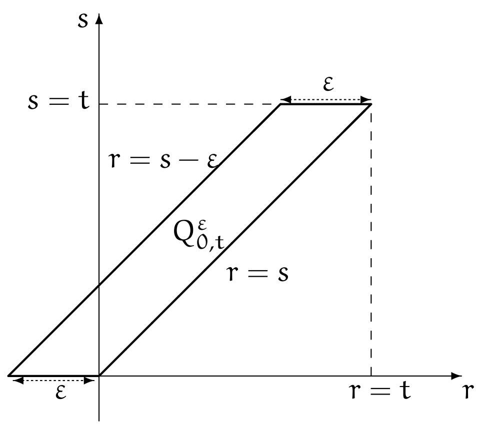

Figure 2: Strip $Q^{\varepsilon}_{0,t}$ corresponding to penalization about the diagonal $\{r = s\}$.

$$\int_{t-\varepsilon}^{t} e^{-2\lambda t} w^2(r,t)\,dr \leqslant \int_{-\varepsilon}^{0} w^2(r,0)\,dr + \iint_{Q^{\varepsilon}_{0,t}} e^{-2\lambda s} f(r,s)\,dr\,ds.$$

Using

$$w(t,t) \leqslant \int_r^t |\dot{\mu}^1_u|\,du + w(r,t) \leqslant \int_{t-\varepsilon}^t |\dot{\mu}^1_u|\,du + w(r,t) \quad \text{if } t-\varepsilon \leqslant r \leqslant t,$$

then, for every $\delta, \delta_\star > 1$ conjugate coefficients ($\delta_\star = \delta/(\delta-1)$), we get

$$w^2(t,t) \leqslant \delta w^2(r,t) + \delta_\star \left(\int_{t-\varepsilon}^t |\dot{\mu}^1_u|\,du \right)^2 . \tag{8.2.7}$$

Integrating (8.2.7) w.r.t. r in the interval $(t-\varepsilon, t)$, we obtain

$$\begin{aligned} e^{-2\lambda t} w^2(t,t) \leqslant{}& \frac{\delta}{\varepsilon} \int_{t-\varepsilon}^t e^{-2\lambda t} w^2(r,t)\,dr \\ &+ \delta_\star \left(\int_{t-\varepsilon}^t |\dot{\mu}^1_u|\,du \right)^2 \max\{1, e^{2|\lambda|T}\}. \end{aligned} \tag{8.2.8}$$

Finally, we have the following inequality

$$\varepsilon^{-1} \iint_{Q^{\varepsilon}_{0,t}} e^{-2\lambda s} f(r,s)\,dr\,ds \leqslant 2|\mathbf{F}|_2(\mu^1_0) \int_0^{\varepsilon} e^{-2\lambda s} w(0,s)\,ds. \tag{8.2.9}$$

Summing up (8.2.8) and (8.2.9) we obtain

$$e^{-2\lambda t} w^2(t) \leqslant \delta \left(w^2(0) + 2|\mathbf{F}|_2(\mu_0^1) \int_0^\varepsilon e^{-2\lambda s} w(0,s)\,ds \right) + \delta_\star \left(\int_{t-\varepsilon}^t |\dot{\mu}_u^1|\,du \right)^2 \max\{1, e^{2|\lambda| T}\},$$

where we have used the notation $w(s) = w(s,s)$. Taking the limit as $\varepsilon \downarrow 0$ and $\delta \downarrow 1$, we obtain the thesis. □

8.3 BARYCENTRIC PROPERTY

If the MPVF $\mathbf{F}$ satisfies additional properties, we are able to show that EVI solutions satisfy the so called *barycentric property* which is strongly related to the notion of evolution treated in [31, 93, 94].

We introduce the following closure of $\mathbf{F}$ along cylindrical functions. We set

$$\exp^{\varphi}(x) := x + \nabla\varphi(x)$$

for every $\varphi \in \mathrm{Cyl}(\mathbb{H})$, and

$$\overline{\mathbf{F}}[\mu] := \left\{ \Phi \in \mathcal{P}_2(\mathbb{H}) \;\middle|\; \begin{array}{l} \exists\, \varphi \in \mathrm{Cyl}(\mathbb{H}),\ (r_n)_{n\in\mathbb{N}} \subset [0,+\infty),\ r_n \downarrow 0, \\ \Phi_n \in \mathbf{F}[\exp^{r_n\varphi}_\sharp \mu] : \Phi_n \to \Phi \text{ in } \mathcal{P}_2^{sw}(\mathsf{T}\mathbb{H}) \end{array} \right\}. \quad (8.3.1)$$

Definition 8.3.1 (Barycentric property)**.** Let $\mathbf{F}$ be a MPVF. We say that a locally absolutely continuous curve $\mu : \mathcal{I} \to D(\mathbf{F})$ satisfies the *barycentric property* (resp. the *relaxed barycentric property*) if for a.e. $t \in \mathcal{I}$ there exists $\Phi_t \in \mathbf{F}[\mu_t]$ (resp. $\Phi_t \in \overline{\mathrm{co}}(\overline{\mathbf{F}}[\mu_t])$) such that

$$\frac{d}{dt}\int_{\mathbb{H}} \varphi(x)\,d\mu_t(x) = \int_{\mathsf{T}\mathbb{H}} \langle \nabla\varphi(x), v\rangle\,d\Phi_t(x,v) \quad \text{for every } \varphi \in \mathrm{Cyl}(\mathbb{H}). \quad (8.3.2)$$

Notice that $\mathbf{F} \subset \overline{\mathbf{F}} \subset \mathrm{cl}(\mathbf{F})$ and $\overline{\mathbf{F}} = \mathbf{F}$ if $\mathbf{F}$ is sequentially closed in $\mathcal{P}_2^{sw}(\mathsf{T}\mathbb{H})$. From Proposition 7.7.3(a) we also get

$$\overline{\mathrm{co}}(\overline{\mathbf{F}}) \subset \hat{\mathbf{F}},$$

so that the relaxed barycentric property implies the corresponding property for the extended MPVF $\hat{\mathbf{F}}$ defined in (7.7.4). In particular, considering the directional closure $\overline{\mathbf{F}}$ in place of the sequential closure $\mathrm{cl}(\mathbf{F})$ not only allows us to obtain a finer result, but it could be easier to compute when one considers specific examples, being $\overline{\mathbf{F}}$ the closure of $\mathbf{F}$ along regular directions.

Remark 8.3.2. If $\mathbb{H} = \mathbb{R}^d$, the property stated in Definition 8.3.1 coincides with the weak definition of solution to (8.1.1) given in [93].

The aim is to prove that the λ-EVI solution of (8.1.1) enjoys the barycentric property of Definition 8.3.1, under suitable mild conditions on $\mathbf{F}$. This is strictly related to the behaviour of $\mathbf{F}$ along the family of smooth deformations induced by cylindrical functions. Let us denote by $\mathbf{pr}_\mu$ the orthogonal projection in $L^2_\mu(\mathbb{H};\mathbb{H})$

onto the tangent space $\mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{H})$ and by $\mathbf{b}_\Phi$ the barycenter of Φ as in Definition 7.1.1.

Before stating the next Theorem, we recall the following characterization of the closed convex hull $\overline{\mathrm{co}}(C)$ of a set C (i.e. the intersection of all the closed convex sets containing C) in a Banach space.

Lemma 8.3.3. *Let Z be a Banach space and let $C \subset Z$ be nonempty. Then $\nu \in \overline{\mathrm{co}}(C)$ if and only if*

$$\langle z^*, \nu\rangle \leqslant \sup_{c\in C} \langle z^*, c\rangle \tag{8.3.3}$$

for all $z^* \in Z^*$. Moreover if C is bounded, it is enough to have (8.3.3) holding for every $z^* \in W$, with W a dense subset of Z^*.

Proof. The result is a direct consequence of Hahn-Banach theorem.

Concerning the last assertion, observe that the function

$$Z^* \ni z^* \mapsto \sup_{c\in C} \langle z^*, c\rangle$$

is Lipschitz continuous if C is bounded. Hence, if (8.3.3) holds only for some $W \subset Z^*$ dense, then it holds for the whole Z^*. □

Theorem 8.3.4. *Let $\mathbf{F}$ be a λ-dissipative MPVF according to (7.5.1). Assume that for every $\mu \in D(\mathbf{F})$ there exist constants $M, \varepsilon > 0$ such that*

$$\exp_\sharp^\varphi \mu \in D(\mathbf{F}) \quad \textit{and} \quad |\mathbf{F}|_2(\exp_\sharp^\varphi \mu) \leqslant M \tag{8.3.4}$$

for every $\varphi \in \mathrm{Cyl}(\mathbb{H})$ such that $\sup_{\mathbb{H}} |\nabla\varphi| \leqslant \varepsilon$. If $\mu : \mathcal{I} \to D(\mathbf{F})$ is a locally absolutely continuous λ-EVI solution of (8.1.1) with Wasserstein velocity field $\boldsymbol{v}$ satisfying (2.4.8) for every t in the subset $A(\mu) \subset \mathcal{I}$ of full Lebesgue measure, then

$$\textit{for every } t \in A(\mu) \textit{ there exists } \Phi_t \in \overline{\mathrm{co}}(\overline{\mathbf{F}})[\mu_t] \textit{ such that} \quad \boldsymbol{v}_t = \mathbf{pr}_{\mu_t} \circ \mathbf{b}_{\Phi_t}. \tag{8.3.5}$$

In particular, μ satisfies the relaxed barycentric property.

If moreover $\overline{\mathbf{F}} = \mathbf{F}$ and, for every $\nu \in D(\mathbf{F})$, the section $\mathbf{F}[\nu]$ is a convex subset of $\mathcal{P}_2(\mathsf{T}\mathbb{H})$, i.e.

$$\mathbf{F}[\nu] = \mathrm{co}(\mathbf{F})[\nu],$$

then μ satisfies the barycentric property (8.3.2).

Proof. We divide the proof of (8.3.5) into two steps.

Claim 1. Let $t \in A(\mu)$ and $M = M_t$ be the constant associated to the measure μ_t in (8.3.4). Then $\boldsymbol{v}_t \in \overline{\mathrm{co}}(K_t)$, where

$$K_t := \left\{ \mathbf{pr}_{\mu_t}(\mathbf{b}_\Phi) \, : \, \Phi \in \overline{\mathbf{F}}[\mu_t], \, |\Phi|_2 \leqslant M_t \right\} \subset \mathrm{Tan}_{\mu_t} \mathcal{P}_2(\mathbb{H}). \tag{8.3.6}$$

Proof of Claim 1. For every $\zeta \in \mathrm{Cyl}(\mathbb{H})$ there exists $\delta = \delta(\zeta) > 0$ such that $\nu^\zeta := \exp_\sharp^{-\delta\zeta} \mu_t \in D(\mathbf{F})$ and $\boldsymbol{\sigma}^\zeta := (\mathbf{i}_{\mathbb{H}}, \exp^{-\delta\zeta})_\sharp \mu_t \in \Gamma_o^{01}(\mu_t, \nu^\zeta | \mathbf{F})$ is the unique

optimal transport plan between μ_t and ν^ζ.
Thanks to Theorem 7.2.1, the map $s \mapsto W_2^2(\mu_s, \nu^\zeta)$ is differentiable at $s = t$, moreover by employing also (8.1.5b), it holds

$$\delta \int_{\mathbb{H}} \langle \boldsymbol{v}_t(x), \nabla\zeta(x)\rangle \,d\mu_t(x) = \frac{d}{dt}\frac{1}{2}W_2^2(\mu_t, \nu^\zeta) \leqslant [\mathbf{F}, \boldsymbol{\sigma}^\zeta]_{0+} = \lim_{s\downarrow 0}\,[\mathbf{F}, \boldsymbol{\sigma}^\zeta]_{l,s}. \quad (8.3.7)$$

We can choose a decreasing vanishing sequence $(s_k)_{k\in\mathbb{N}} \subset (0,1)$, measures $\nu_k^\zeta := x_\sharp^{s_k}\boldsymbol{\sigma}^\zeta$ and $\Phi_k^\zeta \in \mathbf{F}[\nu_k^\zeta]$ such that $\sup_k |\Phi_k^\zeta|_2 \leqslant M_t$ and $\Phi_k^\zeta \to \Phi^\zeta$ in $\mathcal{P}_2^{sw}(\mathbb{TH})$. Then, by (8.3.1), we get $\Phi^\zeta \in \overline{\mathbf{F}}[\mu_t]$ with $|\Phi^\zeta|_2 \leqslant M_t$ and by (8.3.7) and the upper semicontinuity of $[\cdot,\cdot]_l$ (see Lemma 7.3.1) we get

$$\delta \int_{\mathbb{H}} \langle \boldsymbol{v}_t(x), \nabla\zeta(x)\rangle \,d\mu_t(x) \leqslant \left[\Phi^\zeta, \nu^\zeta\right]_l = \delta \int_{\mathbb{TH}} \langle v, \nabla\zeta(x)\rangle \,d\Phi^\zeta(x,v). \qquad (8.3.8)$$

Indeed, notice that, by [5, Lemma 5.3.2], we have $\Lambda(\Phi^\zeta, \nu^\zeta) = \{\Phi^\zeta \otimes \nu^\zeta\}$ with $(x^0, x^1)_\sharp(\Phi^\zeta \otimes \nu^\zeta) = \boldsymbol{\sigma}^\zeta$.

The expression in (8.3.8) can be written as follows

$$\langle \boldsymbol{v}_t, \nabla\zeta\rangle_{L^2_{\mu_t}(\mathbb{H};\mathbb{H})} \leqslant \langle \mathbf{b}_{\Phi^\zeta}, \nabla\zeta\rangle_{L^2_{\mu_t}(\mathbb{H};\mathbb{H})} = \langle \mathbf{pr}_{\mu_t}(\mathbf{b}_{\Phi^\zeta}), \nabla\zeta\rangle_{L^2_{\mu_t}(\mathbb{H};\mathbb{H})}$$

so that

$$\langle \boldsymbol{v}_t, \nabla\zeta\rangle_{L^2_{\mu_t}(\mathbb{H};\mathbb{H})} \leqslant \sup_{\mathbf{b}\in K_t} \langle \mathbf{b}, \nabla\zeta\rangle_{L^2_{\mu_t}(\mathbb{H};\mathbb{H})}$$

for all $\zeta \in \mathrm{Cyl}(\mathbb{H})$, with K_t as in (8.3.6). Applying Lemma 8.3.3 in $\mathrm{Tan}_{\mu_t}\mathcal{P}_2(\mathbb{H}) \subset L^2_{\mu_t}(\mathbb{H};\mathbb{H})$ we obtain that $\boldsymbol{v}_t \in \overline{\mathrm{co}}(K_t)$.

Claim 2. For every $\boldsymbol{w} \in \overline{\mathrm{co}}(K_t)$ there exists $\Psi \in \overline{\mathrm{co}}(\overline{\mathbf{F}})[\mu_t]$ such that $\boldsymbol{w} = \mathbf{pr}_{\mu_t} \circ \mathbf{b}_\Psi$.

Proof of Claim 2. Notice that an element $\boldsymbol{w} \in \mathrm{Tan}_\mu\mathcal{P}_2(\mathbb{H})$ coincides with $\mathbf{pr}_\mu(\mathbf{b}_\Psi)$ for $\Psi \in \mathcal{P}_2(\mathbb{TH}|\mu)$ if and only if

$$\int \langle \boldsymbol{w}, \nabla\zeta\rangle \,d\mu = \int \langle v, \nabla\zeta\rangle \,d\Psi(x,v) \qquad (8.3.9)$$

for every $\zeta \in \mathrm{Cyl}(\mathbb{H})$. It is easy to check that any element $\boldsymbol{w} \in \mathrm{co}(K)_t$ can be represented as $\mathbf{pr}_{\mu_t}(\mathbf{b}_\Psi)$ (and thus as in (8.3.9)) for some $\Psi \in \mathrm{co}(\overline{\mathbf{F}}[\mu_t])$. If $\boldsymbol{w} \in \overline{\mathrm{co}}(K_t)$ we can find a sequence $(\Psi_n)_{n\in\mathbb{N}} \subset \mathrm{co}(\overline{\mathbf{F}}[\mu_t])$ such that $|\Psi_n|_2 \leqslant M_t$ and $\boldsymbol{w}_n = \mathbf{pr}_{\mu_t}(\mathbf{b}_{\Psi_n}) \to \boldsymbol{w}$ in $L^2_{\mu_t}(\mathbb{H};\mathbb{H})$. Since the sequence $(\Psi_n)_{n\in\mathbb{N}}$ is relatively compact in $\mathcal{P}_2^{sw}(\mathbb{TH})$ by Proposition 6.3.3(2), we can extract a (not relabeled) subsequence converging to a limit Ψ in $\mathcal{P}_2^{sw}(\mathbb{TH})$, as $n \to +\infty$. By definition $\Psi \in \overline{\mathrm{co}}(\overline{\mathbf{F}}[\mu_t])$ with $|\Psi|_2 \leqslant M_t$. We can eventually pass to the limit in (8.3.9) written for $\boldsymbol{w}_n$ and Ψ_n thanks to $\mathcal{P}_2^{sw}(\mathbb{TH})$ convergence, obtaining the corresponding identity for $\boldsymbol{w}$ and Ψ in the limit.

The thesis (8.3.5) follows by Claim 1 and Claim 2.

Finally, being μ locally absolutely continuous, it satisfies the continuity equation driven by $\boldsymbol{v}$ in the sense of distributions (see Theorem 2.4.6), so that by (8.3.5) we have

$$\frac{d}{dt}\int_{\mathbb{H}} \zeta(x)\,d\mu_t(x) = \int_{\mathbb{H}} \langle \nabla\zeta(x), \boldsymbol{v}_t(x)\rangle \,d\mu_t(x) = \int_{\mathbb{TH}} \langle \nabla\zeta(x), v\rangle \,d\Phi_t(x,v),$$

for every $\zeta \in \mathrm{Cyl}(\mathbb{H})$ and every $t \in A(\mu)$. □

Remark 8.3.5. Using a standard approximation argument (see for example the proof of Lemma 5.1.12(f) in [5]) it is possible to show that actually the barycentric property (8.3.2) holds for every $\varphi \in C^{1,1}(\mathbb{H};\mathbb{R})$ (indeed, in this case, $\nabla\varphi \in \mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{H})$ for every $\mu \in \mathcal{P}_2(\mathbb{H})$).

Remark 8.3.6. We point out that the result stated in Theorem 8.3.4 is still valid if we replace the convex hull of $\mathbf{F}$ defined in (7.7.1) using the "flat" structure of $\mathcal{P}_2(\mathsf{T}\mathbb{H})$, with the following one which makes use of plan interpolations

$$\widetilde{\mathrm{co}}(\mathbf{F})(\nu) := \left\{ \left(x, \sum_{k_1}^{N} \alpha_k v_k\right)_\sharp \boldsymbol{\Phi} \;\middle|\; \begin{array}{l} \boldsymbol{\Phi} \in \mathcal{P}(\mathbb{H}^{N+1}),\ (x, v_k)_\sharp \boldsymbol{\Phi} = \Phi_k,\ \Phi_k \in \mathbf{F}[\nu], \\ \alpha_k \geqslant 0,\ k = 1, \dots, N,\ \sum_{k=1}^N \alpha_k = 1,\ N \in \mathbb{N} \end{array} \right\},$$

for any $\nu \in D(\mathbf{F})$, where

$$x(x, v_1, \dots, v_N) = x \quad \text{and} \quad v_k(x, v_1, \dots, v_N) = v_k,\ k = 1, \dots, N.$$

Indeed, $\mathrm{co}(\mathbf{F})(\nu)$ and $\widetilde{\mathrm{co}}(\mathbf{F})(\nu)$ share the same barycentric projection. However, while $\mathrm{co}(\mathbf{F})$ preserves dissipativity as proved in Proposition 7.7.2, $\widetilde{\mathrm{co}}(\mathbf{F})(\nu)$ does not satisfy this property in general, as highlighted in the following example: let $\mathbb{H} = \mathbb{R}$ and consider the PVF $\mathbf{F}$, with domain $D(\mathbf{F}) = \left\{\delta_0, \frac{1}{2}\delta_1 + \frac{1}{2}\delta_0\right\}$, defined by

$$\mathbf{F}[\delta_0] := \frac{1}{2}\delta_{(0,3)} + \frac{1}{2}\delta_{(0,-3)}, \qquad \mathbf{F}\left[\frac{1}{2}\delta_1 + \frac{1}{2}\delta_0\right] := \frac{1}{2}\delta_{(1,2)} + \frac{1}{2}\delta_{(0,1)}.$$

Then $\mathbf{F}$ is dissipative, indeed

$$\left[\mathbf{F}[\delta_0], \mathbf{F}\left[\frac{1}{2}\delta_1 + \frac{1}{2}\delta_0\right]\right]_r \leqslant -1 \leqslant 0.$$

However, $\widetilde{\mathrm{co}}(\mathbf{F})$ is not dissipative, indeed, if we take $\delta_{(0,0)} \in \widetilde{\mathrm{co}}(\mathbf{F})[\delta_0]$, we have

$$\left[\delta_{(0,0)}, \mathbf{F}\left[\frac{1}{2}\delta_1 + \frac{1}{2}\delta_0\right]\right]_r = 2 > 0.$$

As a complement to the studies investigated in this section, we prove the converse characterization of Theorem 8.3.4 in the particular case of *regular measures* or *regular vector fields*. We refer to [5, Definitions 6.2.1, 6.2.2] for the definition of $\mathcal{P}_2^r(\mathbb{H})$, that is the space of regular measures on $\mathbb{H}$. When $\mathbb{H} = \mathbb{R}^d$ has finite dimension, $\mathcal{P}_2^r(\mathbb{H})$ is just the subset of measures in $\mathcal{P}_2(\mathbb{H})$ which are absolutely continuous w.r.t. the d-dimensional Lebesgue measure $\mathcal{L}^d$.

Theorem 8.3.7. *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1)*. Let* $\mu : \mathcal{I} \to D(\mathbf{F})$ *be a locally absolutely continuous curve satisfying the relaxed barycentric property of Definition 8.3.1. If for a.e.* $t \in \mathcal{I}$ *at least one of the following properties holds:*

1. $\mu_t \in \mathcal{P}_2^r(\mathbb{H})$,
2. $\overline{\mathbf{F}}[\mu_t]$ *contains a unique element* Φ_t *concentrated on a map, i.e.* $\Phi_t = (\mathbf{i}_{\mathbb{H}}, \mathbf{b}_{\Phi_t})_\sharp \mu_t$

then μ *is* λ*-EVI solution of* (8.1.1).

Proof. Take $\varphi \in \mathrm{Cyl}(\mathbb{H})$ and observe that, since μ has the relaxed barycentric property, then for a.e. $t \in \mathcal{I}$ (recall Theorem 7.2.1) there exists $\Phi_t \in \overline{\mathrm{co}}(\overline{\mathbf{F}}[\mu_t])$ such that

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{H}} \varphi(x)\,d\mu_t(x) &= \int_{\mathsf{T}\mathbb{H}} \langle \nabla\varphi(x), v\rangle\,d\Phi_t\\ &= \int_{\mathbb{H}} \langle \nabla\varphi, \mathbf{pr}_{\mu_t}\circ \mathbf{b}_{\Phi_t}\rangle\,d\mu_t\\ &= \int_{\mathbb{H}} \langle \boldsymbol{v}_t, \nabla\varphi\rangle\,d\mu_t,\end{aligned}$$

hence μ solves the continuity equation $\partial_t\mu_t + \mathrm{div}(\boldsymbol{v}_t\mu_t) = 0$, with $\boldsymbol{v}_t = \mathbf{pr}_{\mu_t}\circ \mathbf{b}_{\Phi_t} \in \mathrm{Tan}_{\mu_t}\mathcal{P}_2(\mathbb{H})$. By Theorem 7.2.1, we also know that

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t,\nu) = \int_{\mathbb{H}^2} \langle \boldsymbol{v}_t(x_0), x_0 - x_1\rangle\,d\boldsymbol{\gamma}_t(x_0,x_1) \tag{8.3.10}$$

for any $t \in A(\mu,\nu)$, $\boldsymbol{\gamma}_t \in \Gamma_o(\mu_t,\nu)$, $\nu \in \mathcal{P}_2(\mathbb{H})$. Possibly disregarding a Lebesgue negligible set, we can decompose the set $A(\mu,\nu)$ in the union $A_1 \cup A_2$, where A_1, A_2 correspond to the times t for which the properties (1) and (2) hold.

If $t \in A_1$ and $\nu \in D(\mathbf{F})$, then by [5, Theorem 6.2.10], since $\mu_t \in \mathcal{P}_2^r(\mathbb{H})$, there exists a unique $\boldsymbol{\gamma}_t \in \Gamma_o(\mu_t,\nu)$ and $\boldsymbol{\gamma}_t = (\mathbf{i}_{\mathbb{H}}, \mathbf{r}_t)_\sharp\mu_t$ for some map $\mathbf{r}_t$ s.t. $\mathbf{i}_{\mathbb{H}} - \mathbf{r}_t \in \mathrm{Tan}_{\mu_t}\mathcal{P}_2(\mathbb{H}) \subset L^2_{\mu_t}(\mathbb{H};\mathbb{H})$ (recall [5, Proposition 8.5.2]), so that

$$\begin{aligned}\int_{\mathbb{H}^2} \langle \boldsymbol{v}_t(x_0), x_0 - x_1\rangle\,d\boldsymbol{\gamma}_t(x_0,x_1) &= \int_{\mathbb{H}} \langle \boldsymbol{v}_t(x_0), x_0 - \mathbf{r}_t(x_0)\rangle\,d\mu_t(x_0)\\ &= \int_{\mathbb{H}} \langle \mathbf{b}_{\Phi_t}, x_0 - \mathbf{r}_t(x_0)\rangle\,d\mu_t(x_0)\\ &= \int_{\mathsf{T}\mathbb{H}} \langle v, x - \mathbf{r}_t(x)\rangle\,d\Phi_t(x,v)\\ &= [\Phi_t,\nu]_r,\end{aligned} \tag{8.3.11}$$

where we also applied Theorem 7.1.8 and Remark 7.4.2, recalling that in this case $\Lambda(\Phi_t,\nu)$ is a singleton.

If $t \in A_2$ we can select the optimal plan $\boldsymbol{\gamma}_t \in \Gamma_o(\mu_t,\nu)$ along which

$$[\Phi_t,\nu]_r = [\Phi_t,\boldsymbol{\gamma}_t]_{r,0} = \int_{\mathbb{H}} \langle \mathbf{b}_{\Phi_t}(x_0), x_0 - x_1\rangle\,d\boldsymbol{\gamma}_t(x_0,x_1).$$

If $\mathbf{r}_t$ is the barycenter of $\boldsymbol{\gamma}_t$ with respect to its first marginal μ_t, recalling that $\mathbf{i}_{\mathbb{H}} - \mathbf{r}_t \in \mathrm{Tan}_{\mu_t}\mathcal{P}_2(\mathbb{H})$ (see also the proof of [5, Thm. 12.4.4]) we also get

$$\begin{aligned}\int_{\mathbb{H}^2} \langle \boldsymbol{v}_t(x_0), x_0 - x_1\rangle\,d\boldsymbol{\gamma}_t(x_0,x_1) &= \int_{\mathbb{H}} \langle \boldsymbol{v}_t(x_0), x_0 - \mathbf{r}_t(x_0)\rangle\,d\mu_t(x_0)\\ &= \int_{\mathbb{H}} \langle \mathbf{b}_{\Phi_t}(x_0), x_0 - \mathbf{r}_t(x_0)\rangle\,d\mu_t(x_0)\\ &= \int_{\mathbb{H}} \langle \mathbf{b}_{\Phi_t}(x_0), x_0 - x_1\rangle\,d\boldsymbol{\gamma}_t(x_0,x_1)\\ &= [\Phi_t,\nu]_r,\end{aligned} \tag{8.3.12}$$

where we still applied Theorem 7.1.8 and Remark 7.4.2.

Combining (8.3.10) with (8.3.11) and (8.3.12) we eventually get

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t,\nu) = [\Phi_t,\nu]_r \leqslant -[\Psi,\mu_t]_r + \lambda W_2^2(\mu_t,\nu)$$

for every $\Psi \in \mathbf{F}[\nu]$, by definition of $\hat{\mathbf{F}}$ and the fact that $\overline{\text{co}}(\overline{\mathbf{F}})[\mu_t] \subset \hat{\mathbf{F}}[\mu_t]$. $\square$

9 EXPLICIT AND IMPLICIT EULER SCHEMES, EXAMPLES

The aim of this chapter is to present two tools to prove existence of EVI solutions: the Explicit Euler scheme and the Implicit one. In particular Section 9.1 presents the Explicit Euler scheme and the conditions under which we can prove its convergence; in Section 9.2 we presents the main consequence of the solvability of the Explicit Euler scheme in terms of stability and uniqueness of EVI solutions; Section 9.3 deals with law invariant dissipative operators in a space of random variables $\mathcal{H}$ and presents their main properties; Section 9.4 contains the procedure that allows to construct a dissipative operator on $\mathcal{H}$ starting from a MPVF on $\mathcal{P}_2(\mathbb{H})$; finally in Section 9.5 are listed a few examples of MPVFs and EVI solutions.

This Chapter is the result of a collaboration with Giulia Cavagnari and Giuseppe Savaré and Sections 9.1, 9.2 and 9.5 appeared in [34].

In this whole chapter, $\mathbb{H}$ is a separable Hilbert space with $\dim(\mathbb{H}) \geqslant 2$.

9.1 EXPLICIT EULER SCHEME

Our first strategy to prove the existence of a λ-EVI solution to (8.1.1), is to define an Explicit Euler scheme.
In the following $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ denote the floor and the ceiling functions respectively, i.e.

$$\lfloor t\rfloor := \max\{m \in \mathbb{Z} \mid m \leqslant t\} \quad \text{and} \quad \lceil t\rceil := \min\{m \in \mathbb{Z} \mid m \geqslant t\}, \tag{9.1.1}$$

for any $t \in \mathbb{R}$.

Definition 9.1.1 (Explicit Euler Scheme)**.** Let $\mathbf{F}$ be a MPVF and suppose we are given a step size $\tau > 0$, an initial datum $\mu_0 \in D(\mathbf{F})$, a bounded interval $[0, T]$, corresponding to the final step $N(T,\tau) := \lceil T/\tau\rceil$, and a stability bound $L > 0$. A sequence $(M^n_\tau, \Phi^n_\tau)_{0\leqslant n\leqslant N(T,\tau)} \subset D(\mathbf{F}) \times \mathbf{F}$ is a L-*stable solution to the Explicit Euler Scheme in* $[0, T]$ *starting from* $\mu_0 \in D(\mathbf{F})$ if

$$\begin{cases} M^0_\tau = \mu_0, \\ \Phi^n_\tau \in \mathbf{F}[M^n_\tau],\ |\Phi^n_\tau|_2 \leqslant L & 0 \leqslant n < N(T,\tau), \\ M^n_\tau = (\exp^\tau)_\sharp \Phi^{n-1}_\tau & 1 \leqslant n \leqslant N(T,\tau). \end{cases} \tag{EE}$$

We define the following two different interpolations of the sequence (M^n_τ, Φ^n_τ):

- the affine interpolation:

$$M_\tau(t) := (\exp^{t-n\tau})_\sharp \Phi^n_\tau \quad \text{if } t \in [n\tau,(n+1)\tau] \text{ for some } n \in \mathbb{N},\ 0 \leqslant n < N(T,\tau), \tag{9.1.2}$$

- the piecewise constant interpolation:

$$\bar{M}_\tau(t) := M_\tau^{\lfloor t/\tau \rfloor}, \quad t \in [0,T], \tag{9.1.3}$$

$$\mathbf{F}_\tau(t) := \Phi_\tau^{\lfloor t/\tau \rfloor}, \quad t \in [0,T]. \tag{9.1.4}$$

We define the following (possibly empty) sets

$$\begin{aligned} \mathscr{E}(\mu_0,\tau,T,L) &:= \Big\{ (M_\tau,\mathbf{F}_\tau) \mid M_\tau, \mathbf{F}_\tau \text{ are as in (9.1.2), (9.1.4) respectively} \Big\}, \\ \mathscr{M}(\mu_0,\tau,T,L) &:= \Big\{ M_\tau \mid M_\tau \text{ is the curve given by (9.1.2)} \Big\}. \end{aligned} \tag{9.1.5}$$

Remark 9.1.2. We immediately notice that, if $(M_\tau, \mathbf{F}_\tau) \in \mathscr{E}(\mu_0,\tau,T,L)$ and $\bar{M}_\tau(\cdot)$ is as in (9.1.3), then the following holds for any $0 \leqslant s \leqslant t \leqslant T$:

1. the affine interpolation can be trivially written as

$$M_\tau(t) = \left(\exp^{t-\lfloor t/\tau \rfloor \tau}\right)_\sharp (\mathbf{F}_\tau(t))\,;$$

2. M_τ satisfies the uniform Lipschitz bound

$$W_2(M_\tau(t), M_\tau(s)) \leqslant L|t-s|; \tag{9.1.6}$$

3. we have the following estimate

$$W_2(\bar{M}_\tau(t), M_\tau(t)) = W_2\left(M_\tau\left(\left\lfloor \frac{t}{\tau} \right\rfloor \tau\right), M_\tau(t)\right) \leqslant L\tau. \tag{9.1.7}$$

The estimate (9.1.7) shows that the stability and convergence results stated for the affine interpolation can be easily adapted to the piecewise constant one.

Notice that, in general, $\mathbf{F}[\mu]$ is not reduced to a singleton, so that $\mathscr{E}(\mu_0,\tau,T,L)$ and $\mathscr{M}(\mu_0,\tau,T,L)$ may contain more than one element.

9.1.1 *The Explicit Euler Scheme: preliminary estimates*

We first prove a simple estimate and a discrete version of (λ-EVI).

Proposition 9.1.3. *Every solution* $(M_\tau, \mathbf{F}_\tau) \in \mathscr{E}(\mu_0,\tau,T,L)$ *of* (EE) *satisfies*

$$W_2(M_\tau(t), \mu_0) \leqslant Lt, \quad |\mathbf{F}_\tau(t)|_2 \leqslant L \quad \textit{for every } t \in [0,T], \tag{9.1.8}$$

$$W_2(M_\tau(t), M_\tau(s)) \leqslant L|t-s| \quad \textit{for every } s,t \in [0,T], \tag{9.1.9}$$

and

$$\frac{d}{dt}\frac{1}{2}W_2^2(M_\tau(t),\nu) \leqslant [\mathbf{F}_\tau(t),\nu]_r + \tau|\mathbf{F}_\tau(t)|_2^2 \leqslant [\mathbf{F}_\tau(t),\nu]_r + \tau L^2 \tag{IEVI}$$

for every $t \in [0,T]$ *and* $\nu \in \mathcal{P}_2(\mathbb{H})$, *with possibly countable exceptions. In particular*

$$\frac{1}{2}W_2^2(M_\tau^{n+1},\nu) - \frac{1}{2}W_2^2(M_\tau^n,\nu) \leqslant \tau\,[\Phi_\tau^n,\nu]_r + \frac{1}{2}\tau^2 L^2 \tag{9.1.10}$$

for every $0 \leqslant n < N(T,\tau)$ *and* $\nu \in \mathcal{P}_2(\mathbb{H})$.

Proof. The second inequality of (9.1.8) is a trivial consequence of the definition of $\mathscr{E}(\mu_0, \tau, T, L)$, the first inequality is a particular case of (9.1.9). The estimate (9.1.9) is immediate if $n\tau \leqslant s < t \leqslant (n+1)\tau$ since

$$\begin{aligned} W_2(M_\tau(s), M_\tau(t)) &= W_2((\exp^{s-n\tau})_\sharp \Phi^n_\tau, (\exp^{t-n\tau})_\sharp \Phi^n_\tau) \\ &\leqslant \sqrt{\int_{\mathrm{T}\mathbb{H}} |(t-s)v)|^2 \, d\Phi^n_\tau} \\ &= (t-s)\sqrt{\int_{\mathrm{T}\mathbb{H}} |v|^2 \, d\Phi^n_\tau} \\ &\leqslant (t-s)L. \end{aligned}$$

This implies that the metric velocity of M_τ is bounded by L in $[0, T]$ and therefore M_τ is L-Lipschitz.

Let us recall that for every $\nu \in \mathcal{P}_2(\mathbb{H})$ and $\Phi \in \mathcal{P}_2(\mathrm{T}\mathbb{H})$ the function $g(t) := \frac{1}{2}W_2^2(\exp^t_\sharp \Phi, \nu)$ satisfies

$$t \mapsto g(t) - \frac{1}{2}t^2|\Phi|_2^2 \text{ is concave}, \quad g'_r(0) = [\Phi, \nu]_r, \quad g'(t) \leqslant [\Phi, \nu]_r + t|\Phi|_2^2 \quad (9.1.11)$$

for $t \geqslant 0$, by Definition 7.1.4 and Proposition 7.1.3. In particular, the concavity yields the differentiability of g with at most countable exceptions. Thus, taking any $n \in \mathbb{N}$, $0 \leqslant n < N(T, \tau)$, $t \in [n\tau, (n+1)\tau)$ and $\Phi = \Phi^n_\tau$ so that $\exp^t_\sharp \Phi = M_\tau(t)$, (9.1.11) yields (IEVI). The inequality in (9.1.10) follows by integration in each interval $[n\tau, (n+1)\tau]$. □

We conclude this subsection with a stability estimate. We introduce the notation

$$I_\kappa(t) := \int_0^t e^{\kappa r} \, dr = \frac{1}{\kappa}(e^{\kappa t} - 1) \quad \text{if } \kappa \neq 0; \quad I_0(t) := t.$$

Notice that for every $t \geqslant 0$

$$I_\kappa(t) \leqslant te^{\kappa t} \quad \text{if } \kappa \geqslant 0. \tag{9.1.12}$$

Proposition 9.1.4. *Let* $M_\tau \in \mathscr{M}(\mu_0, \tau, T, L)$ *and* $M'_\tau \in \mathscr{M}(\mu'_0, \tau, T, L)$. *If* $\lambda_+\tau \leqslant 2$ *then*

$$W_2(M_\tau(t), M'_\tau(t)) \leqslant W_2(\mu_0, \mu'_0)e^{\lambda t} + 8L\sqrt{t\tau}\Big(1 + |\lambda|\sqrt{t\tau}\Big)e^{\lambda_+ t}$$

for every $t \in [0, T]$.

Proof. Let us set $w(t) := W_2(M_\tau(t), M'_\tau(t))$. Since by Proposition 7.1.3(2), in every interval $[n\tau, (n+1)\tau]$ the function $t \mapsto w^2(t) - 4L^2(t - n\tau)^2$ is concave, with

$$\frac{d}{dt} w^2(t)\Big|_{t=n\tau+} = 2\left[F_\tau(t), F'_\tau(t)\right]_r \leqslant 2\lambda W_2^2(\bar{M}_\tau(t), \bar{M}'_\tau(t)),$$

we obtain

$$\frac{d}{dt} w^2(t) \leqslant 2\lambda W_2^2(\bar{M}_\tau(t), \bar{M}'_\tau(t)) + 8L^2\tau$$

for every $t \in [0, T]$, with possibly countable exceptions. Using the identity

$$a^2 - b^2 = 2b(a-b) + |a-b|^2$$

with $a = W_2(\bar{M}_\tau(t), \bar{M}'_\tau(t))$ and $b = W_2(M_\tau(t), M'_\tau(t))$ and observing that

$$|a-b| \leqslant W_2(\bar{M}_\tau(t), M_\tau(t)) + W_2(\bar{M}'_\tau(t), M'_\tau(t)) \leqslant 2L\tau,$$

we eventually get

$$\begin{aligned}\frac{d}{dt}w^2(t) &\leqslant 2\lambda w^2(t) + 8L^2\tau + 8|\lambda|L\tau w(t) + \lambda_+ 8L^2\tau^2 \\ &\leqslant 2\lambda w^2(t) + 8|\lambda|L\tau w(t) + 24L^2\tau,\end{aligned}$$

since $\lambda_+\tau \leqslant 2$ by assumption. The Gronwall estimate in [5, Lemma 4.1.8] and (9.1.12) yield

$$\begin{aligned}w(t) &\leqslant \left(w^2(0)e^{2\lambda t} + 24L^2\tau I_{2\lambda}(t)\right)^{1/2} + 8|\lambda|L\tau I_\lambda(t) \\ &\leqslant w(0)e^{\lambda t} + 8L\sqrt{t\tau}\left(1 + |\lambda|\sqrt{t\tau}\right)e^{\lambda_+ t}.\end{aligned}$$

□

9.1.2 *Error estimates for the Explicit Euler scheme*

In this subsection we prove that the family of affine interpolants is Cauchy, providing estimates under different step sizes and a uniform (optimal, see [106]) error estimate between the affine interpolant and the λ-EVI solution for $\mathbf{F}$.

Theorem 9.1.5. *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF as in* (7.5.1)*. If* $M_\tau \in \mathscr{M}(M^0_\tau, \tau, T, L)$*,* $M_\eta \in \mathscr{M}(M^0_\eta, \eta, T, L)$ *with* $\lambda\sqrt{T(\tau+\eta)} \leqslant 1$*, then for every* $\delta > 1$ *there exists a constant* $C(\delta)$ *such that*

$$W_2(M_\tau(t), M_\eta(t)) \leqslant \left(\sqrt{\delta}W_2(M^0_\tau, M^0_\eta) + C(\delta)L\sqrt{(\tau+\eta)(t+\tau+\eta)}\right)e^{\lambda_+ t}$$

for every $t \in [0, T]$.

Proof. We argue as in the proof of Theorem 8.2.3. Since λ-dissipativity implies λ'-dissipativity for $\lambda' \geqslant \lambda$, it is not restrictive to assume $\lambda > 0$. We set $\sigma := \tau + \eta$. We will extensively use the a priori bounds (9.1.8) and (9.1.9); in particular,

$$W_2(M_\tau(t), \bar{M}_\tau(t)) \leqslant L\tau, \quad W_2(M_\eta(t), \bar{M}_\eta(t)) \leqslant L\eta.$$

We will also extend M_τ and $\bar{M}_\tau$ for negative times by setting

$$M_\tau(t) = \bar{M}_\tau(t) = M^0_\tau, \quad \mathbf{F}_\tau(t) = M^0_\tau \otimes \delta_0 \quad \text{if } t < 0. \tag{9.1.13}$$

The proof is divided into several steps.

1. *Doubling variables.*

We fix a final time $t \in [0, T]$ and two variables $r, s \in [0, t]$ together with the functions

$$\begin{aligned} w(r,s) &:= W_2(M_\tau(r), M_\eta(s)), & w_\tau(r,s) &:= W_2(\bar{M}_\tau(r), M_\eta(s)), \\ w_\eta(r,s) &:= W_2(M_\tau(r), \bar{M}_\eta(s)), & w_{\tau,\eta}(r,s) &:= W_2(\bar{M}_\tau(r), \bar{M}_\eta(s)), \end{aligned} \tag{9.1.14}$$

observing that

$$\max\{|w - w_\tau|, |w_\eta - w_{\tau,\eta}|\} \leqslant L\tau, \quad \max\{|w - w_\eta|, |w_\tau - w_{\tau,\eta}|\} \leqslant L\eta. \tag{9.1.15}$$

By Proposition 9.1.3, we can write (IEVI) for M_τ and get

$$\frac{\partial}{\partial r}\frac{1}{2}W_2^2(M_\tau(r), \nu_1) \leqslant \tau|\mathbf{F}_\tau(r)|_2^2 + [\mathbf{F}_\tau(r), \nu_1]_r \quad \text{for every } \nu_1 \in \mathcal{P}_2(\mathbb{H}), \tag{IEVI$_\tau$}$$

and for M_η obtaining

$$\begin{aligned} &\frac{\partial}{\partial s}\frac{1}{2}W_2^2(M_\eta(s), \nu_2) \\ &\leqslant \eta|\mathbf{F}_\eta(s)|_2^2 + [\mathbf{F}_\eta(s), \nu_2]_r \\ &\leqslant \eta|\mathbf{F}_\eta(s)|_2^2 + \lambda W_2^2(\bar{M}_\eta(s), \nu_2) - \left[\Phi, \bar{M}_\eta(s)\right]_r \quad \text{for } \Phi \in \mathbf{F}[\nu_2],\ \nu_2 \in D(\mathbf{F}). \end{aligned} \tag{IEVI$_\eta$}$$

Apart from possible countable exceptions, (IEVI$_\tau$) holds for $r \in (-\infty, t]$ and (IEVI$_\eta$) for $s \in [0, t]$. Taking $\nu_1 = \bar{M}_\eta(s)$, $\nu_2 = \bar{M}_\tau(r)$, $\Phi = \mathbf{F}_\tau(\max\{r, 0\}) \in \mathbf{F}[\bar{M}_\tau(r)]$, summing the two inequalities (IEVI$_{\tau,\eta}$), setting

$$f(r,s) := \begin{cases} 2LW_2(\bar{M}_\eta(s), M_\tau(0)) = 2Lw_\eta(0,s) & \text{if } r < 0, \\ 0 & \text{if } r \geqslant 0, \end{cases}$$

using (9.1.8) and the λ-dissipativity of $\mathbf{F}$, we obtain

$$\frac{\partial}{\partial r}w_\eta^2(r,s) + \frac{\partial}{\partial s}w_\tau^2(r,s) \leqslant 2\lambda w_{\tau,\eta}^2(r,s) + 2L^2\sigma + f(r,s)$$

in $(-\infty, t] \times [0, t]$ (see also [89, Lemma 6.15]). By multiplying both sides by $e^{-2\lambda s}$, we have

$$\frac{\partial}{\partial r}e^{-2\lambda s}w_\eta^2 + \frac{\partial}{\partial s}e^{-2\lambda s}w_\tau^2 \leqslant \left(2\lambda\left(w_{\tau,\eta}^2 - w_\tau^2\right) + f + 2L^2\sigma\right)e^{-2\lambda s}. \tag{9.1.16}$$

Using (9.1.15), the inequalities

$$\begin{aligned} &w_{\tau,\eta} + w_\tau = w_{\tau,\eta} - w_\tau + 2(w_\tau - w) + 2w \leqslant 2L\sigma + 2w, \\ &|w(r,s) - w(s,s)| \leqslant L|r - s| \end{aligned}$$

and the elementary inequality $a^2 - b^2 \leqslant |a-b||a+b|$, we get

$$2\big(w_{\tau,\eta}^2(r,s) - w_\tau^2(r,s)\big) \leqslant R_{r,s}, \quad \text{if } r, s \leqslant t,$$

where $R_{r,s} := 4L^2\sigma(\sigma + |r - s|) + 4L\sigma w(s,s)$. Thus (9.1.16) becomes

$$\frac{\partial}{\partial r}e^{-2\lambda s}w_\eta^2 + \frac{\partial}{\partial s}e^{-2\lambda s}w_\tau^2 \leqslant Z_{r,s}, \tag{9.1.17}$$

where $Z_{r,s} := \left(R\lambda + f + 2L^2\sigma\right)e^{-2\lambda s}$.

2. *Penalization.*

We fix any $\varepsilon > 0$ and apply the Divergence Theorem to the inequality (9.1.17) in the two-dimensional strip $Q^{\varepsilon}_{0,t}$ as in (8.2.6) and we get

$$\begin{aligned}\int_{t-\varepsilon}^{t} e^{-2\lambda t} w_\tau^2(r,t)\,dr \leqslant \int_{-\varepsilon}^{0} & w_\tau^2(r,0)\,dr + \\ & + \int_0^t e^{-2\lambda s}\left(w_\tau^2(s,s) - w_\eta^2(s,s)\right)\,ds \\ & + \int_0^t e^{-2\lambda s}\left(w_\eta^2(s-\varepsilon,s) - w_\tau^2(s-\varepsilon,s)\right)\,ds \\ & + \iint_{Q^{\varepsilon}_{0,t}} Z_{r,s}\,drds.\end{aligned} \tag{9.1.18}$$

3. *Estimates of the r.h.s..*

We want to estimate the integrals (say I_0, I_1, I_2, I_3) of the right hand side of (9.1.18) in terms of

$$w(s) := w(s,s) \quad \text{and} \quad W(t) := \sup_{0\leqslant s\leqslant t} e^{-\lambda s} w(s).$$

We easily get

$$I_0 = \int_{-\varepsilon}^{0} w_\tau^2(r,0)\,dr = \varepsilon w^2(0).$$

(9.1.15) yields

$$|w_\tau(s,s) - w_\eta(s,s)| \leqslant L(\tau+\eta) = L\sigma$$

and

$$|w_\tau^2(s,s) - w_\eta^2(s,s)| \leqslant L\sigma\Big(L\sigma + 2w(s)\Big);$$

after an integration,

$$I_1 \leqslant L^2\sigma^2 t + 2L\sigma\int_0^t e^{-2\lambda s} w(s)\,ds \leqslant L^2\sigma^2 t + 2L\sigma t W(t).$$

Performing the same computations for the third integral term at the r.h.s. of (9.1.18) we end up with

$$\begin{aligned} I_2 &= \int_0^t e^{-2\lambda s}\left(w_\eta^2(s-\varepsilon,s) - w_\tau^2(s-\varepsilon,s)\right)\,ds \\ &\leqslant L^2 t\sigma^2 + 2L\sigma\int_0^t e^{-2\lambda s} w(s-\varepsilon,s)\,ds \\ &\leqslant L^2\sigma^2 t + 2L^2\sigma\varepsilon t + 2L\sigma\int_0^t e^{-2\lambda s} w(s)\,ds \\ &\leqslant L^2\sigma^2 t + 2L^2\sigma\varepsilon t + 2L\sigma t W(t).\end{aligned}$$

Eventually, using the elementary inequalities,

$$\iint_{Q^{\varepsilon}_{0,t}} \lambda e^{-2\lambda s}\,dr\,ds \leqslant \frac{\varepsilon}{2}, \quad \iint_{Q^{\varepsilon}_{0,t}} e^{-2\lambda s} w(s,s)\,dr\,ds = \varepsilon\int_0^t e^{-2\lambda s} w(s)\,ds,$$

and $f(r,s) \leqslant 2L^2(\eta+s)+2Lw(s)$ for $r<0$ and $f(r,s)=0$ for $r \geqslant 0$, we get

$$\begin{aligned} I_3 = \iint_{Q^{\varepsilon}_{0,t}} Z_{r,s}\,drds &\leqslant 2L^2\sigma\varepsilon(\sigma+\varepsilon)+4L\lambda\sigma\varepsilon\int_0^t e^{-2\lambda s}w(s)\,ds+2L^2\sigma\varepsilon t \\ &\quad + 2\iint_{Q^{\varepsilon}_{0,\min\{\varepsilon,t\}}} (L^2(\eta+s)+Lw(s))e^{-2\lambda s}\,drds \\ &\leqslant 2L^2\sigma\varepsilon(\sigma+\varepsilon)+2L^2\varepsilon^2(\sigma+\varepsilon)+2L^2\sigma\varepsilon t+4L\lambda\sigma\varepsilon tW(t)+2L\varepsilon^2 W(\min\{t,\varepsilon\}). \end{aligned}$$

We eventually get

$$\begin{aligned} \sum_{k=0}^{3} I_k &\leqslant \varepsilon w^2(0)+2L^2\sigma^2 t \\ &\quad + 4L^2\sigma\varepsilon t+2L^2\varepsilon(\sigma+\varepsilon)^2+4L\sigma(1+\lambda\varepsilon)tW(t)+2L\varepsilon^2W(\min\{t,\varepsilon\}). \end{aligned} \tag{9.1.19}$$

4. L.h.s. and penalization

We want to use the first integral term in (9.1.18) to derive a pointwise estimate for $w(t)$;

(9.1.9) and (9.1.14) yield

$$w(t)=w(t,t) \leqslant L(t-r)+w(r,t) \leqslant L(\tau+|t-r|)+w_\tau(r,t). \tag{9.1.20}$$

We then square (9.1.20), use the Young inequality (i.e. $2ab \leqslant \frac{a^2}{\vartheta}+\vartheta b^2$ for any $a,b \geqslant 0$, $\vartheta>0$), multiply the resulting inequality by $\frac{e^{-2\lambda t}}{\varepsilon}$ and integrate over the interval $(t-\varepsilon,t)$. So that, for every $\delta,\delta_\star>1$ conjugate coefficients, we get

$$\begin{aligned} e^{-2\lambda t}w^2(t) &\leqslant \frac{\delta}{\varepsilon}\int_{t-\varepsilon}^{t} e^{-2\lambda t}w_\tau^2(r,t)\,dr+\delta_\star L^2(\tau+\varepsilon)^2 \\ &\leqslant \frac{\delta}{\varepsilon}(I_0+I_1+I_2+I_3)+\delta_\star L^2(\tau+\varepsilon)^2, \end{aligned}$$

with I_0,I_1,I_2,I_3 as in step 3. Using (9.1.19) yields

$$\begin{aligned} e^{-2\lambda t}w^2(t) \leqslant &(2\delta+\delta_\star)L^2(\sigma+\varepsilon)^2+\delta\Big(w^2(0)+2L^2\sigma^2 t/\varepsilon+4L^2\sigma t\Big) \\ &+\frac{4L(1+\lambda\varepsilon)\sigma\delta}{\varepsilon}tW(t)+2L\varepsilon\delta W(\min\{t,\varepsilon\}). \end{aligned}$$

5. Conclusion.

Choosing $\varepsilon := \sqrt{\sigma\max\{\sigma,t\}}$ and assuming $\lambda\sqrt{T\sigma} \leqslant 1$, we obtain

$$e^{-2\lambda t}w^2(t) \leqslant \delta w^2(0)+(14\delta+4\delta_\star)L^2\,\sigma\max\{\sigma,t\}+10\delta L\sqrt{\sigma\max\{\sigma,t\}}W(t). \tag{9.1.21}$$

Since the right hand side of (9.1.21) is an increasing function of t, (9.1.21) holds even if we substitute the left hand side with $e^{-2\lambda s}w^2(s)$ for every $s \in [0,t]$; we thus obtain the inequality

$$W^2(t) \leqslant \delta w^2(0)+(14\delta+4\delta_\star)L^2\,\sigma\max\{\sigma,t\}+10\delta L\sqrt{\sigma\max\{\sigma,t\}}W(t).$$

Using the elementary property for positive a, b

$$W^2 \leqslant a + 2bW \quad \Rightarrow \quad W \leqslant b + \sqrt{b^2 + a} \leqslant 2b + \sqrt{a}, \tag{9.1.22}$$

we eventually obtain

$$\begin{aligned} e^{-\lambda t} w(t) &\leqslant \Big(\delta w^2(0) + (14\delta + 4\delta_\star)L^2\, \sigma \max\{\sigma, t\}\Big)^{1/2} + 10\delta L\sqrt{\sigma \max\{\sigma, t\}} \\ &\leqslant \sqrt{\delta} w(0) + C(\delta) L\sqrt{\sigma \max\{\sigma, t\}}, \end{aligned}$$

with $C(\delta) := (14\,\delta + 4\,\delta_\star)^{1/2} + 10\,\delta$. □

9.1.3 *Error estimates between discrete and EVI solutions*

Theorem 9.1.6. *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1)*. If* $\mu : [0, T] \to \overline{D(\mathbf{F})}$ *is a* λ*-EVI solution and* $M_\tau \in \mathscr{M}(M_\tau^0, \tau, T, L)$*, then for every* $\delta > 1$ *there exists a constant* $C(\delta)$ *such that*

$$W_2(\mu_t, M_\tau(t)) \leqslant \Big(\sqrt{\delta}\, W_2(\mu_0, M_\tau^0) + C(\delta) L\sqrt{\tau(t+\tau)}\Big) e^{\lambda_+ t}$$

for every $t \in [0, T]$*.*

Remark 9.1.7. When $\mu_0 = M_\tau^0$ and $\lambda \leqslant 0$ we obtain the optimal error estimate

$$W_2(\mu_t, M_\tau(t)) \leqslant 13 L\sqrt{\tau(t+\tau)}.$$

Proof. We repeat the same argument of the previous proof, still assuming $\lambda > 0$, extending $M_\tau, \bar{M}_\tau, \mathbf{F}_\tau$ as in (9.1.13) and setting

$$w(r, s) := W_2(M_\tau(r), \mu_s), \quad w_\tau(r, s) := W_2(\bar{M}_\tau(r), \mu_s).$$

We use (λ-EVI) for μ_s with $\nu = \bar{M}_\tau(r)$ and $\Phi = \mathbf{F}_\tau(\max\{r, 0\})$ and (IEVI) for $M_\tau(r)$ with $\nu = \mu_s$ obtaining

$$\begin{aligned} \frac{\partial}{\partial r} \frac{e^{-2\lambda s}}{2} W_2^2(M_\tau(r), \mu_s) &\leqslant e^{-2\lambda s}\Big(\tau|\mathbf{F}_\tau(r)|_2^2 + [\mathbf{F}_\tau(r), \mu_s]_r\Big) \\ &\qquad \text{for every } s \in [0, T], r \in (-\infty, T) \\ \frac{\partial}{\partial s} \frac{e^{-2\lambda s}}{2} W_2^2(\mu_s, \bar{M}_\tau(r)) &\leqslant -e^{-2\lambda s}\, [\mathbf{F}_\tau(\max\{r, 0\}), \mu_s]_r \\ &\qquad \text{in } \mathscr{D}'(0, T),\ r \in (-\infty, T). \end{aligned}$$

Using [89, Lemma 6.15] we can sum the two contributions obtaining

$$\frac{\partial}{\partial r} e^{-2\lambda s} w^2(r, s) + \frac{\partial}{\partial s} e^{-2\lambda s} w_\tau^2(r, s) \leqslant Z_{r,s},$$

where $Z_{r,s} := (2L^2\tau + 2f(r, s))e^{-2\lambda s}$, and

$$f(r, s) := \begin{cases} L W_2(M_\tau(0), \mu_s) = L w(0, s) & \text{if } r < 0, \\ 0 & \text{if } r \geqslant 0. \end{cases}$$

Let $t \in [0,T]$ and $\varepsilon > 0$. Applying the Divergence Theorem in $Q^{\varepsilon}_{0,t}$ (see (8.2.6) and Figure 2), we get

$$\begin{aligned}\int_{t-\varepsilon}^{t} e^{-2\lambda t} w_\tau^2(r,t)\,dr \leqslant \int_{-\varepsilon}^{0} w_\tau^2(r,0)\,dr \\ + \int_0^t e^{-2\lambda s}\left(w_\tau^2(s,s) - w^2(s,s)\right)\,ds \\ + \int_0^t e^{-2\lambda s}\left(w^2(s-\varepsilon,s) - w_\tau^2(s-\varepsilon,s)\right)\,ds \\ + \iint_{Q^{\varepsilon}_{0,t}} Z_{r,s}\,drds.\end{aligned} \tag{9.1.23}$$

Using

$$w(t,t) \leqslant w(r,t) + L(t-r) \leqslant w_\tau(r,t) + L(\tau+\varepsilon) \quad \text{if } t-\varepsilon \leqslant r \leqslant t,$$

we get for every $\delta, \delta_\star > 1$ conjugate coefficients ($\delta_\star = \delta/(\delta-1)$)

$$e^{-2\lambda t} w^2(t) \leqslant \frac{\delta}{\varepsilon}\int_{t-\varepsilon}^{t} e^{-2\lambda t} w_\tau^2(r,t)\,dr + \delta_\star L^2(\tau+\varepsilon)^2. \tag{9.1.24}$$

Similarly to (9.1.15) we have

$$|w_\tau(s,s) - w(s,s)| \leqslant L\tau, \quad |w_\tau^2(s,s) - w^2(s,s)| \leqslant L\tau\Big(L\tau + 2w(s)\Big)$$

and, after an integration,

$$\int_0^t e^{-2\lambda s}\left(w_\tau^2(s,s) - w^2(s,s)\right)\,ds \leqslant L^2 t\tau^2 + 2L\tau\int_0^t e^{-2\lambda s} w(s)\,ds. \tag{9.1.25}$$

Performing the same computations for the third integral term at the r.h.s. of (9.1.23) we end up with

$$\begin{aligned}\int_0^t e^{-2\lambda s}\left(w^2(s-\varepsilon,s) - w_\tau^2(s-\varepsilon,s)\right)\,ds &\leqslant L^2 t\tau^2 + 2L\tau\int_0^t e^{-2\lambda s} w(s-\varepsilon,s)\,ds \\ &\leqslant L^2 t\tau(\tau+2\varepsilon) + 2L\tau\int_0^t e^{-2\lambda s} w(s)\,ds.\end{aligned} \tag{9.1.26}$$

Finally, since if $r < 0$ we have $f(r,s) = Lw(0,s) \leqslant L^2 s + Lw(s,s)$, then

$$\begin{aligned}\varepsilon^{-1}\iint_{Q^{\varepsilon}_{0,t}} Z_{r,s}\,drds &\leqslant 2L^2 t\tau + \varepsilon^{-1}\iint_{Q^{\varepsilon}_{0,\min\{\varepsilon,t\}}} 2f(r,s)e^{-2\lambda s}\,drds \\ &\leqslant 2L^2 t\tau + L^2\varepsilon^2 + 2L\varepsilon \sup_{0\leqslant s\leqslant \min\{\varepsilon,t\}} e^{-\lambda s} w(s).\end{aligned} \tag{9.1.27}$$

Using (9.1.25), (9.1.26), (9.1.27) in (9.1.23), we can rewrite the bound in (9.1.24) as

$$\begin{aligned}e^{-2\lambda t} w^2(t) \leqslant \delta_\star L^2(\tau+\varepsilon)^2 + \\ + \delta\Big(w^2(0) + 2L^2 t\tau^2/\varepsilon + 2L^2 t\tau + L^2\varepsilon^2 \\ + 2L\varepsilon \sup_{0\leqslant s\leqslant \min\{\varepsilon,t\}} e^{-\lambda s} w(s)\Big) \\ + \frac{4\delta L\tau}{\varepsilon}\int_0^t e^{-2\lambda s} w(s)\,ds.\end{aligned}$$

Choosing $\varepsilon := \sqrt{\tau \max\{\tau, t\}}$ we get

$$\begin{aligned} e^{-2\lambda t} w^2(t) &\leqslant 4\,\delta_\star L^2\,\tau \max\{\tau,t\} + \delta\Big(w^2(0) + 5L^2\,\tau\max\{\tau,t\}\Big) \\ &\quad + 6\,\delta L\sqrt{\tau\max\{\tau,t\}}\sup_{0\leqslant s\leqslant t} e^{-\lambda s} w(s). \end{aligned}$$

A further application of (9.1.22) yields

$$\begin{aligned} e^{-\lambda t} w(t) &\leqslant \Big(\delta w^2(0) + (5\delta + 4\delta_\star)L^2\,\tau\max\{\tau,t\}\Big)^{1/2} + 6\delta L\sqrt{\tau\max\{\tau,t\}} \\ &\leqslant \sqrt{\delta} w(0) + C(\delta) L\sqrt{t+\tau}\sqrt{\tau}, \end{aligned}$$

with $C(\delta) := (5\delta + 4\delta_\star)^{1/2} + 6\delta$. □

We show now that the limit curve as $\tau \downarrow 0$ of the family $(M_\tau)_{\tau>0}$ as in (9.1.2) is a λ-EVI solution of (8.1.1).

Theorem 9.1.8. *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1) *and let* $n \mapsto \tau(n)$ *be a vanishing sequence of time steps, let* $(\mu_{0,n})_{n\in\mathbb{N}}$ *be a sequence in* $D(\mathbf{F})$ *converging to* $\mu_0 \in \overline{D(\mathbf{F})}$ *in* $\mathcal{P}_2(\mathbb{H})$ *and let* $M_n \in \mathscr{M}(\mu_{0,n}, \tau(n), T, L)$*. Then* M_n *is uniformly converging to a Lipschitz continuous limit curve* $\mu : [0,T] \to \overline{D(\mathbf{F})}$ *which is a* λ*-EVI solution starting from* μ_0*.*

Proof. Theorem 9.1.5 shows that M_n is a Cauchy sequence in $C([0,T]; \overline{D(\mathbf{F})})$, so that there exists a unique limit curve μ as $n \to \infty$. Moreover, μ is also L-Lipschitz and, recalling (9.1.7), we have that μ is also the uniform limit of $\bar{M}_{\tau(n)}$.
Let us fix a reference measure $\nu \in D(\mathbf{F})$ and $\Phi \in \mathbf{F}[\nu]$. The (IEVI) and the λ-dissipativity of $\mathbf{F}$ yield

$$\begin{aligned} \frac{d}{dt}\frac{1}{2} W_2^2(M_n(t), \nu) &\leqslant \tau(n) |\mathbf{F}_{\tau(n)}(t)|_2^2 + \left[\mathbf{F}_{\tau(n)}, \nu\right]_r \\ &\leqslant \tau(n)\, L^2 + \lambda W_2^2(\bar{M}_{\tau(n)}(t), \nu) - \left[\Phi, \bar{M}_{\tau(n)}(t)\right]_r \end{aligned}$$

for a.e. $t \in [0,T]$. Integrating the above inequality in $(t, t+h) \subset [0,T]$ we get

$$\begin{aligned} &\frac{W_2^2(M_n(t+h), \nu) - W_2^2(M_n(t), \nu)}{2h} \leqslant \\ &\tau(n)L^2 + \frac{1}{h}\int_t^{t+h} \Big(\lambda W_2^2(\bar{M}_{\tau(n)}(s), \nu) - \left[\Phi, \bar{M}_{\tau(n)}(s)\right]_r\Big)\, ds. \end{aligned} \tag{9.1.28}$$

Notice that as $n \to +\infty$, by (9.1.7), we have

$$\liminf_{n\to+\infty} \left[\Phi, \bar{M}_{\tau(n)}(s)\right]_r \geqslant [\Phi, \mu_s]_r$$

for every $s \in [0,T]$, together with the uniform bound given by

$$\left|\left[\Phi, \bar{M}_{\tau(n)}(s)\right]_r\right| \leqslant \frac{1}{2} W_2^2(\bar{M}_{\tau(n)}(s), \nu) + \frac{1}{2}|\Phi|_2^2$$

for every $s \in [0,T]$. Thanks to Fatou's Lemma and the uniform convergence given by Theorem 9.1.5, we can pass to the limit as $n \to +\infty$ in (9.1.28) obtaining

$$\frac{W_2^2(\mu_{t+h}, \nu) - W_2^2(\mu_t, \nu)}{2h} \leqslant \frac{1}{h}\int_t^{t+h} \Big(\lambda W_2^2(\mu_s, \nu) - [\Phi, \mu_s]_r\Big)\, ds.$$

A further limit as $h \downarrow 0$ yields

$$\frac{1}{2}\frac{d^+}{dt} W_2^2(\mu_t, \nu) \leqslant \lambda W_2^2(\mu_t, \nu) - [\Phi, \mu_t]_r$$

which provides (λ-EVI). □

9.2 CONSEQUENCES OF THE SOLVABILITY OF THE EXPLICIT EULER SCHEME

In the following Theorem we collect the results obtained in Subsection 9.1. We stress that in the next statement $A(\delta)$ solely depend on δ (in particular, it is independent of $\lambda, L, T, \tau, \eta, M_\tau, M_\eta$).

Theorem 9.2.1. *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1).

1. *For every* $\mu_0, \mu_0' \in D(\mathbf{F})$, *every* $M_\tau \in \mathscr{M}(\mu_0, \tau, T, L)$, $M_\tau' \in \mathscr{M}(\mu_0', \tau, T, L)$ *with* $\tau\lambda_+ \leqslant 2$ *we have*

$$W_2(M_\tau(t), M_\tau'(t)) \leqslant e^{\lambda t} W_2(\mu_0, \mu_0') + 8L\sqrt{t\tau}\left(1 + |\lambda|\sqrt{t\tau}\right)e^{\lambda_+ t} \quad (9.2.1)$$

   for every $t \in [0, T]$.

2. *For every* $\delta > 1$ *there exists a constant* $A(\delta)$ *such that if* $M_\tau \in \mathscr{M}(M_\tau^0, \tau, T, L)$ *and* $M_\eta \in \mathscr{M}(M_\eta^0, \eta, T, L)$ *with* $\lambda_+(\tau + \eta) \leqslant 1$ *then*

$$W_2(M_\tau(t), M_\eta(t)) \leqslant \left(\delta\, W_2(M_\tau^0, M_\eta^0) + A(\delta) L \sqrt{(\tau + \eta)(t + \tau + \eta)}\right) e^{\lambda_+ t}$$

   for every $t \in [0, T]$.

3. *For every* $\delta > 1$ *there exists a constant* $A(\delta)$ *such that if* $\mu : [0, T] \to \overline{D(\mathbf{F})}$ *is a* λ*-EVI solution and* $M_\tau \in \mathscr{M}(M_\tau^0, \tau, T, L)$ *then*

$$W_2(\mu_t, M_\tau(t)) \leqslant \left(\delta\, W_2(\mu_0, M_\tau^0) + A(\delta) L \sqrt{\tau(t + \tau)}\right) e^{\lambda_+ t} \quad (9.2.2)$$

   for every $t \in [0, T]$.

4. *If* $n \mapsto \tau(n)$ *is a vanishing sequence of time steps,* $(\mu_{0,n})_{n\in\mathbb{N}}$ *is a sequence in* $D(\mathbf{F})$ *converging to* $\mu_0 \in \overline{D(\mathbf{F})}$ *in* $\mathcal{P}_2(\mathbb{H})$ *and* $M_n \in \mathscr{M}(\mu_{0,n}, \tau(n), T, L)$, *then* M_n *is uniformly converging to a Lipschitz continuous limit curve* $\mu : [0, T] \to \overline{D(\mathbf{F})}$ *which is a* λ*-EVI solution starting from* μ_0.

Definition 9.2.2 (Local and global solvability of (EE))**.** We say that the Explicit Euler Scheme (EE) associated to a MPVF $\mathbf{F}$ is *locally solvable* at $\mu_0 \in D(\mathbf{F})$ if there exist strictly positive constants $\boldsymbol{\tau}, T, L$ such that $\mathscr{E}(\mu_0, \tau, T, L)$ is not empty for every $\tau \in (0, \boldsymbol{\tau})$.
We say that (EE) is *globally solvable* at $\mu_0 \in D(\mathbf{F})$ if for every $T > 0$ there exist strictly positive constants $\boldsymbol{\tau}, L$ such that $\mathscr{E}(\mu_0, \tau, T, L)$ is not empty for every $\tau \in (0, \boldsymbol{\tau})$.

If we work under the assumption that the Explicit Euler scheme is locally solvable, then Theorem 9.2.1 is very useful to treat local existence and uniqueness of λ-EVI solutions.

Given $T \in (0, +\infty]$ and $\mu : [0, T) \to \mathcal{P}_2(\mathbb{H})$ we denote by $|\dot{\mu}_t|_+$ the right upper metric derivative

$$|\dot{\mu}_t|_+ := \limsup_{h \downarrow 0} \frac{W_2(\mu_{t+h}, \mu_t)}{h}.$$

Theorem 9.2.3 (Local existence and uniqueness). *Let $\mathbf{F}$ be a λ-dissipative MPVF according to (7.5.1).*

(a) If the Explicit Euler Scheme is locally solvable at $\mu_0 \in D(\mathbf{F})$, then there exists $T > 0$ and a unique Lipschitz continuous λ-EVI solution $\mu : [0, T] \to \overline{D(\mathbf{F})}$ starting from μ_0, satisfying

$$t \mapsto e^{-\lambda t} |\dot{\mu}_t|_+ \quad \textit{is decreasing in } [0, T). \tag{9.2.3}$$

If $\mu' : [0, T'] \to \overline{D(\mathbf{F})}$ is any other λ-EVI solution starting from μ_0 then $\mu_t = \mu'_t$ if $0 \leqslant t \leqslant \min\{T, T'\}$.

(b) If the Explicit Euler Scheme is locally solvable in $D(\mathbf{F})$ and

$$\begin{aligned} &\textit{for any local } \lambda\textit{-EVI solution } \mu \textit{ starting from } \mu_0 \in D(\mathbf{F}) \\ &\quad \textit{there exists } \delta > 0 : \quad t \in [0, \delta] \quad \Rightarrow \quad \mu_t \in D(\mathbf{F}), \end{aligned} \tag{9.2.4}$$

then for every $\mu_0 \in D(\mathbf{F})$ there exist a unique maximal time $T \in (0, \infty]$ and a unique strict locally Lipschitz continuous λ-EVI solution $\mu : [0, T) \to D(\mathbf{F})$ starting from μ_0, which satisfies (9.2.3) and

$$T < \infty \quad \Rightarrow \quad \lim_{t \uparrow T} \mu_t \notin D(\mathbf{F}). \tag{9.2.5}$$

Any other λ-EVI solution $\mu' : [0, T') \to \overline{D(\mathbf{F})}$ starting from μ_0 coincides with μ in $[0, \min\{T, T'\})$.

Proof.
(a) Let $\boldsymbol{\tau}, T, L$ positive constants such that $\mathscr{E}(\mu_0, \tau, T, L)$ is not empty for every $\tau \in (0, \boldsymbol{\tau})$. Thanks to Theorem 9.2.1(2), the family $M_\tau \in \mathscr{E}(\mu_0, \tau, T, L)$ satisfies the Cauchy condition in $C([0, T]; \mathcal{P}_2(\mathbb{H}))$ so that there exists a unique limit curve

$$\mu = \lim_{\tau \downarrow 0} M_\tau$$

which is also Lipschitz in time, thanks to the a-priori bound (9.1.6). Theorem 9.2.1(4) shows that μ is a λ-EVI solution starting from μ_0 and the estimate (9.2.2) of Theorem 9.2.1(3) shows that any other λ-EVI solution in an interval $[0, T']$ starting from μ_0 should coincide with μ in the interval $[0, \min\{T', T\}]$.

Let us now check (9.2.3): we fix s, t such that $0 \leqslant s < t < T$ and $h \in (0, T - t)$, and we set

$$s_\tau := \tau \lfloor s/\tau \rfloor \quad \text{and} \quad h_\tau := \tau \lfloor h/\tau \rfloor.$$

The curves

$$r \mapsto M_\tau(s_\tau + r) \quad \text{and} \quad r \mapsto M_\tau(s_\tau + h_\tau + r)$$

belong to $\mathscr{M}(M_\tau(s_\tau), \tau, t-s, L)$ and $\mathscr{M}(M_\tau(s_\tau + h_\tau), \tau, t-s, L)$, so that (9.2.1) yields

$$W_2(M_\tau(s_\tau + t - s), M_\tau(s_\tau + h_\tau + (t-s))) \leqslant e^{\lambda(t-s)} W_2(M_\tau(s_\tau), M_\tau(s_\tau + h_\tau)) + B\sqrt{\tau},$$

for $B = B(\lambda, L, \boldsymbol{\tau}, T)$. Passing to the limit as $\tau \downarrow 0$ we get

$$W_2(\mu_t, \mu_{t+h}) \leqslant e^{\lambda(t-s)} W_2(\mu_s, \mu_{s+h}).$$

Dividing by h and passing to the limit as $h \downarrow 0$ we get (9.2.3).

(b) Let us call $\mathcal{S}$ the collection of λ-EVI solutions $\mu : [0, S) \to D(\mathbf{F})$ starting from μ_0 with values in $D(\mathbf{F})$ and defined in some interval $[0, S)$, $S = S(\mu)$. Thanks to (9.2.4) and the previous claim the set $\mathcal{S}$ is not empty.

It is also easy to check that two curves $\mu', \mu'' \in \mathcal{S}$ coincide in the common domain $[0, S)$ with

$$S := \min \left\{ S(\mu'), S(\mu'') \right\}.$$

Indeed, the set

$$\left\{ t \in [0, S) : \mu'_r = \mu''_r \text{ if } 0 \leqslant r \leqslant t \right\}$$

contains $t = 0$, is closed since μ', μ'' are continuous, and it is also open since, if $\mu' = \mu''$ in $[0, t]$, then the previous claim and the fact that $\mu'_t = \mu''_t \in D(\mathbf{F})$ show that $\mu' = \mu''$ also in a right neighborhood of t. Since $[0, S)$ is connected, we conclude that $\mu' = \mu''$ in $[0, S)$.

We can thus define

$$T := \sup \left\{ S(\mu) : \mu \in \mathcal{S} \right\},$$

obtaining that there exists a unique λ-EVI solution μ starting from μ_0 and defined in $[0, T)$ with values in $D(\mathbf{F})$.

If $T < \infty$, since μ is Lipschitz in $[0, T)$ thanks to (9.2.3), we know that there exists the limit

$$\bar{\mu} := \lim_{t \uparrow T} \mu_t$$

in $\mathcal{P}_2(\mathbb{H})$. If $\bar{\mu} \in D(\mathbf{F})$ we can extend μ to a λ-EVI solution with values in $D(\mathbf{F})$ and defined in an interval $[0, T')$ with $T' > T$, which contradicts the maximality of T. □

A set A in a metric space X is locally closed if every point of A has a neighborhood U such that $A \cap U = \bar{A} \cap U$. Equivalently, A is the intersection of an open and a closed subset of X. In particular, open or closed sets are locally closed.

We refer to Definition 8.1.1 for the notion of strict EVI solutions, used in the following.

Corollary 9.2.4. *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1) *for which the Explicit Euler Scheme is locally solvable in* $D(\mathbf{F})$*. If* $D(\mathbf{F})$ *is locally closed then for every* $\mu_0 \in D(\mathbf{F})$ *there exists a unique maximal strict and locally Lipschitz continuous* λ*-EVI solution* $\mu : [0, T) \to D(\mathbf{F})$, $T \in (0, +\infty]$, *satisfying* (9.2.5).

We now present a few situation where the Explicit Euler scheme is locally solvable. The constraints in the construction of the explicit Euler scheme are feasible if at each measure M^n_τ, $0 \leqslant n < N(T, \tau)$, the set $\mathrm{Adm}_{\tau,L}(M^n_\tau)$ defined by

$$\mathrm{Adm}_{\tau,L}(\mu) := \left\{\Phi \in \mathbf{F}[\mu] : |\Phi|_2 \leqslant L \quad \text{and} \quad \exp^\tau_\sharp \Phi \in D(\mathbf{F})\right\}, \quad \mu \in \mathcal{P}_2(\mathbb{H}),$$

is not empty. If $D(\mathbf{F})$ is open and $\mathbf{F}$ is locally bounded, then it is easy to check that the Explicit Euler scheme is locally solvable (see Lemma 9.2.5). We will adopt the following notation:

$$|\mathbf{F}|_2(\mu) := \inf\left\{|\Phi|_2 : \Phi \in \mathbf{F}[\mu]\right\} \quad \text{for every } \mu \in D(\mathbf{F}), \tag{9.2.6}$$

and we will also introduce the upper semicontinuous envelope $|\mathbf{F}|_{2\star}$ of the function $|\mathbf{F}|_2$: i.e.

$$\begin{aligned}|\mathbf{F}|_{2\star}(\mu) &:= \inf_{\delta>0} \sup\left\{|\mathbf{F}|_2(\nu) : \nu \in D(\mathbf{F}),\ W_2(\nu, \mu) \leqslant \delta\right\}\\ &= \sup\left\{\limsup_{k\to\infty} |\mathbf{F}|_2(\mu_k) : \mu_k \in D(\mathbf{F}),\ \mu_k \to \mu \text{ in } \mathcal{P}_2(\mathbb{H})\right\}.\end{aligned}$$

Lemma 9.2.5. *If* $\mathbf{F}$ *is a* λ*-dissipative MPVF according to* (7.5.1), $\mu_0 \in \mathrm{Int}(D(\mathbf{F}))$ *and* $\mathbf{F}$ *is bounded in a neighborhood of* μ_0*, i.e. there exists* $\rho > 0$ *such that* $|\mathbf{F}|_2$ *is bounded in* $B(\mu_0, \rho)$*, then the Explicit Euler scheme is locally solvable at* μ_0 *and the locally Lipschitz continuous solution* μ *given by Theorem* 9.2.3*(a) satisfies*

$$|\dot\mu_t|_+ \leqslant e^{\lambda t}|\mathbf{F}|_{2\star}(\mu_0) \quad \textit{for all} \ \ t \in [0, T). \tag{9.2.7}$$

In particular, if $D(\mathbf{F})$ *is open and* $\mathbf{F}$ *is locally bounded, for every* $\mu_0 \in D(\mathbf{F})$ *there exists a unique maximal locally Lipschitz continuous* λ*-EVI solution* $\mu : [0, T) \to \mathcal{P}_2(\mathbb{H})$ *satisfying* (9.2.5) *and* (9.2.7).

Proof. Let $\mu_0 \in \mathrm{Int}(D(\mathbf{F}))$ and let $\rho, L > 0$ so that $|\mathbf{F}|_2(\mu) < L$ for every $\mu \in B(\mu_0, \rho)$. We set

$$T := \rho/(2L) \quad \text{and} \quad \boldsymbol{\tau} := \min\{T, 1\}$$

and we perform a simple induction argument to prove that

$$W_2(M^n_\tau, \mu_0) \leqslant Ln\tau < \rho$$

if $n \leqslant N(T, \tau)$, so that we can always find an element $\Phi^n_\tau \in \mathrm{Adm}_{\tau,L}(M^n_\tau)$. In fact, if $W_2(M^n_\tau, \mu_0) < Ln\tau$ and $n < N(T, \tau)$ then

$$W_2(M^{n+1}_\tau, \mu_0) \leqslant W_2(M^{n+1}_\tau, M^n_\tau) + W_2(M^n_\tau, \mu_0) \leqslant L(n+1)\tau.$$

The property in (9.2.3) shows that $|\dot\mu_t|_+ \leqslant Le^{\lambda t}$ for every $L > |\mathbf{F}|_{2\star}(\mu_0)$, so that we obtain (9.2.7). □

Another example is related to measures with bounded support.

Proposition 9.2.6. *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1)*. Assume that* $D(\mathbf{F}) \subset \mathcal{P}_b(\mathbb{H})$ *and for every* $\mu_0 \in D(\mathbf{F})$ *there exist* $\rho > 0$, $L > 0$ *such that, for every* $\mu \in \mathcal{P}_b(\mathbb{H})$ *with* $\text{supp}(\mu) \subset \text{supp}(\mu_0) + B(0, \rho)$*, there exists* $\Phi \in \mathbf{F}[\mu]$ *such that*

$$\text{supp}(v_\sharp \Phi) \subset B(0, L).$$

Then for every $\mu_0 \in D(\mathbf{F})$ *there exists* $T \in (0, +\infty]$ *and a unique maximal strict and locally Lipschitz continuous* λ*-EVI solution* $\mu : [0, T) \to D(\mathbf{F})$ *satisfying* (9.2.5)*.*

Proof. Arguing as in the proof of Lemma 9.2.5, it is easy to check that setting $T := \rho/4L$, $\boldsymbol{\tau} = \min\{T, 1\}$, we can find a discrete solution $(M_\tau, F_\tau) \in \mathscr{E}(\mu_0, \tau, T, L)$ satisfying the more restrictive condition

$$\text{supp}(M_\tau^n) \subset \text{supp}(\mu_0) + B(0, Ln\tau) \subset \text{supp}(\mu_0) + B(0, \rho/2), \ \text{supp}(v_\sharp \Phi_\tau^n) \subset B(0, L).$$

So that the Explicit Euler scheme is locally solvable and M_τ satisfies the uniform bound

$$\text{supp}(M_\tau(t)) \subset \text{supp}(\mu_0) + B(0, \rho/2) \tag{9.2.8}$$

for every $t \in [0, T]$. Theorem 9.2.3 then yields the existence of a local solution, and Theorem 9.2.1(3) shows that the local solution satisfies the same bound (9.2.8) on the support, so that (9.2.4) holds. □

9.2.1 *Stability and uniqueness*

The following stability result assumes that the Explicit Euler scheme is locally solvable in $D(\mathbf{F})$.

Theorem 9.2.7 (Uniqueness and Stability)**.** *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1) *such that the Explicit Euler scheme is locally solvable in* $D(\mathbf{F})$*, and let* $\mu^1, \mu^2 : [0, T) \to \overline{D(\mathbf{F})}$, $T \in (0, +\infty]$*, be* λ*-EVI solutions to* (8.1.1)*. If* μ^1 *is strict, then*

$$W_2(\mu_t^1, \mu_t^2) \leqslant W_2(\mu_0^1, \mu_0^2)\, e^{\lambda_+ t} \quad \textit{for every } t \in [0, T). \tag{9.2.9}$$

In particular, if $\mu_0^1 = \mu_0^2$ *then* $\mu^1 \equiv \mu^2$ *in* $[0, T)$*.*
If μ^1, μ^2 *are both strict, then*

$$W_2(\mu_t^1, \mu_t^2) \leqslant W_2(\mu_0^1, \mu_0^2)\, e^{\lambda t} \quad \textit{for every } t \in [0, T). \tag{9.2.10}$$

Proof. In order to prove (9.2.9), let us fix $t \in (0, T)$. Since the Explicit Euler scheme is locally solvable and $\mu_t^1 \in D(\mathbf{F})$, there exist $\boldsymbol{\tau}, \delta, L$ such that $\mathscr{M}(\mu_t^1, \tau, \delta, L)$ is not empty for every $\tau \in (0, \boldsymbol{\tau})$. If $M_\tau^1 \in \mathscr{M}(\mu_t^1, \tau, \delta, L)$, then (9.2.2) yields

$$\begin{aligned} W_2(\mu_{t+h}^1, \mu_{t+h}^2) &\leqslant W_2(M_\tau^1(h), \mu_{t+h}^2) + W_2(M_\tau^1(h), \mu_{t+h}^1) \\ &\leqslant \delta\, W_2(\mu_t^1, \mu_t^2) e^{\lambda_+ h} + B\sqrt{\tau} \quad \text{if } 0 \leqslant h \leqslant \delta, \end{aligned}$$

for $B = B(\lambda, L, \boldsymbol{\tau}, \delta)$ Passing to the limit as $\tau \downarrow 0$ we obtain

$$W_2(\mu_{t+h}^1, \mu_{t+h}^2) \leqslant \delta\, W_2(\mu_t^1, \mu_t^2) e^{\lambda_+ h}$$

and a further limit as $\delta \downarrow 1$ yields

$$W_2(\mu^1_{t+h}, \mu^2_{t+h}) \leqslant W_2(\mu^1_t, \mu^2_t) e^{\lambda_+ h}$$

for every $h \in [0, \delta]$, which implies that the map $t \mapsto e^{-\lambda_+ t} W_2(\mu^1_t, \mu^2_t)$ is decreasing in $[t, t+\delta]$. Since t is arbitrary, we obtain (9.2.9).

In order to prove the estimate (9.2.10) (which is better than (9.2.9) when $\lambda < 0$), we argue in a similar way, using (9.2.1).

As before, for a given $t \in (0, T)$, since the Explicit Euler scheme is locally solvable and $\mu^1_t, \mu^2_t \in D(\mathbf{F})$, there exist $\boldsymbol{\tau}, \delta, L$ such that $\mathscr{M}(\mu^1_t, \tau, \delta, L)$ and $\mathscr{M}(\mu^2_t, \tau, \delta, L)$ are not empty for every $\tau \in (0, \boldsymbol{\tau})$. If $M^i_\tau \in \mathscr{M}(\mu^i_t, \tau, \delta, L)$, for $i = 1, 2$, (9.2.1) and (9.2.2) then yield

$$\begin{aligned} W_2(\mu^1_{t+h}, \mu^2_{t+h}) &\leqslant W_2(\mu^1_{t+h}, M^1_\tau(h)) + W_2(M^1_\tau(h), M^2_\tau(h)) + W_2(\mu^2_{t+h}, M^2_\tau(h)) \\ &\leqslant e^{\lambda h} W_2(\mu^1_t, \mu^2_t) + B\sqrt{\tau} \end{aligned}$$

if $0 \leqslant h \leqslant \delta$, with $B = B(\lambda, L, \boldsymbol{\tau}, \delta)$. Passing to the limit as $\tau \downarrow 0$ we obtain

$$W_2(\mu^1_{t+h}, \mu^2_{t+h}) \leqslant e^{\lambda h} W_2(\mu^1_t, \mu^2_t)$$

which implies that the map $t \mapsto e^{-\lambda t} W_2(\mu^1_t, \mu^2_t)$ is decreasing in $(0, T)$. □

Corollary 9.2.8 (Local Lipschitz estimate). *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1) *and let* $\mu : (0, T) \to \overline{D(\mathbf{F})}$, $T \in (0, +\infty]$, *be a* λ*-EVI solution to* (8.1.1). *If at least one of the following two conditions holds*

(a) μ *is strict and* (EE) *is locally solvable in* $D(\mathbf{F})$,

(b) μ *is locally absolutely continuous and* (7.7.6) *holds,*

then μ *is locally Lipschitz and*

$$t \mapsto e^{-\lambda t} |\dot{\mu}_t|_+ \quad \text{is decreasing in } (0, T). \tag{9.2.11}$$

Proof. Since for every $h > 0$ the curve $t \mapsto \mu_{t+h}$ is a λ-EVI solution, (9.2.10) yields

$$e^{-\lambda(t-s)} W_2(\mu_{t+h}, \mu_t) \leqslant W_2(\mu_{s+h}, \mu_s)$$

for every $0 < s < t$. Dividing by h and taking the limsup as $h \downarrow 0$, we get (9.2.11), which in turn shows the local Lipschitz character of μ. □

9.2.2 *Global existence and generation of* λ*-flows*

Here we treat the existence of global solutions and the generation of λ-flows.

Theorem 9.2.9 (Global existence). *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1). *If the Explicit Euler Scheme is globally solvable at* $\mu_0 \in D(\mathbf{F})$, *then there exists a unique global and locally Lipschitz continuous* λ*-EVI solution* $\mu : [0, \infty) \to \overline{D(\mathbf{F})}$ *starting from* μ_0.

Proof. We can argue as in the proof of Theorem 9.2.3(a), observing that the global solvability of (EE) allows for the construction of a limit solution on every interval $[0, T]$, $T > 0$. □

Let us provide a simple condition ensuring global solvability; we will use the following discrete Gronwall estimate.

Lemma 9.2.10 (Discrete Gronwall inequality). *Let $\alpha \geqslant 0$, $y \geqslant 0$, $\tau > 0$ and $N \in \mathbb{N}$ with $N > 0$. If a sequence $(x_n)_{n\in\mathbb{N}}$ of positive real numbers satisfies*

$$x_{n+1} - x_n \leqslant \tau y + \tau\alpha x_n, \tag{9.2.12}$$

for any $0 \leqslant n \leqslant N$, then

$$x_n \leqslant (x_0 + \tau n y)e^{\alpha n \tau},$$

for any $0 \leqslant n \leqslant N + 1$.

Proof. We treat only the non trivial case $n \geqslant 1$ and $\alpha > 0$; we will repeatedly use the elementary inequality

$$1 + x \leqslant e^x \tag{9.2.13}$$

for every $x \in \mathbb{R}$. Multiplying (9.2.12) written for $n = k \in \{0, \dots, N\}$ by $e^{-\alpha\tau(k+1)}$, we obtain

$$e^{-\alpha\tau(k+1)}x_{k+1} \leqslant \tau y e^{-\alpha\tau(k+1)} + x_k(1 + \tau\alpha)e^{-\alpha\tau(k+1)} \leqslant \tau y e^{-\alpha\tau(k+1)} + x_k e^{-\alpha\tau k},$$

where the last inequality comes from (9.2.13) with $x = \alpha\tau$. Let $n \in \{0, \dots, N+1\}$; we sum the previous inequality written for $k \in \{0, \dots, n-1\}$ obtaining

$$e^{-\alpha\tau n}x_n - x_0 \leqslant \tau y e^{-\alpha\tau} \sum_{k=0}^{n-1} \left(e^{-\alpha\tau}\right)^k = \tau y e^{-\alpha\tau} \frac{1 - e^{-\alpha\tau n}}{1 - e^{-\alpha\tau}}.$$

Then we get

$$\begin{aligned} x_n &\leqslant x_0 e^{\alpha\tau n} + \tau y \frac{e^{\alpha\tau n} - 1}{e^{\alpha\tau} - 1} \\ &= x_0 e^{\alpha\tau n} + \tau y n \frac{e^{\alpha\tau n} - 1}{\alpha\tau n} \frac{\alpha\tau}{e^{\alpha\tau} - 1} \\ &\leqslant x_0 e^{\alpha\tau n} + \tau y n e^{\alpha\tau n}, \end{aligned}$$

where we used again (9.2.13) in the last step. □

Proposition 9.2.11. *Let $\mathbf{F}$ be a λ-dissipative MPVF according to (7.5.1). Assume that for every $R > 0$ there exist $M = M(R) > 0$ and $\bar\tau = \bar\tau(R) > 0$ such that, for every $\mu \in D(\mathbf{F})$ with $m_2(\mu) \leqslant R$ and every $0 < \tau \leqslant \bar\tau$,*

$$\text{there exists } \Phi \in \mathbf{F}[\mu] \text{ s.t. } |\Phi|_2 \leqslant M(R) \text{ and } \exp_\sharp^\tau \Phi \in D(\mathbf{F}). \tag{9.2.14}$$

Then the Explicit Euler scheme is globally solvable in $D(\mathbf{F})$. More precisely, if for a given $\mu_0 \in D(\mathbf{F})$ with $\Psi_0 \in \mathbf{F}[\mu_0]$, $m_0 := m_2(\mu_0)$, and we set

$$R := m_0 + \left(|\Psi_0|_2 + 1\right)\sqrt{2T}e^{(1+2\lambda_+)T}, \quad L := M(R), \quad \boldsymbol{\tau} = \min\left\{\frac{1}{L^2}, \bar\tau(R), T\right\}, \tag{9.2.15}$$

then for every $\tau \in (0, \boldsymbol{\tau}]$ the set $\mathscr{E}(\mu_0, \tau, T, L)$ is not empty.

Proof. We want to prove by induction that for every integer $N \leqslant N(T,\tau)$, (EE) has a solution up to the index N satisfying the upper bound

$$m_2(M_\tau^N) \leqslant R, \tag{9.2.16}$$

corresponding to the constants R, L given by (9.2.15). For $N = 0$ the statement is trivially satisfied. Assuming that $0 \leqslant N < N(T,\tau)$ and elements (M_τ^n, Φ_τ^n), $0 \leqslant n < N$, M_τ^N, are given satisfying (EE) and (9.2.16), we want to show that we can perform a further step of the Euler Scheme so that (EE) is solvable up to the index $N+1$ and $m_2(M_\tau^{N+1}) \leqslant R$.

Notice that by the induction hypothesis, for $n = 0, \dots, N-1$, we have $|\Phi_\tau^n|_2 \leqslant L$; since $m_2(M_\tau^N) \leqslant R$, by (9.2.14) we can select $\Phi_\tau^N \in \mathbf{F}[M_\tau^N]$ with $|\Phi_\tau^N|_2 \leqslant L$ such that $M_\tau^{N+1} = \exp_\sharp^\tau \Phi_\tau^N \in D(\mathbf{F})$. Using (9.1.10) with $\nu = \mu_0$, the λ-dissipativity with $\Psi_0 \in \mathbf{F}[\mu_0]$

$$[\Phi_\tau^n, \mu_0]_r \leqslant \lambda W_2^2(M_\tau^n, \mu_0) - [\Psi_0, M_\tau^n]_r\,,$$

and the bound

$$-[\Psi_0, M_\tau^n]_r \leqslant \frac{1}{2} W_2^2(M_\tau^n, \mu_0) + \frac{1}{2}|\Psi_0|_2^2,$$

we end up with

$$\frac{1}{2} W_2^2(M_\tau^{n+1}, \mu_0) - \frac{1}{2} W_2^2(M_\tau^n, \mu_0) \leqslant \frac{\tau^2}{2} L^2 + \tau \left(\frac{1}{2} + \lambda_+\right) W_2^2(M_\tau^n, \mu_0) + \frac{\tau}{2}|\Psi_0|_2^2,$$

for every $n \leqslant N$. Using the Gronwall estimate of Lemma 9.2.10 we get

$$\begin{aligned} W_2(M_\tau^n, \mu_0) &\leqslant \sqrt{T+\tau}\Big(|\Psi_0|_2 + \sqrt{\tau} L\Big) e^{(\frac{1}{2}+\lambda_+)(T+\tau)} \\ &\leqslant \sqrt{2T}\Big(|\Psi_0|_2 + 1\Big) e^{(1+2\lambda_+)T} \end{aligned}$$

for every $n \leqslant N+1$, so that

$$m_2(M_\tau^{N+1}) \leqslant m_0 + \sqrt{2T}\Big(|\Psi_0|_2 + 1\Big) e^{(1+2\lambda_+)T} \leqslant R. \qquad \square$$

Let us consider now the generation of λ-flows.

Definition 9.2.12. We say that the λ-dissipative MPVF $\mathbf{F}$, according to (7.5.1), *generates a λ-flow* if for every $\mu_0 \in \overline{D(\mathbf{F})}$ there exists a unique λ-EVI solution $\mu = S[\mu_0]$ starting from μ_0 and the maps $\mu_0 \mapsto S_t[\mu_0] = (S[\mu_0])_t$ induce a semigroup of Lipschitz transformations $(S_t)_{t \geqslant 0}$ of $\overline{D(\mathbf{F})}$ satisfying

$$W_2(S_t[\mu_0], S_t[\mu_1]) \leqslant e^{\lambda t} W_2(\mu_0, \mu_1) \quad \text{for every } t \geqslant 0. \tag{9.2.17}$$

Theorem 9.2.13 (Generation of a λ-flow)**.** *Let $\mathbf{F}$ be a λ-dissipative MPVF according to (7.5.1). If at least one of the following properties is satisfied:*

(a) the Explicit Euler Scheme is globally solvable for every μ_0 in a dense subset of $D(\mathbf{F})$;

(b) *the Explicit Euler Scheme is locally solvable in* $D(\mathbf{F})$ *and, for every* μ_0 *in a dense subset of* $D(\mathbf{F})$*, there exists a strict global* λ*-EVI solution starting from* μ_0*;*

(c) *the Explicit Euler Scheme is locally solvable in* $D(\mathbf{F})$ *and* $D(\mathbf{F})$ *is closed;*

(d) *for every* $\mu_0 \in D(\mathbf{F})$*,* $\mu_1 \in \overline{D(\mathbf{F})}$ *we have* $\Gamma_o^0(\mu_0, \mu_1|\mathbf{F}) \neq \emptyset$ *and, for every* μ_0 *in a dense subset of* $D(\mathbf{F})$*, there exists a locally absolutely continuous strict global* λ*-EVI solution starting from* μ_0*;*

(e) *for every* μ_0 *in a dense subset of* $D(\mathbf{F})$*, there exists a locally absolutely continuous solution of* (8.1.2) *starting from* μ_0*,*

then $\mathbf{F}$ *generates a* λ*-flow.*

Proof. (a) Let D be the dense subset of $D(\mathbf{F})$ for which (EE) is globally solvable. For every $\mu_0 \in D$ we define $S_t[\mu_0]$, $t \geqslant 0$, as the value at time t of the unique λ-EVI solution starting from μ_0, whose existence is guaranteed by Theorem 9.2.9.

If $\mu_0, \mu_1 \in D$, $T > 0$, we can find $\boldsymbol{\tau}, L$ such that $\mathscr{M}(\mu_0, \tau, T, L)$ and $\mathscr{M}(\mu_1, \tau, T, L)$ are not empty for every $\tau \in (0, \boldsymbol{\tau})$. We can then pass to the limit in the uniform estimate (9.2.1) for every choice of $M^i_\tau \in \mathscr{M}(\mu_i, \tau, T, L)$, $i = 0, 1$, obtaining (9.2.17) for every $\mu_0, \mu_1 \in D$.

We can then extend the map S_t to $\overline{D} = \overline{D(\mathbf{F})}$ still preserving the same property. Proposition 8.1.6 shows that for every $\mu_0 \in \overline{D(\mathbf{F})}$ the continuous curve $t \mapsto S_t[\mu_0]$ is a λ-EVI solution starting from μ_0.

Finally, if $\mu : [0, T') \to \overline{D(\mathbf{F})}$ is any λ-EVI solution starting from μ_0, we can apply (9.2.2) to get

$$W_2(\mu_t, M^1_\tau(t)) \leqslant \Big(2W_2(\mu_0, \mu_1) + C(\boldsymbol{\tau}, L, T)\sqrt{\tau}\Big)e^{\lambda_+ t} \tag{9.2.18}$$

for every $t \in [0, T]$, $T < T'$ and $\tau < \boldsymbol{\tau}$, where $C(\boldsymbol{\tau}, L, T) > 0$ is a suitable constant. Passing to the limit as $\tau \downarrow 0$ in (9.2.18) we obtain

$$W_2(\mu_t, S_t[\mu_1]) \leqslant 2W_2(\mu_0, \mu_1)e^{\lambda_+ t} \quad \text{for every } t \in [0, T]. \tag{9.2.19}$$

Choosing now a sequence $\mu_{1,n}$ in D converging to μ_0 and observing that we can choose arbitrary $T < T'$, we eventually get $\mu_t = S_t[\mu_0]$ for every $t \in [0, T')$.

(b) Let D be the dense subset of $D(\mathbf{F})$ such that there exists a global strict λ-EVI solution starting from D. By Theorem 9.2.7 such a solution is unique and the corresponding family of solution maps $S_t : D \to D(\mathbf{F})$ satisfies (9.2.17). Arguing as in the previous claim, we can extend S_t to $\overline{D(\mathbf{F})}$ still preserving (9.2.17) and the fact that $t \mapsto S_t[\mu_0]$ is a λ-EVI solution.

If μ is λ-EVI solution starting from μ_0, Theorem 9.2.7 shows that (9.2.19) holds for every $\mu_1 \in D$. By approximation we conclude that $\mu_t = S_t[\mu_0]$.

(c) Corollary 9.2.4 shows that for every initial datum $\mu_0 \in D(\mathbf{F})$ there exists a global λ-EVI solution. We can then apply Claim (b).

(d) Let D be the dense subset of $D(\mathbf{F})$ such that there exists a locally absolutely continuous strict global λ-EVI solution starting from D. By Theorem 8.2.1 such a solution is the unique locally absolutely continuous solution starting from μ_0

and the corresponding family of solution maps $S_t : D \to D(\mathbf{F})$ satisfies (9.2.17). Arguing as in the previous claim (b), we can extend S_t to $\overline{D(\mathbf{F})}$ still preserving (9.2.17) (again thanks to Theorem 8.2.1) and the fact that $t \mapsto S_t[\mu_0]$ is a λ-EVI solution.

If μ is a λ-EVI solution starting from $\mu_0 \in \overline{D(\mathbf{F})}$ and $(\mu_0^n)_{n\in\mathbb{N}} \subset D$ is a sequence converging to μ_0, we can apply Theorem 8.2.3(1) and conclude that $\mu_t = S_t[\mu_0]$.

(e) The proof follows by the same argument of the previous claim, eventually applying Theorem 8.2.3(2). □

By Lemma 9.2.5 we immediately get the following result.

Corollary 9.2.14. *If* $\mathbf{F}$ *is locally bounded* λ*-dissipative MPVF according to (7.5.1), with* $D(\mathbf{F}) = \mathcal{P}_2(\mathbb{H})$*, then for every* $\mu_0 \in \mathcal{P}_2(\mathbb{H})$ *there exists a unique global* λ*-EVI solution starting from* μ_0*.*

We show now that our notion of solution is consistent with the Hilbertian theory.

Corollary 9.2.15 (Consistency with the theory of contraction semigroups in Hilbert spaces)**.** *Let* $F \subset \mathbb{H} \times \mathbb{H}$ *be a dissipative maximal subset generating the semigroup* $(R_t)_{t\geqslant 0}$ *of nonlinear contractions* [26, Theorem 3.1]*. Let* $\mathbf{F}$ *be the dissipative MPVF according to (7.5.1), defined by*

$$\mathbf{F} := \Big\{ \Phi \in \mathcal{P}_2(\mathsf{T}\mathbb{H}) \mid \Phi \text{ is concentrated on } F \Big\}.$$

The semigroup $\mu_0 \mapsto S_t[\mu_0] := (R_t)_\sharp \mu_0$, $t \geqslant 0$*, is the 0-flow generated by* $\mathbf{F}$ *in* $\overline{D(\mathbf{F})}$*.*

Proof. Let D be the set of discrete measures $\frac{1}{n}\sum_{j=1}^n \delta_{x_j}$ with $x_j \in D(F)$. Since every $\mu_0 \in \overline{D(\mathbf{F})}$ is supported in $\overline{D(F)}$, D is dense in $\overline{D(\mathbf{F})}$. Our thesis follows by applying Theorem 9.2.13(e) if we show that for every $\mu_0^n = \frac{1}{n}\sum_{j=1}^n \delta_{x_{j,0}} \in D$ there exists a locally absolutely continuous solution $\mu^n : [0,\infty) \to D$ of (8.1.2) starting from μ_0^n.

It can be directly checked that

$$\mu_t^n := (R_t)_\sharp \mu_0^n = \frac{1}{n}\sum_{j=1}^n \delta_{x_{j,t}}, \quad x_{j,t} := R_t(x_{j,0})$$

satisfies the continuity equation with Wasserstein velocity vector $\boldsymbol{v}_t$ (defined on the finite support of μ_t^n) satisfying

$$\boldsymbol{v}_t(x_{j,t}) = \dot{x}_{j,t} = F^\circ(x_{j,t}) \quad \text{and} \quad |\boldsymbol{v}_t(x_{j,t})| \leqslant |F^\circ(x_{j,0})|$$

for every $j = 1, \cdots, n$, and a.e. $t > 0$, where F° is the minimal selection of F. It follows that

$$(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t^n \in \mathbf{F}[\mu_t^n] \quad \text{for a.e. } t > 0,$$

so that μ^n is a Lipschitz EVI solution for $\mathbf{F}$ starting from μ_0^n. We can thus conclude observing that the map $\mu_0 \mapsto (R_t)_\sharp \mu_0$ is a contraction in $\mathcal{P}_2(\mathbb{H})$ and the curve $\mu_t^n = (R_t)_\sharp \mu_0^n$ is continuous with values in $\overline{D(\mathbf{F})}$. □

We conclude this sections with a result which is the natural refinement of Proposition 9.2.6

Theorem 9.2.16 (Generation of λ-flow). *Let* $\mathbf{F}$ *be a* λ*-dissipative MPVF according to* (7.5.1). *Assume that* $\mathcal{P}_b(\mathbb{H}) \subset D(\mathbf{F})$ *and for every* $\mu_0 \in \mathcal{P}_b(\mathbb{H})$ *there exist* $\rho > 0$ *and* $L > 0$ *such that, for every* μ *with* $\operatorname{supp}(\mu) \subset \operatorname{supp}(\mu_0) + B(0,\rho)$,

$$\textit{there exists } \Phi \in \mathbf{F}[\mu] \textit{ s.t. } \operatorname{supp}(\nu_\sharp \Phi) \subset B(0, L). \tag{9.2.20}$$

Let $\mathbf{F}_b := \mathbf{F} \cap \mathcal{P}_b(\mathsf{T}\mathbb{H})$. *If there exists* $a \geqslant 0$ *such that for every* $\Phi \in \mathbf{F}_b$

$$\operatorname{supp}(\Phi) \subset \Big\{(x,v) \in \mathsf{T}\mathbb{H} : \langle v, x\rangle \leqslant a(1+|x|^2)\Big\}, \tag{9.2.21}$$

then $\mathbf{F}$ *generates a* λ*-flow.*

Proof. It is enough to prove that $\mathbf{F}_b$ generates a λ-flow. Applying Proposition 9.2.6 to the MPVF $\mathbf{F}_b$, we know that for every $\mu_0 \in D(\mathbf{F}_b)$ there exists a unique maximal strict and locally Lipschitz continuous λ-EVI solution $\mu : [0,T) \to \mathcal{P}_b(\mathbb{H})$ driven by $\mathbf{F}_b$ and satisfying (9.2.5). We argue by contradiction, and we assume that $T < +\infty$. Notice that by (9.2.20) $\mathbf{F}$ satisfies (8.3.4), so that μ is a relaxed barycentric solution for $\mathbf{F}_b$. Since $\mu_0 \in \mathcal{P}_b(\mathbb{H})$, we know that $\operatorname{supp}(\mu_0) \subset B(0, r_0)$ for some $r_0 > 1$.

It is easy to check that (9.2.21) holds also for every $\Phi \in \overline{\operatorname{co}}(\overline{\mathbf{F}}_b)$. Moreover, setting $b := 2a$, condition (9.2.21) yields

$$\langle v, x\rangle \leqslant b|x|^2 \quad \text{for every} \quad (x,v) \in \operatorname{supp}\Phi \in \mathbf{F}_b,\ |x| \geqslant 1. \tag{9.2.22}$$

Let $\phi(r) : \mathbb{R} \to \mathbb{R}$ be any smooth increasing function such that $\phi(r) = 0$ if $r \leqslant r_0$ and $\phi(r) = 1$ if $r \geqslant r_0 + 1$, and let $\varphi(t,x) := \phi(|x|e^{-bt})$. Clearly $\varphi \in C^{1,1}(\mathbb{H} \times [0,+\infty))$, with

$$\begin{aligned}
&\nabla\varphi(t,x) = \frac{x}{|x|}\phi'(|x|e^{-bt})e^{-bt} \ \text{ if } x \neq 0,\\
&\nabla\varphi(t,0) = 0,\\
&\partial_t\varphi(t,x) = -b\phi'(|x|e^{-bt})|x|e^{-bt}.
\end{aligned}$$

We thus have for a.e. $t \in [0,T)$

$$\begin{aligned}
\frac{d}{dt}\int_{\mathbb{H}} \varphi(t,x)\,d\mu_t &= e^{-bt}\int_{\mathsf{T}\mathbb{H}} \Big(-b\phi'(|x|e^{-bt})|x| \\
&\qquad + \langle v,x\rangle|x|^{-1}\phi'(|x|e^{-bt})\Big)d\Phi_t(v,x)\\
&\leqslant e^{-bt}\int_{\mathsf{T}\mathbb{H}} \Big(-b\phi'(|x|e^{-bt})|x| \\
&\qquad + b|x|\phi'(|x|e^{-bt})\Big)d\Phi_t(v,x)\\
&= 0
\end{aligned}$$

where in the last inequality we used (9.2.22) and the fact that the integrand vanishes if $|x| \leqslant 1$. We get

$$\int_{\mathbb{H}} \varphi(t,x)\,d\mu_t = 0 \quad \text{in } [0,T);$$

this implies that $\mathrm{supp}(\mu_t) \subset B(0, (r_0 + 1)e^{bt})$ so that the limit measure μ_T belongs to $\mathcal{P}_b(\mathbb{H})$ as well, leading to a contradiction with (9.2.5) for $\mathbf{F}_b$.

We deduce that μ is a global strict λ-EVI solution for $\mathbf{F}_b$. We can then apply Theorem 9.2.13(b) to $\mathbf{F}_b$. □

9.3 LAW INVARIANT DISSIPATIVE OPERATORS IN HILBERT SPACES

The aim of this section is to study the properties of a dissipative operator defined on a Hilbert space of random variables, that is invariant w.r.t. measure preserving maps. The results obtained in this section will be applied to Section 9.4.

In this section, $(\Omega, \mathcal{B}, \mathbb{P})$ is an arbitrary fixed standard Borel probability space (see Definition 6.2.1). We denote by $S(\Omega)$ the class of $\mathcal{B}$-measurable maps $g : \Omega \to \Omega$ which are essentially injective and measure preserving, meaning that there exists a full $\mathbb{P}$-measure set $\Omega_0 \in \mathcal{B}$ such that g is injective on Ω_0 and $g_\sharp \mathbb{P} = \mathbb{P}$. Every $g \in S(\Omega)$ has an inverse $g^{-1} \in S(\Omega)$ (defined up to a $\mathbb{P}$-negligible set) such that $g^{-1} \circ g = g \circ g^{-1} = \mathbf{i}_\Omega$ $\mathbb{P}$-a.e. in Ω. Finally, we set

$$\mathcal{H} := L^2((\Omega, \mathcal{B}, \mathbb{P}); \mathbb{H}). \tag{9.3.1}$$

Definition 9.3.1. We say that a set $\mathbf{L} \subset \mathcal{H} \times \mathcal{H}$ is *invariant* if for every $g \in S(\Omega)$ it holds

$$(X, V) \in \mathbf{L} \Rightarrow (X \circ g, V \circ g) \in \mathbf{L}.$$

We refer to Section 6.1 for a review on maximal dissipative operators $\mathbf{L}$ on Hilbert spaces, in particular for the definitions of $\mathbf{J}_\tau$ and $\mathbf{L}^\circ$, denoting respectively the resolvent operator and the minimal selection of $\mathbf{L}$.

Lemma 9.3.2. *Let $\mathbf{L} \subset \mathcal{H} \times \mathcal{H}$ be an invariant maximal dissipative operator and let $g \in S(\Omega)$. Then*

1. *if $X \in \mathcal{H}$ and $\tau > 0$ it holds*

$$\mathbf{J}_\tau(X \circ g) = (\mathbf{J}_\tau X) \circ g;$$

2. *if $X \in D(\mathbf{L})$ we have*

$$\mathbf{L}^\circ[X \circ g] = \mathbf{L}^\circ[X] \circ g.$$

In particular

$$|\mathbf{J}_\tau X \circ g - \mathbf{J}_\tau X|_{\mathcal{H}} \leqslant |X \circ g - X|_{\mathcal{H}}, \tag{9.3.2}$$

for every $\tau > 0$ and $X \in \mathcal{H}$.

Proof. The identities $\mathbf{J}_\tau(X \circ g) = (\mathbf{J}_\tau X) \circ g$ and $\mathbf{L}^\circ[X \circ g] = \mathbf{L}^\circ[X] \circ g$ come from the invariance of $\mathbf{L}$ and the uniqueness property of the resolvent operator while (9.3.2) follows from the contractivity of the resolvent operator and (1). □

Since $\mathbf{L}$ is a maximal operator, there exists (see [26, Thm. 3.1]) a semigroup of contractions $(S_t)_{t\geqslant 0}$ with $S_t : D(\mathbf{L}) \to D(\mathbf{L})$ s.t. the curve $t \mapsto S_t X_0$ is the unique Lipschitz continuous solution of the evolution equation

$$\begin{cases} \dot{X}_t \in \mathbf{L}[X_t] & \text{a.e. } t > 0, \\ X_t|_{t=0} = X_0. \end{cases} \tag{9.3.3}$$

By [26, Thm. 3.1], we have

$$\frac{d}{dt}(S_t X_0) = \mathbf{L}^\circ[S_t X_0], \quad \text{for a.e. } t > 0. \tag{9.3.4}$$

Theorem 9.3.3. *Let $\mathbf{L} \subset \mathcal{H} \times \mathcal{H}$ be an invariant maximal dissipative operator, $X_0 \in D(\mathbf{L})$ and $Y_0 \in \mathcal{H}$ be such that $(X_0)_\sharp \mathbb{P} = (Y_0)_\sharp \mathbb{P}$. Then $Y_0 \in D(\mathbf{L})$ and*

$$(X_0, \mathbf{J}_\tau X_0, \mathbf{L}^\circ[X_0], S_t X_0)_\sharp \mathbb{P} = (Y_0, \mathbf{J}_\tau Y_0, \mathbf{L}^\circ[Y_0], S_t Y_0)_\sharp \mathbb{P} \quad \forall \tau > 0,\ t \geqslant 0. \tag{9.3.5}$$

Moreover

(a) for every $X \in \mathcal{H}$ and $\tau > 0$, there exists a 1-Lipschitz map $f_{X,\tau} : \mathbb{H} \to \mathbb{H}$ such that $\mathbf{J}_\tau X = f_{X,\tau} \circ X$ in $\mathcal{H}$;

(b) for every $X \in D(\mathbf{L})$, there exists a Borel function $h_X \in L^2(\mathbb{H}, X_\sharp \mathbb{P}; \mathbb{H})$ such that $\mathbf{L}^\circ[X] = h_X \circ X$ in $\mathcal{H}$.

Proof. Let $\mathfrak{N} := \{2^n \mid n \in \mathbb{N}\}$ and let $(\mathfrak{P}_N)_{N\in\mathfrak{N}}$ be a $\mathfrak{N}$-segmentation of $(\Omega, \mathcal{B}, \mathbb{P})$ as in Definition 6.2.9, whose existence is granted by Lemma 6.2.12. Let us define

$$\mathcal{H}_N := L^2((\Omega, \sigma(\mathfrak{P}_N), \mathbb{P}); \mathbb{H}), \quad N \in \mathfrak{N}, \quad \mathcal{H}_\infty := \cup_{N\in\mathfrak{N}} \mathcal{H}_N.$$

We divide the proof in several steps.
(1) *If $\tau > 0$ and $X \in \mathcal{H}_N$ for some $N \in \mathfrak{N}$, then (there exists a unique representative of) $\mathbf{J}_\tau X$ (that) belongs to $\mathcal{H}_N$ and*

$$|\mathbf{J}_\tau X(\omega') - \mathbf{J}_\tau X(\omega'')| \leqslant |X(\omega') - X(\omega'')| \quad \textit{for every } \omega', \omega'' \in \Omega. \tag{9.3.6}$$

Let $\Omega' \subset \Omega$ be a full $\mathbb{P}$-measure subset of Ω where both (6.2.3) and Lemma 6.2.14 hold for the increasing sequence $b_n = 2^n$, $n \in \mathbb{N}$ and the $L^2((\Omega, \mathcal{B}, \mathbb{P}); \mathbb{H})$ function $\mathbf{J}_\tau X$. Let us fix $k \in I_N := \{0, \dots, N-1\}$ and show that (a representative of) $\mathbf{J}_\tau X$ is almost everywhere constant on $\Omega'_{N,k} := \Omega_{N,k} \cap \Omega'$, where $\mathfrak{P}_N := \{\Omega_{N,k}\}_{k\in I_N}$ for every $N \in \mathfrak{N}$. Let $\omega', \omega'' \in \Omega'_{N,k}$ with $\omega' \neq \omega''$. For every $n \in \mathbb{N}$ there exist $k(n;\omega'), k(n;\omega'') \in I_{b_n}$ such that $\omega' \in \Omega_{b_n,k(n;\omega')}$ and $\omega'' \in \Omega_{b_n,k(n;\omega'')}$. By Lemma 6.2.14 we know that for $n \in \mathbb{N}$ sufficiently large $\Omega_{b_n,k(n;\omega')}, \Omega_{b_n,k(n;\omega'')} \subset \Omega_{N,k}$ and $\Omega_{b_n,k(n;\omega')} \cap \Omega_{b_n,k(n;\omega'')} = \emptyset$. Thus, since $\mathbb{P}(\Omega_{b_n,k(n;\omega')}) = \mathbb{P}(\Omega_{b_n,k(n;\omega'')}) = 2^{-n}$ for every $n \in \mathbb{N}$ (see Definition 6.2.9), by Corollary 6.2.7 we can find a measure preserving map $g_n \in S(\Omega)$ such that

$$(g_n)_\sharp \mathbb{P}|_{\Omega_{b_n,k(n;\omega')}} = \mathbb{P}|_{\Omega_{b_n,k(n;\omega'')}}$$

and g_n is the identity outside $\Omega_{b_n,k(n;\omega')} \cup \Omega_{b_n,k(n;\omega'')}$. By (9.3.2) we have

$$|\mathbf{J}_\tau X \circ g_n - \mathbf{J}_\tau X|_{\mathcal{H}} \leqslant |X \circ g_n - X|_{\mathcal{H}} = 0, \quad \forall n \in \mathbb{N} \text{ sufficiently large}$$

since X is constant on the whole $\Omega_{N,k}$. This implies that

$$2^{-n}\int_{\Omega_{b_n,k(n;\omega')}} J_\tau X\, d\mathbb{P} = 2^{-n}\int_{\Omega_{b_n,k(n;\omega'')}} J_\tau X\, d\mathbb{P} \quad \forall n \in \mathbb{N} \text{ sufficiently large.}$$

By definition of conditional expectation, this means that

$$\mathbb{E}_{\mathbb{P}}\left[J_\tau X \mid \sigma\left(\mathfrak{P}_{b_n}\right)\right](\omega') = \mathbb{E}_{\mathbb{P}}\left[J_\tau X \mid \sigma\left(\mathfrak{P}_{b_n}\right)\right](\omega'') \quad \forall n \in \mathbb{N} \text{ sufficiently large.}$$

Passing to the limit as $n \to +\infty$ we get by (6.2.3) that $J_\tau(\omega') = J_\tau(\omega'')$. This proves that $J_\tau X$ is $\mathbb{P}$-almost everywhere constant on $\Omega_{N,k}$; being $k \in I_N$ arbitrary, we can find a representative of $J_\tau X$ belonging to $\mathcal{H}_N$. If $\omega', \omega'' \in \Omega$ and $\omega' \in \Omega_{N,i}$, $\omega'' \in \Omega_{N,j}$, $i, j \in I_N$ we choose as $g \in S(\Omega)$ a measure preserving map induced by the permutation $\sigma \in \mathrm{Sym}(I_N)$ that swaps i and j (see Corollary 6.2.6), so that we get

$$\frac{2}{N}|J_\tau X(\omega') - J_\tau X(\omega'')|^2 \leqslant \frac{2}{N}|X(\omega') - X(\omega'')|^2$$

which yields (9.3.6).
(2) *If* $X \in D(\mathbf{L})$, *there exists a sequence* $(X_n)_n \subset D(\mathbf{L}) \cap \mathcal{H}_\infty$ *such that* $X_n \to X$ *as* $n \to +\infty$. By Theorem 6.2.8 we can find a sequence $(Y_n)_n \subset \mathcal{H}_\infty$ such that $Y_n \to X$. Define

$$X_n := J_{1/n} Y_n, \quad n \in \mathbb{N};$$

since $Y_n \in \mathcal{H}_\infty$, we have by (1) that $X_n \in D(\mathbf{L}) \cap \mathcal{H}_\infty$. Moreover

$$|X_n - X|_{\mathcal{H}} \leqslant |J_{1/n}Y_n - J_{1/n}X|_{\mathcal{H}} + |J_{1/n}X - X|_{\mathcal{H}} \leqslant |Y_n - X|_{\mathcal{H}} + |J_{1/n}X - X|_{\mathcal{H}} \to 0,$$

where we used that the resolvent operator is a contraction.
(3) *For every* $\tau > 0$ *it holds*

$$(X_0, J_\tau X_0)_\sharp \mathbb{P} = (Y_0, J_\tau Y_0)_\sharp \mathbb{P}. \tag{9.3.7}$$

By (2) we can find $(X_n)_n \subset \mathcal{H}_\infty \cap D(\mathbf{L})$ such that $X_n \to X_0$. By Proposition 6.2.18 we can find $(Y_n)_n \subset \mathcal{H}$ such that $(X_n)_\sharp \mathbb{P} = (Y_n)_\sharp \mathbb{P}$ for every $n \in \mathbb{N}$ and $Y_n \to Y_0$. Having X_n and Y_n the same discrete law, there exist measure preserving maps $g_n \in S(\Omega)$ such that $Y_n = X_n \circ g_n$ so that by Lemma 9.3.2 we get

$$(Y_n, J_\tau Y_n) = (X_n, J_\tau X_n) \circ g_n$$

which implies

$$(Y_n, J_\tau Y_n)_\sharp \mathbb{P} = (X_n, J_\tau X_n)_\sharp \mathbb{P}.$$

Passing to the limit as $n \to +\infty$ we obtain (9.3.7), by the continuity of the resolvent operator.
(4) $Y_0 \in D(\mathbf{L})$ *and* (9.3.5) *holds.* By (9.3.7) we have in particular that

$$\left|\frac{J_\tau Y_0 - Y_0}{\tau}\right|_{\mathcal{H}} = \left|\frac{J_\tau X_0 - X_0}{\tau}\right|_{\mathcal{H}} \quad \forall \tau > 0$$

so that

$$\lim_{\tau\downarrow 0}\left|\frac{\mathbf{J}_\tau Y_0 - Y_0}{\tau}\right|_{\mathcal{H}} = |\mathbf{L}^\circ[X_0]|_{\mathcal{H}} < +\infty.$$

This gives that $Y_0 \in D(\mathbf{L})$ by e.g. [26, Proposition 2.6(iv)]. By Lemma 9.3.2 we have that, for every $\tau, \sigma > 0$, $t \geqslant 0$ and $m, n \in \mathbb{N}$, it holds

$$\left(X_n, \mathbf{J}_\tau X_n, \frac{\mathbf{J}_\sigma X_n - X_n}{\sigma}, (\mathbf{J}_{t/m})^m X_n\right) = \left(Y_n, \mathbf{J}_\tau Y_n, \frac{\mathbf{J}_\sigma Y_n - Y_n}{\sigma}, (\mathbf{J}_{t/m})^m Y_n\right) \circ g_n$$

so that

$$\left(X_n, \mathbf{J}_\tau X_n, \frac{\mathbf{J}_\sigma X_n - X_n}{\sigma}, (\mathbf{J}_{t/m})^m X_n\right)_\sharp \mathbb{P} = \left(Y_n, \mathbf{J}_\tau Y_n, \frac{\mathbf{J}_\sigma Y_n - Y_n}{\sigma}, (\mathbf{J}_{t/m})^m Y_n\right)_\sharp \mathbb{P}.$$

Using the continuity of the resolvent operator, we can first pass to the limit as $n \to +\infty$ and then as $\sigma \downarrow 0$ and $m \to +\infty$, obtaining (9.3.5).
(4)*Claim (a).* Let $X \in \mathcal{H}$ and let $\tau > 0$; by Theorem 6.2.8 we can find $(X_n)_n \subset \mathcal{H}_\infty$ s.t. $X_n \to X$. By (9.3.6) we have

$$|\mathbf{J}_\tau X_n(\omega') - \mathbf{J}_\tau X_n(\omega)| \leqslant |X_n(\omega') - X_n(\omega)| \text{ for every } \omega, \omega' \in \Omega,\ n \in \mathbb{N}.$$

Let us consider two representatives of $\mathbf{J}_\tau X$ and X, a full measures set $\Omega_0 \subset \Omega$ and a subsequence $(X_{n_k})_k$ s.t. $X_{n_k}(\omega) \to X(\omega)$ and $\mathbf{J}_\tau X_{n_k}(\omega) \to \mathbf{J}_\tau X(\omega)$ for every $\omega \in \Omega_0$. Passing to the limit in the above inequality for every pair $(\omega, \omega') \in \Omega_0^2$, we obtain that

$$|\mathbf{J}_\tau X(\omega') - \mathbf{J}_\tau X(\omega)| \leqslant |X(\omega') - X(\omega)| \text{ for every } \omega, \omega' \in \Omega_0.$$

This gives the existence of a 1-Lipschitz function $f_{X,\tau} : X(\Omega_0) \to \mathbb{H}$ s.t. $(\mathbf{J}_\tau X)(\omega) = f_{X,\tau}(X(\omega))$ for every $\omega \in \Omega_0$. By Kirszbraum theorem we can extend it to the whole $\mathbb{H}$ and it is easy to check that $f_{X,\tau}$ does not depend on the chosen representative of X.
(5)*Claim (b).* Since $\tau^{-1}(\mathbf{J}_\tau X - X) \to \mathbf{L}^\circ[X]$ as $\tau \downarrow 0$, we can find a vanishing subsequence $(\tau_k)_k$, representatives X and $\mathbf{L}^\circ[X]$ and a full measure set $\Omega_0 \subset \Omega$ s.t.

$$\lim_k \frac{(\mathbf{J}_{\tau k} X)(\omega) - X(\omega)}{\tau_k} = \lim_k \frac{f_{X,\tau_k}(X(\omega)) - X(\omega)}{\tau_k} = \mathbf{L}^\circ[X](\omega) \text{ for every } \omega \in \Omega_0.$$

This means that the Borel set

$$E_X := \left\{ x \in \mathbb{H} \mid \exists \lim_k \frac{f_{X,\tau_k}(x) - x}{\tau_k} \right\}$$

contains $X(\Omega_0)$. We can thus define $h_X : \mathbb{H} \to \mathbb{H}$ as[1]

$$h_X(x) = \begin{cases} \lim_k \frac{f_{X,\tau_k}(x) - x}{\tau_k} & \text{if } x \in E_X, \\ 0 & \text{if } x \in \mathbb{H} \setminus E_X. \end{cases} \tag{9.3.8}$$

1 One can alternatively notice that $X(\Omega_0)$ is ι_X-measurable since it is a Souslin set (see [20, Theorem 7.4.1]); this means that we can find $E_1, E_2 \in \mathcal{B}(\mathbb{H})$ s.t. $E_1 \subset X(\Omega_0) \subset E_2$ and $\iota_X(E_2 \setminus E_1) = 0$. Then h_X can be defined as in (9.3.8) with E_1 in place of E_X. The equality in (9.3.9) holds then in the full $\mathbb{P}$-measure set $X^{-1}(E_1)$.

Of course h_X satisfies

$$\mathbf{L}^\circ[X](\omega) = h_X(X(\omega)) \text{ for every } \omega \in \Omega_0 \tag{9.3.9}$$

and it is easy to check that h_X does not depend on the choice of the representative X and that $h_X \in L^2(\mathbb{H}, X_\sharp\mathbb{P}; \mathbb{H})$. □

Let us conclude this section with the following simple remark: if $\mathbf{L} \subset \mathcal{H} \times \mathcal{H}$ is a maximal dissipative operator we can define the associated MPVF

$$\mathbf{F_L} := \left\{(X,V)_\sharp\mathbb{P} \mid (X,V) \in \mathbf{L}\right\} \subset \mathcal{P}_2(\mathsf{T}\mathbb{H}).$$

It is clear that $\mathbf{F_L}$ is dissipative: if $\Phi, \Psi \in \mathbf{F_L}$, we can find $(X,V),(Y,W) \in \mathbf{L}$ such that $(X,V)_\sharp\mathbb{P} = \Phi$ and $(Y,W)_\sharp\mathbb{P} = \Psi$ so that

$$[\Phi,\Psi]_r \leqslant \langle V-W, X-Y\rangle \leqslant 0$$

by dissipativity of $\mathbf{L}$.

9.4 DISSIPATIVE OPERATORS: FROM WASSERSTEIN TO HILBERT

In this section we will work under the following assumptions.

Hypothesis 9.4.1. We assume that:

(a) $\mathfrak{N}$ is a fixed unbounded directed subset of the integers w.r.t. the order relation $\prec$ as in (6.2.4).

(b) $(\Omega, \mathcal{B}, \mathbb{P}, (\mathfrak{P}_N)_{N\in\mathfrak{N}})$, with $\mathfrak{P}_N = \{\Omega_{N,n}\}_{n\in I_N}$ and $I_N := \{0,\dots,N-1\}$, is a fixed $\mathfrak{N}$-refined standard Borel probability space as in Definition 6.2.9.

(c) $\mathbf{F}$ is a fixed dissipative MPVF as in Definition 7.5.3 with $\lambda = 0$.

(d) $C \subset D(\mathbf{F}) \cap \mathcal{P}_{\mathfrak{N}}(\mathbb{H})$ is a fixed nonempty set such that $C \cap \mathcal{P}_N(\mathbb{H})$ is W_∞-relatively open in $\mathcal{P}_N(\mathbb{H})$ and convex along couplings in $\mathcal{P}_N(\mathbb{H}\times\mathbb{H})$ for every $N \in \mathfrak{N}$, where we recall that

$$\mathcal{P}_N(\mathbb{H}) := \left\{\mu \in \mathcal{P}_f(\mathbb{H}) : N\mu(A) \in \mathbb{N}\ \forall A \subset \mathbb{H}\right\} \tag{9.4.1}$$

and we define

$$\mathcal{P}_{\mathfrak{N}}(\mathbb{H}) := \bigcup_{N\in\mathfrak{N}} \mathcal{P}_N(\mathbb{H}). \tag{9.4.2}$$

We call such a set C a *$\mathfrak{N}$-core for* $\mathbf{F}$.

We denote by $\mathcal{B}_N := \sigma(\mathfrak{P}_N)$, $N \in \mathfrak{N}$, and set

$$\mathcal{H} := L^2((\Omega,\mathcal{B},\mathbb{P});\mathbb{H}), \quad \mathcal{H}_N := L^2((\Omega,\mathcal{B}_N,\mathbb{P});\mathbb{H}), \quad N \in \mathfrak{N}, \quad \mathcal{H}_\infty := \cup_{N\in\mathfrak{N}}\mathcal{H}_N \tag{9.4.3}$$

and we recall that $\mathcal{H}_\infty$ is dense in $\mathcal{H}$ by Theorem 6.2.8.

Thanks to Corollary 6.2.3 we can parametrize measures in $\mathcal{P}(\mathbb{H})$ by random variables in $(\Omega, \mathcal{B}, \mathbb{P})$. On the other hand, every element $X \in \mathcal{H}$ induces a measure $\iota_X := X_\sharp \mathbb{P} \in \mathcal{P}_2(\mathbb{H})$: the map $\iota : \mathcal{H} \to \mathcal{P}_2(\mathbb{H})$, $X \to \iota_X$ is 1-Lipschitz, since

$$W_2(\iota_X, \iota_Y) \leqslant |X - Y|_{\mathcal{H}} \quad \text{for every } X, Y \in \mathcal{H}. \tag{9.4.4}$$

Similarly, to every pair $(X, V) \in \mathcal{H} \times \mathcal{H}$ we can associate the measure $\iota^2_{X,V} := (X, V)_\sharp \mathbb{P} \in \mathcal{P}_2(\mathsf{T}\mathbb{H})$. We can identify $\mathcal{H}_N$ with the space $\mathbb{H}_N$ of maps $\boldsymbol{x} : I_N \to \mathbb{H}$ such that $X(\omega) = \boldsymbol{x}(n)$ whenever $\omega \in \Omega_{N,n}$. In this case we set $X = \mathscr{I}_N(\boldsymbol{x})$. Clearly $\iota(\mathcal{H}_N) = \mathcal{P}_N(\mathbb{H})$.

The isomorphism $\mathscr{I}_N$ preserves the scalar product on $\mathbb{H}_N$

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbb{H}_N} := N^{-1} \sum_{n=0}^{N-1} \langle \boldsymbol{x}(n), \boldsymbol{y}(n) \rangle = \mathbb{E}\big[\langle \mathscr{I}_N(\boldsymbol{x}), \mathscr{I}_N(\boldsymbol{y}) \rangle\big] = \langle \mathscr{I}_N(\boldsymbol{x}), \mathscr{I}_N(\boldsymbol{y}) \rangle_{\mathcal{H}} \tag{9.4.5}$$

for every $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{H}_N$. The conditional expectation $\Pi_N = \mathbb{E}[\cdot|\mathcal{B}_N]$ provides the orthogonal projection of an arbitrary map $X \in \mathcal{H}$ onto $\mathcal{H}_N$:

$$\Pi_N[X](\omega) = N \int_{\Omega_{N,n}} X \, d\mathbb{P} \quad \text{if } \omega \in \Omega_{N,n}. \tag{9.4.6}$$

Notice that

$$\text{if } M \mid N \text{ then } \mathcal{B}_M \subset \mathcal{B}_N \text{ and } \Pi_M = \Pi_M \circ \Pi_N. \tag{9.4.7}$$

For every $X = \mathscr{I}_N(\boldsymbol{x}) \in \mathcal{H}_N$ the probability measure $\iota_X = X_\sharp \mathbb{P}$ takes the form $\iota_X = N^{-1} \sum_{n=0}^{N-1} \delta_{\boldsymbol{x}(n)} \in \mathcal{P}_N(\mathbb{H})$ and we have

$$W_\infty(\iota_X, \iota_Y) \leqslant \max_{n \in I_N} |\boldsymbol{x}(n) - \boldsymbol{y}(n)| \leqslant N \, |X - Y|_{\mathcal{H}}. \tag{9.4.8}$$

We denote by $O_N \subset \mathbb{H}_N$ the subset of the injective maps and by $\mathcal{O}_N := \mathscr{I}_N(O_N) \subset \mathcal{H}_N$. Since the complement of O_N is the union of a finite number of proper closed subspaces with empty interior $S_{ij} := \{\boldsymbol{x} \in \mathbb{H}_N : \boldsymbol{x}(i) = \boldsymbol{x}(j)\}$, $i \neq j$, of $\mathbb{H}_N$, then O_N is open and dense in $\mathbb{H}_N$.

Every permutation $\sigma \in \mathrm{Sym}(I_N)$ acts on $\mathbb{H}_N$ via $\sigma\boldsymbol{x}(n) := \boldsymbol{x}(\sigma(n))$ and can be thus extended to $\mathcal{H}_N$ via $\sigma(\mathscr{I}_N(\boldsymbol{x})) := \mathscr{I}_N(\sigma(\boldsymbol{x}))$. It is not difficult to see that, for every $X, Y \in \mathcal{H}_N$, $\iota_X = \iota_Y$ is equivalent to $Y = \sigma X$ for some $\sigma \in \mathrm{Sym}(I_N)$.

As in Section 9.3, we denote by $S(\Omega)$ the class of $\mathcal{B}$-$\mathcal{B}$-measurable maps $g : \Omega \to \Omega$ which are essentially injective and measure preserving, meaning that there exists a full $\mathbb{P}$-measure set $\Omega_0 \in \mathcal{B}$ such that g is injective on Ω_0 and $g_\sharp \mathbb{P} = \mathbb{P}$. Moreover, for every $N \in \mathfrak{N}$, we denote by $S_N(\Omega) := S(\Omega, \mathcal{B}, \mathbb{P}; \mathcal{B}_N)$, i.e. the subset of $S(\Omega)$ of $\mathcal{B}_N$-$\mathcal{B}_N$ measurable maps.

Remark 9.4.2. Clearly, if $X = \mathscr{I}_N(\boldsymbol{x}) \in \mathcal{H}_N$ and $g \in S_N(\Omega)$ then $X \circ g \in \mathcal{H}_N$ and there exists a unique permutation $\sigma = \sigma_g \in \mathrm{Sym}(I_N)$ such that $X \circ g = \sigma_g X = \mathscr{I}_N(\boldsymbol{x} \circ \sigma_g)$. Conversely, if $\sigma \in \mathrm{Sym}(I_N)$ there exists $g \in S_N(\Omega)$ such that $\sigma = \sigma_g$, as shown in Corollary 6.2.6.

We set

$$\begin{aligned} \mathcal{D}_\infty &:= \Big\{X \in \mathcal{H}_\infty : \iota_X \in C\Big\}, \\ \mathcal{D}_N &:= \mathcal{D}_\infty \cap \mathcal{H}_N = \Big\{X \in \mathcal{H}_N : \iota_X \in C \cap \mathcal{P}_N(\mathbb{H})\Big\}. \end{aligned} \tag{9.4.9}$$

Lemma 9.4.3. *Under Hypothesis 9.4.1, we have that for every $N \in \mathfrak{N}$ it holds:*

1. *$\mathcal{D}_N$ and $\mathcal{O}_N \cap \mathcal{D}_N$ are relatively open subsets of $\mathcal{H}_N$, invariant with respect to the action of permutations of* $\mathrm{Sym}(I_N)$*.*
2. *$\mathcal{O}_N \cap \mathcal{D}_N$ is dense in $\mathcal{D}_N$, and $\mathcal{D}_N = \mathrm{co}(\mathcal{O}_N \cap \mathcal{D}_N)$.*
3. *$\overline{\mathcal{O}_N \cap \mathcal{D}_N} = \overline{\mathcal{D}_N}$ is convex and its relative interior in $\mathcal{H}_N$ coincides with $\mathcal{D}_N = \mathrm{co}(\mathcal{O}_N \cap \mathcal{D}_N)$.*
4. *If $M \in \mathfrak{N}$ and $M \mid N$ then $\mathcal{D}_M = \mathcal{D}_N \cap \mathcal{H}_M = \Pi_M(\mathcal{D}_N)$, $\overline{\mathcal{D}_M} = \overline{\mathcal{D}_N} \cap \mathcal{H}_M$.*

Proof. (1) $\mathcal{D}_N$ is open in $\mathcal{H}_N$ since $C \cap \mathcal{P}_N(\mathbb{H})$ is relatively open in $\mathcal{P}_N(\mathbb{H})$ with respect to W_∞, and the map $X \mapsto \iota_X$ is Lipschitz from $\mathcal{H}_N$ to $\mathcal{P}_N(\mathbb{H})$, thanks to (9.4.8). $\mathcal{O}_N \cap \mathcal{D}_N$ is also open, since O_N is open in $\mathbb{H}_N$ and $\mathcal{O}_N = \mathscr{I}_N(O_N)$ is relatively open in $\mathcal{H}_N$.

(2) $\mathcal{O}_N \cap \mathcal{D}_N$ is dense in $\mathcal{D}_N$ since O_N is dense in $\mathbb{H}_N$ and $\mathcal{D}_N$ is relatively open.

On the other hand $\mathcal{D}_N$ is convex, since $C \cap \mathcal{P}_N(\mathbb{H})$ is convex along couplings of $\mathcal{P}_N(\mathbb{H} \times \mathbb{H})$ and for every $X, Y \in \mathcal{H}_N$ the coupling $\boldsymbol{\mu} = \iota^2_{X,Y}$ belongs to $\mathcal{P}_N(\mathbb{H} \times \mathbb{H})$. Clearly the displacement interpolation $\mu_t = x^t_\sharp \boldsymbol{\mu}$ corresponds to $((1-t)X + tY)_\sharp \mathbb{P}$ and belongs to $C \cap \mathcal{P}_N(\mathbb{H})$.

Let us now show that $\mathcal{D}_N$ coincides with $\mathrm{co}(\mathcal{O}_N \cap \mathcal{D}_N)$. If $\boldsymbol{z} : I_N \to \mathbb{H}$ is an arbitrary injective map with $|\boldsymbol{z}(n)| \leqslant 1$ for every $n \in I_N$, it is easy to see that for every $\boldsymbol{x} \in \mathbb{H}_N$ there exists a sufficiently small $\varepsilon > 0$ such that $\boldsymbol{x} + t\boldsymbol{z} \in O_N$ whenever $|t| < \varepsilon$. It follows that for every $X = \mathscr{I}_N(\boldsymbol{x}) \in \mathcal{D}_N$ we can find $t > 0$ such that $\mathscr{I}_N(\boldsymbol{x} + t\boldsymbol{z}), \mathscr{I}_N(\boldsymbol{x} - t\boldsymbol{z}) \in \mathcal{O}_N \cap \mathcal{D}_N$ and therefore $X \in \mathrm{co}(\mathcal{O}_N \cap \mathcal{D}_N)$.

(3) $\overline{\mathcal{O}_N \cap \mathcal{D}_N} = \overline{\mathcal{D}_N}$ is the closure of a convex set and therefore it is convex. By convexity, its relative interior coincides with $\mathcal{D}_N$.

(4) The identity $\mathcal{D}_M = \mathcal{D}_N \cap \mathcal{H}_M$ is trivial. Since $\mathcal{D}_M \subset \mathcal{D}_N$, in order to prove that $\mathcal{D}_M = \Pi_M \mathcal{D}_N$ it is sufficient to prove the inclusion $\Pi_M \mathcal{D}_N \subset \mathcal{D}_M$. Let $X \in \mathcal{D}_N$, $K := N/M$, and let $\sigma : I_N \to I_N$ be the cyclic permutation defined by

$$\sigma(n) := \begin{cases} mK + k + 1 & \text{if } n = mK + k,\ m \in I_M,\ 0 \leqslant k < K - 1; \\ mK & \text{if } n = mK + K - 1,\ m \in I_M \end{cases}$$

and let us consider its powers σ^p, $p \in I_K$. It is not difficult to check that $\sigma^K = \sigma^0 = \mathrm{Id}_{I_N}$ and for every $Y \in \mathcal{H}_M$ we have $\sigma^p Y = Y$ for every $p \in I_K$. Therefore for every $X \in \mathcal{H}_N$ we obtain the representation

$$\Pi_M(X) = \frac{1}{K} \sum_{p=0}^{K-1} \sigma^p X. \tag{9.4.10}$$

Since $\mathcal{D}_N$ is invariant with respect to permutations and convex, we get $\Pi_M(X) \in \mathcal{D}_N$ for every $X \in \mathcal{D}_N$, so that $\Pi_M(X) \in \mathcal{D}_N \cap \mathcal{H}_M = \mathcal{D}_M$.

Concerning the last property, it is obvious that $\overline{\mathcal{D}_M} \subset \overline{\mathcal{D}_N} \cap \mathcal{H}_M$. On the other hand, if $X \in \overline{\mathcal{D}_N}$ then $\Pi_M X \in \overline{\mathcal{D}_M}$ so that $\overline{\mathcal{D}_N} \cap \mathcal{H}_M \subset \overline{\mathcal{D}_M}$. □

Let us now study the Lagrangian representations of $\Phi \in \mathcal{P}_2(\mathsf{T}\mathbb{H})$. We can consider the (not empty) set of all the maps $(X, V) \in \mathcal{H}^2$ such that $(X, V)_\sharp \mathbb{P} = \Phi$. A particular case arises when the first marginal $\mu = x_\sharp \Phi$ of Φ belongs to $\mathcal{P}_N(\mathbb{H})$. In this case, X has the form $X = \mathscr{I}_N(\mathbf{x}) \in \mathcal{H}_N$, so that $\mu = X_\sharp \mathbb{P} = \frac{1}{N}\sum_{k \in I_N} \delta_{\mathbf{x}(k)}$, and we can construct V from the representation of Φ given by

$$\Phi = \frac{1}{N}\sum_{k \in I_N} \Phi_k, \quad x_\sharp \Phi_k = \delta_{\mathbf{x}(k)}, \tag{9.4.11}$$

for a family $(\Phi_k)_{k \in I_N} \subset \mathcal{P}(\mathsf{T}\mathbb{H})$, by setting $V(\omega) := V_k(\omega)$ if $\omega \in \Omega_{N,k}$, where $V_k \in L^2(\Omega_{N,k}, \mathbb{P}|_{\Omega_{N,k}}; \mathsf{H})$ are maps such that $(V_k)_\sharp \mathbb{P}|_{\Omega_{N,k}} = \frac{1}{N} v_\sharp \Phi_k$.

In the general case, when $\Phi \in \mathcal{P}_2(\mathsf{T}\mathbb{H})$, it is easy to check that if $(X, V)_\sharp \mathbb{P} = \Phi$ and $Y \in \mathcal{H}$ then

$$[\Phi, \iota^2_{X,Y}]_{r,0} \leqslant \langle V, X - Y \rangle_{\mathcal{H}}. \tag{9.4.12}$$

A particular important case arises when $X \in \mathcal{O}_N$ and $Y \in \mathcal{H}_N$: in this case Φ_k is uniquely determined by the disintegration of Φ w.r.t. μ, and $V|_{\Omega_{N,k}}$ coincides with V_k, with V_k as above, and

$$\langle V, X - Y \rangle_{\mathcal{H}} = \langle \Pi_N V, X - Y \rangle_{\mathcal{H}}, \quad \Pi_N V(\omega) = \mathbf{b}_\Phi(\mathbf{x}(k)) \quad \text{if } \omega \in \Omega_{N,k}. \tag{9.4.13}$$

It is easy to check that

$$[\Phi, \iota^2_{X,Y}]_{r,0} = \langle V, X - Y \rangle_{\mathcal{H}} = \langle \Pi_N V, X - Y \rangle_{\mathcal{H}} \quad \text{if } (X, V)_\sharp \mathbb{P} = \Phi,\ X \in \mathcal{O}_N,\ Y \in \mathcal{H}_N, \tag{9.4.14}$$

since $\iota^2_{X,Y}$ is concentrated on a map. We thus set

$$\mathbf{F}_N := \Big\{ (X, \Pi_N V) \in (\mathcal{O}_N \cap \mathcal{D}_N) \times \mathcal{H}_N : (X, V)_\sharp \mathbb{P} \in \mathbf{F} \Big\}. \tag{9.4.15}$$

We will also adopt the notation $V \in \mathbf{F}_N[X]$ if $(X, V) \in \mathbf{F}_N$. It is easy to check that for every $\sigma \in \mathrm{Sym}(N)$ and $(X, V) \in \mathbf{F}_N$ we have $(\sigma X, \sigma V) \in \mathbf{F}_N$. Indeed, if $(X, V) \in \mathbf{F}_N$ then there exists $W \in \mathcal{H}$ such that $(X, W)_\sharp \mathbb{P} \in \mathbf{F}$ and $V = \Pi_N W$. By Corollary 6.2.6, we can write $\sigma X = X \circ g \in \mathcal{O}_N \cap \mathcal{D}_N$ and $(X \circ g, W \circ g)_\sharp \mathbb{P} \in \mathbf{F}$. To conclude, it suffices to notice that $\Pi_N(W \circ g) = \sigma V$.

Moreover (9.4.14) and the dissipativity of $\mathbf{F}$ along couplings in $\mathcal{P}_N(\mathbb{H} \times \mathbb{H})$ given by Lemma 7.8.2 yields

$$(X, V),\ (Y, W) \in \mathbf{F}_N \quad \Rightarrow \quad \langle W - V, Y - X \rangle_{\mathcal{H}} \leqslant 0, \tag{9.4.16}$$

so that $\mathbf{F}_N$ is a dissipative set in $\mathcal{H}_N \times \mathcal{H}_N$ with open domain $D(\mathbf{F}_N) = \mathcal{O}_N \cap \mathcal{D}_N$.

Proposition 9.4.4. *Let us assume Hypothesis 9.4.1. For every* $N \in \mathfrak{N}$ *the dissipative set* $\mathbf{F}_N$ *admits a unique maximal extension* $\hat{\mathbf{F}}_N$ *in* $\mathcal{H}_N \times \mathcal{H}_N$ *with* $\mathcal{D}_N \subset D(\hat{\mathbf{F}}_N) \subset \overline{\mathcal{D}_N}$. $\hat{\mathbf{F}}_N$ *can be equivalently characterized by*

$$(X, V) \in \hat{\mathbf{F}}_N \quad \Leftrightarrow \quad X \in \overline{\mathcal{D}_N},\ V \in \mathcal{H}_N,\ \langle V - W, X - Y\rangle_{\mathcal{H}} \leqslant 0 \quad \forall (Y, W) \in \mathbf{F}_N, \tag{9.4.17}$$

and, whenever $X \in \mathcal{D}_N$, $\hat{\mathbf{F}}_N[X] = \overline{\mathrm{co}}\left(\bar{\mathbf{F}}_N[X]\right)$, *where*

$$\bar{\mathbf{F}}_N[X] := \Big\{ V \in \mathcal{H}_N : \exists\, (X_n, V_n) \in \mathbf{F}_N : X_n \to X,\ V_n \rightharpoonup V \Big\}. \tag{9.4.18}$$

$\hat{\mathbf{F}}_N$ *is invariant with respect to permutations*

$$(X, V) \in \hat{\mathbf{F}}_N,\ \sigma \in \mathrm{Sym}(I_N) \quad \Rightarrow \quad (\sigma X, \sigma V) \in \hat{\mathbf{F}}_N \tag{9.4.19}$$

and for every $X, Y \in \mathcal{D}_N$, *we have*

$$V \in \hat{\mathbf{F}}_N[X],\ \Psi \in \mathbf{F}[\iota_Y] \quad \Rightarrow \quad \langle V, X - Y\rangle + [\Psi, \iota^2_{Y,X}]_{r,0} \leqslant 0. \tag{9.4.20}$$

Finally, if $M \mid N = KM$, $X \in \overline{\mathcal{D}_M}$, *and* $(X, V) \in \hat{\mathbf{F}}_N$ *then* $\Pi_M V \in \hat{\mathbf{F}}_M[X]$. *Conversely, if* $X \in \mathcal{D}_M$ *and* $W \in \hat{\mathbf{F}}_M[X]$ *then there exists* $V \in \mathcal{H}_N$ *such that*

$$(X, V) \in \hat{\mathbf{F}}_N, \quad W = \Pi_M V. \tag{9.4.21}$$

Proof. (9.4.17) and (9.4.18) follow by the fact that $\mathcal{D}_N$ is convex and open and the domain of $\mathbf{F}_N$ is dense in $\mathcal{D}_N$, see Lemma 9.4.3 and Theorem 6.1.3.

Using (9.4.17) it is immediate to check that $\hat{\mathbf{F}}_N$ satisfies (9.4.19), since for every $(X, V) \in \hat{\mathbf{F}}_N$ and $(Y, W) \in \mathbf{F}_N$

$$\langle \sigma V - W, \sigma X - Y\rangle_{\mathcal{H}} = \langle V - \sigma^{-1} W, X - \sigma^{-1} Y\rangle_{\mathcal{H}} \leqslant 0,$$

since $\mathbf{F}_N$ and the scalar product in $\mathcal{H}_N$ are invariant by the action of permutations in $\mathrm{Sym}(I_N)$.

If $(X, V) \in \mathbf{F}_N$, (9.4.20) follows immediately since there exists $W \in \mathcal{H}$ such that $\Phi := (X, W)_\sharp \mathbb{P} \in \mathbf{F}$, $V = \Pi_N W$, and (9.4.14) yields $\langle V, X - Y\rangle_{\mathcal{H}} = [\Phi, \iota^2_{X,Y}]_{r,0}$ so that

$$\langle V, X - Y\rangle_{\mathcal{H}} + [\Psi, \iota^2_{Y,X}]_{r,0} = [\Phi, \iota^2_{X,Y}]_{r,0} + [\Psi, \iota^2_{Y,X}]_{r,0} \leqslant 0 \tag{9.4.22}$$

by (7.8.7).

If $X \in \mathcal{D}_N$ and $V \in \bar{\mathbf{F}}_N[X]$ according to (9.4.18), then there exist $(X_n, V_n) \in \mathbf{F}_N$, $X_n \in \mathcal{O}_N \cap \mathcal{D}_N$, such that $X_n \to X$ and $V_n \rightharpoonup V$. We can pass to the limit in (9.4.22) written for (X_n, V_n) and using Lemma 7.3.1 we obtain that (X, V) satisfies (9.4.22) as well. Finally, since (9.4.22) holds for every $V \in \bar{\mathbf{F}}_N[X]$, it also holds for every $V \in \overline{\mathrm{co}}\left(\bar{\mathbf{F}}_N[X]\right)$, hence (9.4.20).

Let us now suppose that $M \mid N$, $(X, V) \in \hat{\mathbf{F}}_N$ and $X \in \mathcal{D}_M$. We want to show that $W := \Pi_M V$ belongs to $\hat{\mathbf{F}}_M[X]$ by using (9.4.17). If $(Y, U) \in \mathbf{F}_M$ with $Y \in \mathcal{O}_M \cap \mathcal{D}_M$, we have $U = \Pi_M U'$ with $(Y, U')_\sharp \mathbb{P} =: \Phi \in \mathbf{F}$, so that (9.4.20) yields

$$\langle V, X - Y\rangle_{\mathcal{H}} + [\Phi, \iota^2_{Y,X}]_{r,0} \leqslant 0. \tag{9.4.23}$$

Since $Y \in \mathcal{O}_M$ and $X \in \mathcal{H}_M$ we have $[\Phi, \iota^2_{Y,X}]_{r,0} = \langle U, Y - X\rangle_{\mathcal{H}}$ by (9.4.14); since $X - Y \in \mathcal{H}_M$ we also have $\langle V, X - Y\rangle_{\mathcal{H}} = \langle \Pi_M V, X - Y\rangle_{\mathcal{H}}$ and we get

$$\langle W, X - Y\rangle_{\mathcal{H}} + \langle U, Y - X\rangle_{\mathcal{H}} = \langle V, X - Y\rangle_{\mathcal{H}} + [\Phi, \iota^2_{Y,X}]_{r,0} \leqslant 0. \tag{9.4.24}$$

Hence, by (9.4.17) $(X, W) \in \hat{\mathbf{F}}_M$. In particular, the above property shows that if $\mathbf{G} : \mathcal{D}_N \to \mathcal{H}_N$ is an arbitrary single valued selection of $\hat{\mathbf{F}}_N$, the restriction of $\Pi_M \circ \mathbf{G}$ to $\mathcal{D}_M$ is a selection of $\hat{\mathbf{F}}_M$. To conclude we need to prove that the property holds also if $X \in \overline{\mathcal{D}_M}$. Recall that by Lemma 9.4.3(3), $\overline{D(\mathbf{F}_M)} = \overline{\mathcal{D}_M}$. Then if $X \in \overline{\mathcal{D}_M}$, by Corollary 6.1.4 we have that W belongs to $\hat{\mathbf{F}}_M[X]$ if and only if

$$\langle W - (\Pi_M \circ \mathbf{G})|_{\mathcal{D}_M}[Y], X - Y\rangle_{\mathcal{H}} \leqslant 0 \quad \text{for every } Y \in \mathcal{D}_M, \tag{9.4.25}$$

i.e., if and only if

$$\langle W - \mathbf{G}[Y], X - Y\rangle_{\mathcal{H}} \leqslant 0 \quad \text{for every } Y \in \mathcal{D}_M. \tag{9.4.26}$$

If $V \in \hat{\mathbf{F}}_N[X]$, then using Corollary 6.1.4 we have

$$\langle V - \mathbf{G}[Y], X - Y\rangle_{\mathcal{H}} \leqslant 0 \quad \text{for every } Y \in \mathcal{D}_N \supset \mathcal{D}_M,$$

hence by (9.4.26) we get $\Pi_M V \in \hat{\mathbf{F}}_M[X]$.

Let us now show the converse implication. If $X \in \mathcal{D}_M$ and $W \in \hat{\mathbf{F}}_M[X]$, we need to prove that $W \in \Pi_M\left(\hat{\mathbf{F}}_N[X]\right)$. Since $\overline{D(\mathbf{G})} = \overline{\mathcal{D}_N}$, by Corollary 6.1.4 and Theorem 6.1.3 applied to $\mathbf{B} := \mathbf{G}$, we get $\Pi_M\left(\hat{\mathbf{F}}_N[X]\right) = \Pi_M\left(\tilde{\mathbf{G}}[X]\right) = \Pi_M\left(\overline{\text{co}}\left(\bar{\mathbf{G}}[X]\right)\right)$, where

$$\bar{\mathbf{G}}[X] := \left\{Z \in \mathcal{H}_N : \exists X_n \in \mathcal{D}_N : X_n \to X,\ \mathbf{G}(X_n) \rightharpoonup Z\right\}.$$

Similarly, denoting by $\mathcal{G} := (\Pi_M \circ \mathbf{G})|_{\mathcal{D}_M}$, by Corollary 6.1.4 and Theorem 6.1.3 we get

$$\begin{aligned}\hat{\mathbf{F}}_M[X] = \tilde{\mathcal{G}}[X] &= \overline{\text{co}}\left(\bar{\mathcal{G}}[X]\right)\\ &= \overline{\text{co}}\left(\{Z \in \mathcal{H}_M : \exists X_n \in \mathcal{D}_M : X_n \to X,\ \mathcal{G}(X_n) \rightharpoonup Z\}\right)\\ &= \Pi_M\left(\overline{\text{co}}\left(\bar{\mathbf{G}}[X]\right)\right),\end{aligned}$$

where the proof of the last equality can be pursued as follows. We first observe that

$$\begin{aligned}&\{Z \in \mathcal{H}_M : \exists X_n \in \mathcal{D}_M : X_n \to X,\ \mathcal{G}(X_n) \rightharpoonup Z\}\\ &= \Pi_M(\{W \in \mathcal{H}_N : \exists X_n \in \mathcal{D}_N : X_n \to X,\ \mathbf{G}(X_n) \rightharpoonup W\}) = \Pi_M(\bar{\mathbf{G}}[X]),\end{aligned}$$

by using the local boundedness of $\mathbf{G}$ as a selection of $\tilde{\mathbf{G}}$ (see [26, Prop. 2.9]) and the fact that Π_M is a linear and continuous operator. Then we notice that

$$\overline{\text{co}}\left(\Pi_M(\bar{\mathbf{G}}[X])\right) = \overline{\Pi_M(\text{co}(\bar{\mathbf{G}}[X]))} = \Pi_M(\overline{\text{co}}\left(\bar{\mathbf{G}}[X]\right)),$$

where the first equality follows by linearity of Π_M and, for the second, we exploit again the local boundedness of $\bar{\mathbf{G}}$ as a selection of $\tilde{\mathbf{G}}$ and the linearity and continuity of Π_M. Hence the conclusion. □

We can improve (9.4.20) with the following result.

Lemma 9.4.5. *Let us assume Hypothesis 9.4.1 and let* $N \in \mathfrak{N}$. *Then*

$$\langle V, X - Y\rangle + [\Psi, \iota^2_{Y,X}]_{r,0} \leqslant 0 \tag{9.4.27}$$

for every $(X, V) \in \hat{\mathbf{F}}_N$, $Y \in D(\hat{\mathbf{F}}_N)$ *and every* $\Psi \in \mathbf{F}[\iota_Y]$.

Proof. We start by proving (9.4.27) in case $X \in \mathcal{D}_N$. Let

$$Y_s := (1-s)Y + sX \in \mathcal{D}_N$$

for every $s \in (0, 1]$; then, by (9.4.20), we have

$$\langle V, X - Y_s\rangle + [\mathbf{F}, \iota^2_{Y_s,X}]_{r,0} \leqslant 0.$$

Using (7.4.5) we can rewrite the above equation as

$$\langle V, X - Y_s\rangle + (1-s)[\mathbf{F}, \iota^2_{Y,X}]_{r,s} \leqslant 0,$$

which, together with (7.8.7), gives

$$\langle V, X - Y_s\rangle + (1-s)[\mathbf{F}, \iota^2_{Y,X}]_{r,0} \leqslant 0.$$

Passing to the limit as $s \downarrow 0$, we obtain

$$\langle V, X - Y\rangle + [\Psi, \iota^2_{Y,X}]_{r,0} \leqslant 0 \quad \forall X \in \mathcal{D}_N,\ V \in \hat{\mathbf{F}}_N[X], Y \in D(\hat{\mathbf{F}}_N), \Psi \in \mathbf{F}[\iota_Y]. \tag{9.4.28}$$

We come now to the general case; let $(X, V) \in \hat{\mathbf{F}}_N$, $Y \in D(\hat{\mathbf{F}}_N)$ and $\Psi \in \mathbf{F}[\iota_Y]$. We define $Z = (X + Y)/2 \in \overline{\mathcal{D}_N}$ and, given $T \in \mathcal{D}_N$ and $V_T \in \hat{\mathbf{F}}_N[T]$, we set $Z_t := (1-t)Z + Tt \in \mathcal{D}_N$ for every $t \in (0, 1]$; we take, for every $t \in (0, 1]$, some $V_t \in \hat{\mathbf{F}}_N[Z_t]$. Clearly

$$(X, V), (Z_t, V_t), (T, V_T) \in \hat{\mathbf{F}}_N \quad \forall t \in (0, 1].$$

We compute

$$\begin{aligned}
&\langle V, X - Y\rangle + [\Psi, \iota^2_{Y,X}]_{r,0} = \\
&= \langle V - V_t, X - Y\rangle - \langle V_t, Y - X\rangle + [\Psi, \iota^2_{Y,X}]_{r,0} \\
&= 2\langle V - V_t, X - Z\rangle - 2\langle V_t, Y - Z\rangle + 2[\Psi, \iota^2_{Y,Z}]_{r,0} \\
&= 2\langle V - V_t, X - Z_t\rangle - 2\langle V_t, Y - Z_t\rangle + 2[\Psi, \iota^2_{Y,Z_t}]_{r,0} \\
&\quad + 2\langle V - V_t, Z_t - Z\rangle - 2\langle V_t, Z_t - Z\rangle - 2[\Psi, \iota^2_{Y,Z_t}]_{r,0} + 2[\Psi, \iota^2_{Y,Z}]_{r,0} \\
&= 2\langle V - V_t, X - Z_t\rangle - 2\langle V_t, Y - Z_t\rangle + 2[\Psi, \iota^2_{Y,Z_t}]_{r,0} \\
&\quad + 4\langle V_T - V_t, Z_t - Z\rangle + 2\langle V - 2V_T, Z_t - Z\rangle - 2[\Psi, \iota^2_{Y,Z_t}]_{r,0} + 2[\Psi, \iota^2_{Y,Z}]_{r,0} \\
&= 2\langle V - V_t, X - Z_t\rangle + 2\langle V_t, Z_t - Y\rangle + 2[\Psi, \iota^2_{Y,Z_t}]_{r,0} \\
&\quad + \frac{4t}{1-t}\langle V_T - V_t, T - Z_t\rangle + 2t\langle V - 2V_T, T - Z\rangle - 2[\Psi, \iota^2_{Y,Z_t}]_{r,0} \\
&\quad + 2[\Psi, \iota^2_{Y,Z}]_{r,0} \\
&\leqslant 2t\langle V - 2V_T, T - Z\rangle - 2[\Psi, \iota^2_{Y,Z_t}]_{r,0} + 2[\Psi, \iota^2_{Y,Z}]_{r,0},
\end{aligned}$$

where we have used again (7.4.5), the dissipativity of $\hat{\mathbf{F}}_N$ and (9.4.28) applied to $Z_t \in \mathcal{D}_N$, $V_t \in \hat{\mathbf{F}}_N[Z_t]$. Passing to the lim sup as $t \downarrow 0$, we get

$$\langle V, X - Y \rangle + [\Psi, \iota^2_{Y,X}]_{r,0} \leqslant 2[\Psi, \iota^2_{Y,Z}]_{r,0} - 2 \liminf_{t \downarrow 0} [\Psi, \iota^2_{Y,Z_t}]_{r,0} \leqslant 0$$

by Lemma 7.3.1. □

Proposition 9.4.6. *Let us assume Hypothesis 9.4.1. Then for every $X \in \mathcal{H}_\infty$ and every $\tau > 0$ there exists a unique $X_\tau \in \mathcal{H}_\infty$ such that*

$$X \in \mathcal{H}_N \Rightarrow X_\tau \in D(\hat{\mathbf{F}}_N) \subset \mathcal{H}_N \text{ and } X_\tau - X \in \tau \hat{\mathbf{F}}_N[X_\tau]. \tag{9.4.29}$$

Moreover

$$|X_\tau(\omega') - X_\tau(\omega'')| \leqslant |X(\omega') - X(\omega'')| \quad \text{for every } \omega', \omega'' \in \Omega. \tag{9.4.30}$$

Proof. Since $X \in \mathcal{H}_\infty$, there exits $N \in \mathfrak{N}$ such that $X \in \mathcal{H}_N$. Since $\hat{\mathbf{F}}_N$ is maximal, recalling (6.1.3), there exists a unique solution $X_{\tau,N} \in D(\hat{\mathbf{F}}_N)$ of

$$X_{\tau,N} - X \in \tau \hat{\mathbf{F}}_N[X_{\tau,N}].$$

The invariance of $\hat{\mathbf{F}}_N$ by permutations shows that $(\sigma X)_{\tau,N} = \sigma(X_{\tau,N})$ for every $\sigma \in \mathrm{Sym}(I_N)$. In particular, by dissipativity of $\hat{\mathbf{F}}_N$ we have

$$\langle \sigma X_{\tau,N} - \sigma X - (X_{\tau,N} - X), \sigma X_{\tau,N} - X_{\tau,N} \rangle_{\mathcal{H}} \leqslant 0$$

so that

$$|\sigma X_{\tau,N} - X_{\tau,N}|_{\mathcal{H}} \leqslant |\sigma X - X|_{\mathcal{H}} \quad \text{for every } \sigma \in \mathrm{Sym}(I_N).$$

If $\omega' \in \Omega_{N,i}$, $\omega'' \in \Omega_{N,j}$, $i, j \in I_N$, and we choose as σ the transposition which shifts i with j, we get

$$\frac{2}{N}|X_{\tau,N}(\omega') - X_{\tau,N}(\omega'')|^2 \leqslant \frac{2}{N}|X(\omega') - X(\omega'')|^2$$

which yields (9.4.30).

Let us now suppose that $X \in \mathcal{H}_M$ with $M \mid N$. $X_{\tau,N}$ belongs to $\mathcal{H}_M$ by (9.4.30), so that $X_{\tau,N} \in \overline{\mathcal{D}_N} \cap \mathcal{H}_M = \overline{\mathcal{D}_M}$ by Lemma 9.4.3(4). By Proposition 9.4.4, for every $Y \in \mathcal{D}_M$ and $W \in \hat{\mathbf{F}}_M[Y]$ we can find $V \in \hat{\mathbf{F}}_N[Y]$ such that $W = \Pi_M V$, so that by dissipativity of $\hat{\mathbf{F}}_N$ we have

$$\langle X_{\tau,N} - X - \tau V, X_{\tau,N} - Y \rangle_{\mathcal{H}} \leqslant 0. \tag{9.4.31}$$

Since $X_{\tau,N} - Y \in \mathcal{H}_M$, we can replace V with $W = \Pi_M V$ in (9.4.31), thus obtaining that $X_{\tau,N} - X \in \tau \hat{\mathbf{F}}_M[X_{\tau,N}]$ by Corollary 6.1.4, i.e. $X_{\tau,N} = X_{\tau,M}$. If M, N are arbitrary and $X \in \mathcal{H}_M \cap \mathcal{H}_N$, then setting $R := MN$ the previous argument shows that $X_{\tau,M} = X_{\tau,R} = X_{\tau,N}$. □

Corollary 9.4.7. *Let us assume Hypothesis 9.4.1, let $M \in \mathfrak{N}$ and let $X \in D(\hat{\mathbf{F}}_M)$. Then*

1. $X \in D(\hat{\mathbf{F}}_N)$ *for every* $N \in \mathfrak{N}$ *s.t.* $M \mid N$.

2. $\mathbf{F}^\circ[X] := \lim_{\tau\downarrow 0} \frac{X_\tau - X}{\tau} \in \hat{\mathbf{F}}_M[X]$. *In particular* $\mathbf{F}^\circ[X] \in \hat{\mathbf{F}}_N[X]$ *for every* $N \in \mathfrak{N}$ *s.t.* $M \mid N$.

3. $|\mathbf{F}^\circ[X]|_{\mathcal{H}} \leqslant |V|_{\mathcal{H}}$ *for every* $V \in \hat{\mathbf{F}}_N[X]$ *and for every* $N \in \mathfrak{N}$ *s.t.* $M \mid N$.

4. $|X_\tau - X|_{\mathcal{H}} \leqslant \tau |\mathbf{F}^\circ[X]|_{\mathcal{H}}$ *for every* $\tau > 0$.

Moreover, for every $X, Y \in \bigcup_{N\in\mathfrak{N}} D(\hat{\mathbf{F}}_N)$, *we have*

$$\langle \mathbf{F}^\circ[X] - \mathbf{F}^\circ[Y], X - Y\rangle_{\mathcal{H}} \leqslant 0. \tag{9.4.32}$$

Proof. By e.g. [26, Prop. 2.6(iii)] there exists the limit

$$\lim_{\tau\downarrow 0} \frac{X_\tau - X}{\tau} = \mathbf{F}^\circ[X] \in \hat{\mathbf{F}}_M[X]$$

and (4) holds. If $N \in \mathfrak{N}$ is s.t. $M \mid N$, then $X \in D(\hat{\mathbf{F}}_M) \subset \overline{\mathcal{D}_M} \subset \overline{\mathcal{D}_N}$, by Lemma 9.4.3. Moreover by Proposition 9.4.6, we have that

$$\frac{X_\tau - X}{\tau} \in \hat{\mathbf{F}}_N[X_\tau] \quad \forall \tau > 0.$$

In particular

$$\langle \frac{X_\tau - X}{\tau} - W, X_\tau - Y\rangle \leqslant 0 \quad \forall (Y, W) \in \mathbf{F}_N \quad \forall \tau > 0,$$

so that, passing to the limit as $\tau \downarrow 0$, we get

$$\langle \mathbf{F}^\circ[X] - W, X - Y\rangle \leqslant 0 \quad \forall (Y, W) \in \mathbf{F}_N,$$

since $X_\tau \to X$ as $\tau \downarrow 0$ by [26, Theorem 2.2]. This proves that $(X, \mathbf{F}^\circ[X]) \in \hat{\mathbf{F}}_N$ and, in particular, that $X \in D(\hat{\mathbf{F}}_N)$. This proves (1) and (2), while (3) immediately follows, also using [26, Prop. 2.6(iii)].
Finally, if $X, Y \in \bigcup_{N\in\mathfrak{N}} D(\hat{\mathbf{F}}_N)$, then there exist $N, M \in \mathfrak{N}$ s.t. $X \in D(\hat{\mathbf{F}}_N)$ and $Y \in D(\hat{\mathbf{F}}_M)$ so that, taking $R := MN$, we have

$$(X, \mathbf{F}^\circ[X]), (Y, \mathbf{F}^\circ[Y]) \in \hat{\mathbf{F}}_R$$

by (3) and the dissipativity of $\hat{\mathbf{F}}_R$ gives (9.4.32). □

We can therefore define the operator $\mathbf{F}_\infty \subset \mathcal{H} \times \mathcal{H}$

$$\mathbf{F}_\infty := \Big\{(X, V) \in \mathcal{H}_\infty \times \mathcal{H}_\infty : \exists M \in \mathfrak{N} : (X, V) \in \hat{\mathbf{F}}_N \ \forall N \in \mathfrak{N},\ M \mid N\Big\}. \tag{9.4.33}$$

By the previous results, $\mathbf{F}_\infty$ is well defined and dissipative with domain $D(\mathbf{F}_\infty) = \bigcup_{N\in\mathfrak{N}} D(\hat{\mathbf{F}}_N)$, $\mathbf{F}^\circ$ provides the minimal selection and, by Proposition 9.4.6, for every $X \in \mathcal{H}_\infty$ there exists a unique $X_\tau \in D(\mathbf{F}_\infty)$ such that $X_\tau - X \in \tau\mathbf{F}_\infty[X_\tau]$. We can then apply Lemma 6.1.5 and find the unique maximal extension $\mathbf{F}$ with domain $D(\mathbf{F}) \subset \overline{D(\mathbf{F}_\infty)}$ and characterized by

$$(X, V) \in \mathbf{F} \quad \Leftrightarrow \quad X \in \overline{D(\mathbf{F}_\infty)}, \quad \langle V - W, X - Y\rangle \leqslant 0 \quad \text{for every } (Y, W) \in \mathbf{F}_\infty. \tag{9.4.34}$$

Moreover, since $\mathcal{H}_\infty$ is dense in $\mathcal{H}$, we have by Lemma 6.1.5 that

$$\mathbf{F} = \overline{\mathbf{F}_\infty}^{\mathcal{H}\times\mathcal{H}}. \tag{9.4.35}$$

Remark 9.4.8. Notice that the notation $\mathbf{F}^\circ$ used in Corollary 9.4.7 and in the above discussion is coherent with the one used in Section 6.1 since, as highlighted in the proof of Lemma 6.1.5, the minimal selection of $\mathbf{F}$, when restricted to $D(\mathbf{F}_\infty)$, coincides with $\mathbf{F}^\circ$ defined as in Corollary 9.4.7. Moreover, if $X \in \mathcal{H}_\infty$, the resolvent operator $\mathbf{J}_\tau$ of $\mathbf{F}$ applied to X coincides with X_τ as in Proposition 9.4.6.

Proposition 9.4.9. *Let us assume Hypothesis 9.4.1, let $(X, V) \in \mathbf{F}$ and let $g \in S(\Omega)$. Then $(X \circ g, V \circ g) \in \mathbf{F}$; in particular, we have that $\mathbf{F}$ is an invariant maximal dissipative operator. Moreover, if $Y \in D(\mathbf{F}_\infty)$ and $\Psi \in \mathbf{F}[\iota_Y]$, we have*

$$\langle V, X - Y\rangle + [\Psi, \iota^2_{Y,X}]_{r,0} \leqslant 0. \tag{9.4.36}$$

Finally, if $X \in \mathcal{D}_M$ for some $M \in \mathfrak{N}$ and $\Phi \in \mathbf{F}[\iota_X]$, then

$$|\mathbf{F}^\circ[X]|^2_{\mathcal{H}} \leqslant \int_{\mathbb{H}} |\mathbf{b}_\Phi|^2 \, d\iota_X. \tag{9.4.37}$$

Proof. To show that $(X \circ g, V \circ g) \in \mathbf{F}$, by (9.4.35), it is enough to prove that there exist $((Z_n, W_n))_n \subset \mathbf{F}_\infty$ s.t. $Z_n \to X \circ g$ and $W_n \to V \circ g$. By (9.4.35), we can find a sequence $((X_n, V_n))_n \subset \mathbf{F}_\infty$ and an increasing sequence $(N_n)_n \subset \mathbb{N}$ s.t. $X_n \to X$, $V_n \to V$ and $(X_n, V_n) \in \hat{\mathbf{F}}_N$ for every $N \in \mathfrak{N}$ such that $N_n \prec N$. Let $(b_n)_n \subset \mathfrak{N}$ be the sequence given by Lemma 6.2.13; by Theorem 6.2.16 applied to $(\Omega, \mathcal{B}, \mathbb{P}, (\mathfrak{P}_{b_n})_{n\in\mathbb{N}})$ and $\boldsymbol{\gamma} := (\mathbf{i}_\Omega, g)_\sharp\mathbb{P}$, we can find a strictly increasing sequence $(M_j)_j \subset \mathbb{N}$ and maps $g_j \in S_{b_{M_j}}(\Omega)$ such that

$$(U, W)_\sharp(\mathbf{i}_\Omega, g_j)_\sharp\mathbb{P} \to (U, W)_\sharp(\mathbf{i}_\Omega, g)_\sharp\mathbb{P} \text{ in } \mathcal{P}_2(\mathbb{H}^2)$$

for every $U, V \in \mathcal{H}$. Since M_j is increasing and (6.2.5) holds, then for every $n \in \mathbb{N}$ there exists $j = j(n) \in \mathbb{N}$ such that $g_{j(n)} \in S_{N_n}(\Omega)$. Thus setting $g'_n := g_{j(n)}$, $n \in \mathbb{N}$, by Remark 9.4.2 and (9.4.19) we get that $(X_n \circ g'_n, V_n \circ g'_n) \in \hat{\mathbf{F}}_N$ $\forall N \in \mathfrak{N}$ s.t. $N_n \prec N$, for every $n \in \mathbb{N}$. In particular, $(X_n \circ g'_n, V_n \circ g'_n) \in \mathbf{F}_\infty$ for any $n \in \mathbb{N}$ and

$$(U, W)_\sharp(\mathbf{i}_\Omega, g'_n)_\sharp\mathbb{P} \to (U, W)_\sharp(\mathbf{i}_\Omega, g)_\sharp\mathbb{P} \text{ in } \mathcal{P}_2(\mathbb{H}^2) \tag{9.4.38}$$

for every $U, V \in \mathcal{H}$. We are left to show that $X_n \circ g'_n \to X \circ g$ in $\mathcal{H}$ (the case of V is completely analogous). Since $|X \circ g'_n - X_n \circ g'_n|_{\mathcal{H}} = |X - X_n|_{\mathcal{H}}$ it is enough to show that $X \circ g'_n \to X \circ g$ which, on the other hand, is implied by $X \circ g'_n \rightharpoonup X \circ g$, since $|X \circ g'_n|_{\mathcal{H}} = |X|_{\mathcal{H}} = |X \circ g|_{\mathcal{H}}$. Let $Y \in \mathcal{H}$ and let us take $U = Y, V = X$ in (9.4.38) so that

$$\begin{aligned}
\langle X \circ g'_n, Y\rangle_{\mathcal{H}} &= \int_{\mathbb{H}^2} \langle x, y\rangle \, d((Y, X) \circ (\mathbf{i}_\Omega, g'_n))_\sharp\mathbb{P} \\
&\to \int_{\mathbb{H}^2} \langle x, y\rangle \, d((Y, X) \circ (\mathbf{i}_\Omega, g))_\sharp\mathbb{P} \\
&= \langle X \circ g, Y\rangle_{\mathcal{H}},
\end{aligned}$$

since $\varphi(x, y) := \langle x, y\rangle$ is a real valued function on $\mathbb{H}^2$ with less than quadratic growth (see e.g.[5, Proposition 7.1.5, Lemma 5.1.7]). This shows that $X \circ g'_n \rightharpoonup X \circ g$. We conclude that $(X \circ g, V \circ g) \in \mathbf{F}$.

We now prove (9.4.36). Let $((X_n, V_n))_n$ as before. If $Y \in D(\mathbf{F}_\infty)$ and $\Psi \in \mathbf{F}[\iota_Y]$, for every $n \in \mathbb{N}$ we can find $M_n \in \mathfrak{N}$ such that

$$(X_n, V_n) \in \hat{\mathbf{F}}_N, Y \in D(\hat{\mathbf{F}}_N) \quad \forall N \in \mathfrak{N}, M_n \prec N.$$

By Lemma 9.4.5, we have

$$\langle V_n, X_n - Y\rangle + [\Psi, \iota^2_{Y,X_n}]_{r,0} \leqslant 0 \quad \forall n \in \mathbb{N}.$$

Passing to the liminf as $n \to +\infty$ and using Lemma 7.3.1 we obtain (9.4.36).
Let now $X \in \mathcal{D}_M$ for some $M \in \mathfrak{N}$, and observe that, since $\mathcal{D}_M$ is open by Proposition 9.4.3, $\mathbf{J}_\tau X \in \mathcal{D}_M$ for $\tau > 0$ sufficiently small, since $\mathbf{J}_\tau X \to X$ as $\tau \downarrow 0$. We can thus apply (9.4.20) and get

$$\frac{1}{\tau}\langle \mathbf{J}_\tau X - X, \mathbf{J}_\tau X - X\rangle_{\mathcal{H}} + [\Phi, \iota^2_{X,\mathbf{J}_\tau X}]_{r,0} \leqslant 0.$$

Since we have shown that $\mathbf{F}$ is an invariant maximal dissipative operator, by Theorem 9.3.3, there exists a Lipschitz function $f_{X,\tau}$ such that $\mathbf{J}_\tau X = X \circ f_{X,\tau}$; thus $\iota^2_{X,\mathbf{J}_\tau X}$ is concentrated on a map so that, by Remark 7.4.2, we have

$$[\Phi, \iota^2_{X,\mathbf{J}_\tau X}]_{r,0} = \langle \boldsymbol{b}_\Phi, X - \mathbf{J}_\tau X\rangle_{\mathcal{H}}.$$

We hence get

$$\frac{1}{\tau^2}|\mathbf{J}_\tau X - X|^2_{\mathcal{H}} \leqslant |\boldsymbol{b}_\Phi|^2_{\mathcal{H}} = \int_{\mathbb{H}} |\boldsymbol{b}_\Phi|^2 \, d\iota_X$$

and passing to the limit as $\tau \downarrow 0$ we obtain (9.4.37).

□

Thanks to Proposition 9.4.9 and Theorem 9.3.3, for every $X \in D(\mathbf{F})$, the law of $(X, \mathbf{F}^\circ[X])$ only depends on the law of X so that we can give the following definition.

Definition 9.4.10. For every $\mu \in \iota(D(\mathbf{F}))$, we define $\mathbf{F}^\circ[\mu]$ as

$$\mathbf{F}^\circ[\mu] := (X, \mathbf{F}^\circ[X])_\sharp \mathbb{P},$$

where $X \in D(\mathbf{F})$ is such that $\iota_X = \mu$.

9.4.1 *Lagrangian EVI solutions*

In this section we show that the curve $t \mapsto S_t X_0$ solving (9.3.3) for the maximal dissipative operator $\mathbf{F}$ constructed starting from $\mathbf{F}$ induces a 0-EVI solution for $\mathbf{F}$. We recall that by Theorem 9.3.3 the curve $t \mapsto S_t X_0$ only depends on the law of X_0.

In addition to Hypothesis 9.4.1, we adopt the following additional compatibility property for the core C.

Hypothesis 9.4.11. We assume that

$$\text{for every } \mu \in D(\mathbf{F}) \text{ there exists } \mu_n \in C \text{ and } \Phi_n \in \mathbf{F}[\mu_n] \text{ such that}$$
$$W_2(\mu_n, \mu) \to 0, \quad \sup_n |\Phi_n|_2 < +\infty. \tag{9.4.39}$$

Lemma 9.4.12. *Under Hypotheses 9.4.1 and 9.4.11, it holds*

$$D(\mathbf{F}) \subset \iota(D(\mathbf{F})) \subset \overline{D(\mathbf{F})}.$$

Proof. Let $\mu \in D(\mathbf{F})$. By (9.4.39), we can find a sequence $(\mu_n)_n \subset C \subset D(\mathbf{F}) \cap \mathcal{P}_{\mathfrak{N}}(\mathbb{H})$ s.t. $\mu_n \to \mu$ and a sequence $\Phi_n \in \mathbf{F}[\mu_n]$ with $\sup_n |\Phi_n|_2 < +\infty$. We can assume without loss of generality that

$$\sum_{n=1}^{+\infty} W_2(\mu_n, \mu_{n+1}) < +\infty$$

and find a sequence $(X_n)_n \subset \mathcal{H}$ s.t. $(X_n, X_{n+1})_\sharp \mathbb{P} \in \Gamma_o(\mu_n, \mu_{n+1})$ for every $n \in \mathbb{N}$ (see [5, Lemma 5.3.4] and Corollary 6.2.3). Then $(X_n)_n$ is a Cauchy sequence in $\mathcal{H}$ and thus there exists some $X \in \mathcal{H}$ s.t. $X_n \to X$ and $X_\sharp \mathbb{P} = \mu$. Moreover, since $\mu_n \in C \cap \mathcal{P}_{2^{N_n}}(\mathbb{H})$ for some $N_n \in \mathbb{N}$, we can find $Z_n \in \mathcal{D}_{N_n}$ s.t. $\iota_{Z_n} = \mu_n$ for every $n \in \mathbb{N}$ and, by (9.4.37), we also have

$$\sup_n |\mathbf{F}^\circ[Z_n]|_{\mathcal{H}} < +\infty.$$

Having X_n and Z_n the same law, Proposition 9.4.9 and Theorem 9.3.3 give that $(X_n, \mathbf{F}^\circ[X_n]) \in \mathbf{F}$ with

$$\sup_n |\mathbf{F}^\circ[X_n]|_{\mathcal{H}} < +\infty.$$

We can thus find an (unrelabeled) subsequence s.t. $\mathbf{F}^\circ[X_n] \rightharpoonup V \in \mathcal{H}$. By (9.4.34) we get that $(X, V) \in \mathbf{F}$ and, in particular, that $X \in D(\mathbf{F})$. This proves that $D(\mathbf{F}) \subset \iota(D(\mathbf{F}))$. Let us come to the other inclusion: if $X \in D(\mathbf{F})$, we can find a sequence $(X_n)_n \subset D(\mathbf{F}_\infty)$ s.t. $X_n \to X$. This means that $X_n \in \overline{\mathcal{D}_{N_n}}$ for some $N_n \in \mathbb{N}$ and we can thus find some $X_n' \in \mathcal{D}_{N_n}$ s.t. $|X_n - X_n'|_{\mathcal{H}} < 1/n$ for every $n \in \mathbb{N}$. It is clear that $\iota_{X_n'} \in D(\mathbf{F})$ and that $\iota_{X_n'} \to \iota_X$ in $\mathcal{P}_2(\mathbb{H})$. □

Corollary 9.4.13. *Let us assume Hypotheses 9.4.1 and 9.4.11 and let*

$$\mathbf{G} := \left\{ \Psi \in \mathcal{P}_2(\mathsf{T}\mathbb{H}) : \exists \Psi_n \in \mathbf{F} : \Psi_n \to \Psi \textit{ in } \mathcal{P}_2^{sw}(\mathsf{T}\mathbb{H}),\, x_\sharp \Psi_n \in C \right\} \subset \mathrm{cl}(\mathbf{F}), \tag{9.4.40}$$

where $\mathrm{cl}(\mathbf{F})$ *is as in Proposition 7.7.1. Then for every* $(X, V) \in \mathbf{F}$, $\Psi \in \mathbf{G}$ *and* $Y \in \mathcal{H}$ *such that* $Y_\sharp \mathbb{P} = x_\sharp \Psi$, *we have*

$$\langle V, X - Y \rangle + [\Psi, \iota^2_{Y,X}]_{r,0} \leqslant 0. \tag{9.4.41}$$

Proof. Let X, V, Ψ, Y be as in the statement and set $\nu := x_\sharp \Psi = Y_\sharp \mathbb{P}$. Then, there exists $(\Psi_n)_{n \in \mathbb{N}} \subset \mathbf{F}$ such that $\Psi_n \to \Psi$ in $\mathcal{P}_2^{sw}(\mathsf{T}\mathbb{H})$ and $\nu_n := x_\sharp \Psi_n \in C$. By Proposition 6.2.18, there exists $(Y_n)_{n \in \mathbb{N}} \subset \mathcal{H}$ such that $(Y_n)_\sharp \mathbb{P} = \nu_n$ and $Y_n \to Y$ in $\mathcal{H}$. Moreover, since $\nu_n \in C \cap \mathcal{P}_{N_n}(\mathbb{H})$ for some $N_n \in \mathbb{N}$, we can find $\tilde{Y}_n \in \mathcal{D}_{N_n}$ s.t. $\iota_{\tilde{Y}_n} = \nu_n$ for every $n \in \mathbb{N}$. Thus, having Y_n and $\tilde{Y}_n$ the same discrete law, there exists some $g_n \in S(\Omega)$ s.t. $\tilde{Y}_n = Y_n \circ g_n$. In particular we get by (9.4.19) that $Y_n \circ g_n \in D(\mathbf{F}_\infty)$ and, by Proposition 9.4.9, $(X \circ g_n, V \circ g_n) \in \mathbf{F}$ for all $n \in \mathbb{N}$. By (9.4.36), we get

$$\langle V \circ g_n, X \circ g_n - Y_n \circ g_n \rangle + [\Psi_n, \iota^2_{Y_n \circ g_n, X \circ g_n}]_{r,0} \leqslant 0 \qquad \forall n \in \mathbb{N}.$$

Being g_n measure preserving maps, the above relation is equivalent to

$$\langle V, X - Y_n \rangle + [\Psi_n, \iota^2_{Y_n, X}]_{r,0} \leqslant 0 \qquad \forall n \in \mathbb{N}.$$

Taking the lim inf as $n \to +\infty$ and using Lemma 7.3.1, we conclude. □

Corollary 9.4.14. *Let us assume Hypotheses 9.4.1 and 9.4.11 and let* $\mathbf{G}$ *be as in (9.4.40). Then for every* $\nu \in \iota(D(\mathbf{F}))$, $\Psi \in \mathbf{G}$ *and* $\boldsymbol{\gamma} \in \Gamma(\nu, x_\sharp \Psi)$ *it holds*

$$[\mathbf{F}^\circ[\nu], \boldsymbol{\gamma}]_{r,0} \leqslant [\Psi, \boldsymbol{\gamma}]_{l,1}, \tag{9.4.42}$$

where $\mathbf{F}^\circ[\nu]$ *is as in Definition 9.4.10.*

Proof. Let $X, Y \in \mathcal{H}$ be such that $(X, Y)_\sharp \mathbb{P} = \boldsymbol{\gamma}$. By Proposition 9.4.9 and Theorem 9.3.3 we have that $X \in D(\mathbf{F})$. By Corollary 9.4.13, we have that

$$[\mathbf{F}^\circ[\nu], \boldsymbol{\gamma}]_{r,0} = [(X, \mathbf{F}^\circ[X])_\sharp \mathbb{P}, \boldsymbol{\gamma}]_{r,0} \leqslant \langle \mathbf{F}^\circ[X], X - Y \rangle \leqslant [\Psi, \iota^2_{X,Y}]_{l,1} = [\Psi, \boldsymbol{\gamma}]_{l,1}.$$

□

We recall that, given $\mu, \nu \in \mathcal{P}_2(\mathbb{H})$, a generalized geodesic ([5, Definition 9.2.2]) connecting μ to ν is a curve $(\mu_t)_{t \in [0,1]} \subset \mathcal{P}_2(\mathbb{H})$ such that there exists $\boldsymbol{\gamma} \in \mathcal{P}(\mathbb{H}^3)$ with $\pi^{1,2}_\sharp \boldsymbol{\gamma}$ and $\pi^{2,3}_\sharp \boldsymbol{\gamma}$ optimal and

$$\mu_0 = \mu, \quad \mu_1 = \nu, \quad \mu_t = x^t_\sharp(\pi^{1,3}_\sharp \boldsymbol{\gamma}), \quad t \in [0,1].$$

We denote by $G(\mu, \nu)$ the set of generalized geodesic connecting μ to ν.

Definition 9.4.15. For every $\nu \in \overline{D(\mathbf{F})}$ we define the set of *measures that see* ν, denoted by $\mathcal{S}_{\mathbf{F}}(\nu)$, as

$$\mathcal{S}_{\mathbf{F}}(\nu) := \left\{ \mu \in D_f(\mathbf{F}) \;\middle|\; \begin{array}{l} \mu_t \in D(\mathbf{F}) \text{ for every } t \in (0,1] \\ \text{and every } (\mu_t)_{t \in [0,1]} \in G(\nu, \mu) \end{array} \right\},$$

Theorem 9.4.16. *Let us assume Hypotheses 9.4.1 and 9.4.11, let* $\mu \in \overline{D(\mathbf{F})}$ *and let* $\Phi \in \mathcal{P}_2(\mathsf{T}\mathbb{H}|\mu)$ *be such that*

$$\left[\Phi, x_\sharp \Psi\right]_r + [\Psi, \mu]_r \leqslant 0 \quad \forall \Psi \in \mathbf{G}, \tag{9.4.43}$$

where $\mathbf{G}$ *is as in (9.4.40). If for every* $\mu' \in \overline{D(\mathbf{F})}$ *the set* $\mathcal{S}_{\mathbf{F}}(\mu')$ *is non-empty and open in* $\mathcal{P}_f(\mathbb{H})$ *with respect to the* W_∞*-topology, then for any* $\nu \in \iota(D(\mathbf{F}))$ *we have*

$$[\mathbf{F}^\circ[\nu], \Phi]_r \leqslant 0. \tag{9.4.44}$$

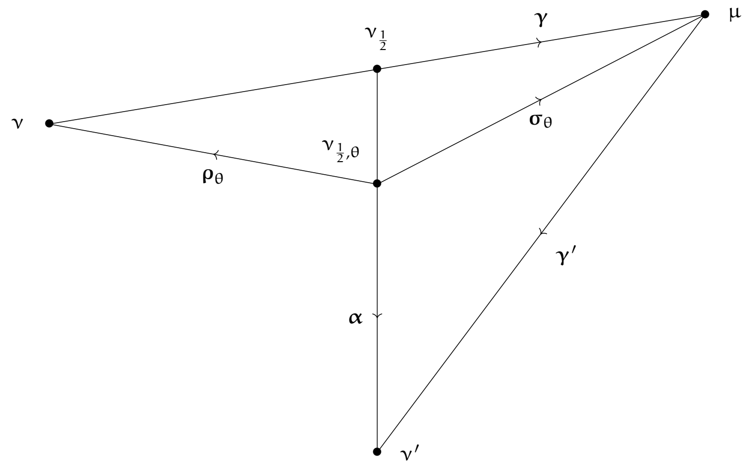

Figure 3: The plans used in the proof of Theorem 9.4.16.

Proof. We start by proving the theorem in case $\nu \in \iota(D(\mathbf{F}_\infty))$. By the geodesic convexity of $\overline{D(\mathbf{F})}$ (given by (9.4.39) and the convexity of C along any discrete plan coming from Hypothesis 9.4.1(d)), we can find $\boldsymbol{\gamma} \in \Gamma_o(\nu, \mu)$ such that $x^t_\sharp \boldsymbol{\gamma} \in \overline{D(\mathbf{F})}$ for every $t \in [0,1]$; in particular $\nu_{1/2} := x^{1/2}_\sharp \boldsymbol{\gamma} \in \overline{D(\mathbf{F})}$. Since $\mathcal{S}_{\mathbf{F}}\left(\nu_{1/2}\right)$ is open in $\mathcal{P}_f(\mathbb{H})$ with respect to the W_∞-topology and $\operatorname{supp}(\nu)$ is a finite set, we can use Proposition 6.4.3 to find $\nu' \in \mathcal{S}_{\mathbf{F}}\left(\nu_{1/2}\right)$ such that

$$(\operatorname{supp}\nu' - \operatorname{supp}\nu') \cap \operatorname{dir}(\operatorname{supp}\nu) = \{0\}. \tag{9.4.45}$$

Let $\boldsymbol{\gamma}' \in \Gamma_o(\mu, \nu')$ and let $\boldsymbol{\lambda} \in \mathcal{P}(\mathbb{H}^3)$ be such that $\pi^{1,2}_\sharp \boldsymbol{\lambda} = \boldsymbol{\gamma}$ and $\pi^{2,3}_\sharp \boldsymbol{\lambda} = \boldsymbol{\gamma}'$. By Proposition 6.4.2, we get that $x^{1/2,\theta}$ is injective on $\operatorname{supp}\boldsymbol{\lambda}$ for every $\theta \in (0,1)$. For every $\theta \in (0,1)$ we define

$$\begin{aligned}
\nu_{1/2,\theta} &:= x^{1/2,\theta}_\sharp \boldsymbol{\lambda}, \\
\boldsymbol{\alpha} &:= (x^{1/2} \circ \pi^{1,2}, \pi^3)_\sharp \boldsymbol{\lambda}, \\
\boldsymbol{\sigma}_\theta &:= (x^{1/2,\theta}, \pi^2)_\sharp \boldsymbol{\lambda}, \\
\boldsymbol{\rho}_\theta &:= (x^{1/2,\theta}, \pi^1)_\sharp \boldsymbol{\lambda}
\end{aligned}$$

and we take $\Psi_\theta \in \mathbf{G}[\nu_{1/2,\theta}]$ and $\Sigma_\theta \in \mathcal{P}(\mathsf{T}\mathbb{H} \times \mathbb{H})$ with $\pi^{1,2}_\sharp \Sigma_\theta = \Psi_\theta$, $\pi^{1,3}_\sharp \Sigma_\theta = \boldsymbol{\rho}_\theta$ such that

$$\int_{\mathsf{T}\mathbb{H}\times\mathbb{H}} \langle x_2, x_1 - x_3 \rangle \, d\Sigma_\theta(x_1, x_2, x_3) = [\Psi_\theta, \boldsymbol{\rho}_\theta]_{r,0}.$$

Observe that, by Proposition 6.4.1, $\boldsymbol{\sigma}_\theta$ (resp. $s_\sharp \boldsymbol{\sigma}_\theta$) is the unique optimal transport plan from $\nu_{1/2,\theta}$ to μ (resp. from μ to $\nu_{1/2,\theta}$), so that by (9.4.43) we get

$$[\Phi, s_\sharp \boldsymbol{\sigma}_\theta]_{r,0} + [\Psi_\theta, \boldsymbol{\sigma}_\theta]_{r,0} = \left[\Phi, \nu_{1/2,\theta}\right]_r + [\Psi_\theta, \mu]_r \leqslant 0. \tag{9.4.46}$$

By construction, we can find $\boldsymbol{\lambda}_\theta \in \mathcal{P}(\mathbb{H}^5)$ such that

$$\pi^{1,2,3}_\sharp \boldsymbol{\lambda}_\theta = \Sigma_\theta, \quad \pi^{1,3,4,5}_\sharp \boldsymbol{\lambda}_\theta = (x^{1/2,\theta}, \boldsymbol{i}_{\mathbb{H}^3})_\sharp \boldsymbol{\lambda}.$$

Let us define $h_\theta : \mathbb{H}^3 \to \mathbb{H}$ as

$$h_\theta(x_1, x_2, x_3) := \frac{2}{1-\theta}(x_1 - \theta x_3) - x_2, \quad (x_1, x_2, x_3) \in \mathbb{H}^3$$

and $\Xi_\theta \in \mathcal{P}(\mathbb{H}^4)$ as

$$\Xi_\theta := (\pi^{2,3}, h_\theta \circ \pi^{1,3,5}, \pi^5)_\sharp \lambda_\theta.$$

It can be easily checked that

$$(x^{1/2,\theta} \circ \pi^{2,3,4}, \pi^{1,2})_\sharp \Xi_\theta = \Sigma_\theta, \tag{9.4.47}$$

$$(x^{1/2} \circ \pi^{2,3}, \pi^1, x^{1/2,\theta} \circ \pi^{2,3,4})_\sharp \Xi_\theta \in \Gamma_1(\Psi_\theta, (x^0, x^\theta)_\sharp \boldsymbol{\alpha}), \tag{9.4.48}$$

$$(x^{1/2,\theta} \circ \pi^{2,3,4}, \pi^{1,3})_\sharp \Xi_\theta \in \Gamma_0(\Psi_\theta, \boldsymbol{\sigma}_\theta). \tag{9.4.49}$$

By (9.4.42), we have

$$[\mathbf{F}^\circ[\nu], s_\sharp \boldsymbol{\rho}_\theta]_{r,0} \leqslant -[\Psi_\theta, \boldsymbol{\rho}_\theta]_{r,0}. \tag{9.4.50}$$

Using (9.4.47) we can thus compute

$$\begin{aligned}
[\mathbf{F}^\circ[\nu], s_\sharp \boldsymbol{\rho}_\theta]_{r,0} &\leqslant -[\Psi_\theta, \boldsymbol{\rho}_\theta]_{r,0} \\
&= \int_{\mathsf{T}\mathbb{H}\times\mathbb{H}} \langle x_2, x_3 - x_1 \rangle \, d\Sigma_\theta \\
&= \int_{\mathbb{H}^4} \langle x_1, x_2 - x^{1/2,\theta}(x_2, x_3, x_4) \rangle \, d\Xi_\theta \\
&= \int_{\mathbb{H}^4} \langle x_1, x_2 - \frac{x_2 + x_3}{2} \rangle \, d\Xi_\theta \\
&\quad + \int_{\mathbb{H}^4} \langle x_1, \frac{x_2 + x_3}{2} - x^{1/2,\theta}(x_2, x_3, x_4) \rangle \, d\Xi_\theta \\
&= \int_{\mathbb{H}^4} \langle x_1, \frac{x_2 + x_3}{2} - x_3 \rangle \, d\Xi_\theta \\
&\quad + \int_{\mathbb{H}^4} \langle x_1, \frac{x_2 + x_3}{2} - x^{1/2,\theta}(x_2, x_3, x_4) \rangle \, d\Xi_\theta \\
&= 2\int_{\mathbb{H}^4} \langle x_1, \frac{x_2 + x_3}{2} - x^{1/2,\theta}(x_2, x_3, x_4) \rangle \, d\Xi_\theta \\
&\quad + \int_{\mathbb{H}^4} \langle x_1, x^{1/2,\theta}(x_2, x_3, x_4) - x_3 \rangle \, d\Xi_\theta \\
&= 2\int_{\mathsf{T}\mathbb{H}\times\mathbb{H}} \langle x_2, x_1 - x_3 \rangle \, d(x^{1/2} \circ \pi^{2,3}, \pi^1, x^{1/2,\theta} \circ \pi^{2,3,4})_\sharp \Xi_\theta \\
&\quad + \int_{\mathsf{T}\mathbb{H}\times\mathbb{H}} \langle x_2, x_1 - x_3 \rangle \, d(x^{1/2,\theta} \circ \pi^{2,3,4}, \pi^{1,3})_\sharp \Xi_\theta \\
&\leqslant 2[\Psi_\theta, (x^0, x^\theta)_\sharp \boldsymbol{\alpha}]_{l,1} + [\Psi_\theta, \boldsymbol{\sigma}_\theta]_{r,0},
\end{aligned}$$

where we have used (9.4.48), (9.4.49) and the fact that $\Gamma_0(\Psi_\theta, \boldsymbol{\sigma}_\theta)$ contains a unique element, being $\boldsymbol{\sigma}_\theta$ induced by a map w.r.t. $\nu_{1/2,\theta}$, as previously remarked. By (9.4.46) we get

$$[\mathbf{F}^\circ[\nu], s_\sharp \boldsymbol{\rho}_\theta]_{r,0} \leqslant -[\Phi, s_\sharp \boldsymbol{\sigma}_\theta]_{r,0} + 2[\Psi_\theta, (x^0, x^\theta)_\sharp \boldsymbol{\alpha}]_{l,1} = -[\Phi, s_\sharp \boldsymbol{\sigma}_\theta]_{r,0} + 2\theta[\Psi_\theta, \boldsymbol{\alpha}]_{r,\theta},$$

$$(9.4.51)$$

thanks to (7.4.5), Remark 7.4.2 and the α-essential injectivity of x^t for every $t \in (0,1)$ given by Proposition 6.4.2. From (9.4.42) with $\gamma = (x^1, x^\theta)_\sharp \alpha$, $\Psi = \Psi_\theta$ and $\nu = \nu'$, we get

$$[\mathbf{F}^\circ[\nu'], (x^1, x^\theta)_\sharp \alpha]_{r,0} \leqslant [\Psi_\theta, (x^1, x^\theta)_\sharp \alpha]_{l,1}$$

which gives, using (7.4.3) and (7.4.5), that

$$[\Psi_\theta, \alpha]_{r,\theta} \leqslant [\mathbf{F}^\circ[\nu'], \alpha]_{l,1}. \tag{9.4.52}$$

Using (9.4.52) in (9.4.51) we finally get

$$[\mathbf{F}^\circ[\nu], s_\sharp \rho_\theta]_{r,0} + [\Phi, s_\sharp \sigma_\theta]_{r,0} \leqslant 2\theta[\mathbf{F}^\circ[\nu'], \alpha]_{l,1}.$$

Passing to the lim inf as $\theta \downarrow 0$ and using Lemma 7.3.1 we get that

$$[\mathbf{F}^\circ[\nu], (x^0, x^{1/2})_\sharp \gamma]_{r,0} + [\Phi, (x^1, x^{1/2})_\sharp \gamma]_{r,0} \leqslant 0.$$

Thus

$$\frac{1}{2}[\mathbf{F}^\circ[\nu], \Phi]_r \leqslant [\mathbf{F}^\circ[\nu], (x^0, x^{1/2})_\sharp \gamma]_{r,0} + [\Phi, (x^1, x^{1/2})_\sharp \gamma]_{r,0} \leqslant 0.$$

This proves the Theorem in case $\nu \in \iota(D(\mathrm{F}_\infty))$. Let us prove the general case; let $X \in D(\mathbf{F})$ be such that $\iota_X = \nu$. By (9.4.35), we can find $(X_n, V_n)_n \subset \mathrm{F}_\infty$ such that $X_n \to X$ and $V_n \to \mathrm{F}^\circ[X]$. Since $(V_n)_n$ is bounded, also $(\mathrm{F}^\circ[X_n])_n$ is bounded and thus weakly converges, up to an unrelabeled subsequence, to $V' \in \mathrm{F}[X]$, where we used (9.4.34) to conclude that $(X, V') \in \mathrm{F}$. By the weak lower semicontinuity of the norm, we have that

$$|V'|_{\mathcal{H}} \leqslant \liminf_n |\mathrm{F}^\circ[X_n]|_{\mathcal{H}} \leqslant \limsup_n |\mathrm{F}^\circ[X_n]|_{\mathcal{H}} \leqslant \limsup_n |V_n|_{\mathcal{H}} = |\mathrm{F}^\circ[X]|_{\mathcal{H}},$$

and then $V' = \mathrm{F}^\circ[X]$ and $\mathrm{F}^\circ[X_n] \to \mathrm{F}^\circ[X]$. Writing (9.4.44) for ι_{X_n} and observing that $\mathbf{F}^\circ[\iota_{X_n}] \to \mathbf{F}^\circ[\nu]$ in $\mathcal{P}_2(\mathsf{TH})$, we conclude using Lemma 7.3.1. □

Remark 9.4.17. Notice that if

$$\text{for every } \mu' \in \overline{D(\mathbf{F})} \text{ there exists } x_0 \in \mathbb{H} \text{ s.t. } \delta_{x_0} \in \mathcal{S}_{\mathbf{F}}\left(\mu'\right),$$

then, in order to prove Theorem 9.4.16, there is no need to assume that $\mathcal{S}_{\mathbf{F}}(\mu')$ is open in $\mathcal{P}_f(\mathbb{H})$ with respect to the W_∞-topology. Indeed, condition (9.4.45) is automatically satisfied for $\nu' = \delta_{x_0}$ so that the use of Proposition 6.4.3 is not necessary.

In the following Theorem, given $\mu_0 \in D(\mathbf{F})$, we denote

$$S_t\mu_0 := (S_t X_0)_\sharp \mathbb{P} \quad \text{for } t \geqslant 0, \tag{9.4.53}$$

where $X_0 \in \mathcal{H}$ is such that $\iota_{X_0} = \mu_0$. Notice that this definition is well posed since, by Theorem 9.3.3, $S_t\mu_0$ only depends on μ_0.

Theorem 9.4.18. *Let us assume Hypotheses 9.4.1 and 9.4.11 and let $\mu_0, \nu_0 \in D(\mathbf{F})$. Assume that for every $\mu' \in \overline{D(\mathbf{F})}$ the set $\mathcal{S}_{\mathbf{F}}(\mu')$ is non-empty and open in $\mathcal{P}_f(\mathbb{H})$ with respect to the W_∞-topology. Let $(\mu_t)_{t\geqslant 0}$ be the Lipschitz curve defined by $\mu_t := S_t\mu_0$ for every $t \in [0, +\infty)$ and let $(\nu_t)_{t\geqslant 0}$ be a locally absolutely continuous 0-EVI solution for $\mathbf{F}$ starting from ν_0. Then*

$$W_2(\mu_t, \nu_t) \leqslant W_2(\mu_0, \nu_0) \quad \forall t \in [0, +\infty), \tag{9.4.54}$$

In particular $(\mu_t)_{t\geqslant 0}$ is the unique locally absolutely continuous 0-EVI solution for $\mathbf{F}$ starting from μ_0.

Proof. Let $X_0 \in \mathcal{H}$ be such that $\iota_{X_0} = \mu_0$. By Proposition 9.4.9, Theorem 9.3.3 and Lemma 9.4.12, we have that $X_0 \in D(\mathbf{F})$ and $(\mu_s)_{s\geqslant 0}$ is independent of X_0. By Lemma 9.4.12, since $\iota(D(\mathbf{F})) \subset \overline{D(\mathbf{F})}$ and $S_sX_0 \in D(\mathbf{F})$, we have that $(\mu_s)_{s\geqslant 0} \in \mathrm{Lip}([0, +\infty); \overline{D(\mathbf{F})})$. Let us define for every $s \geqslant 0$, $\boldsymbol{v}_s : \mathbb{H} \to \mathbb{H}$ as $\boldsymbol{v}_s(x) := h_{S_sX_0}(x)$ (see Theorem 9.3.3(b)); by (9.3.4) we get

$$\partial_s\mu_s + \nabla \cdot (\mu_s\boldsymbol{v}_s) = 0 \qquad \text{for every } s \geqslant 0$$

in the sense of distributions.

(1) $(\mu_s)_{s\geqslant 0}$ *is a 0-EVI solution for* $\mathbf{F}$. Let $\Phi \in \mathbf{F}$ and $t \in A((\mu_s)_{s\geqslant 0}, x_\sharp\Phi)$ be fixed, where $A((\mu_s)_{s\geqslant 0}, x_\sharp\Phi)$ is the subset of $A((\mu_s)_{s\geqslant 0})$ coming from in Theorem 7.2.1. Observe that, since $\mathbf{G} \subset \mathrm{cl}(\mathbf{F})$ (cf. (9.4.40)) and the latter is dissipative by Proposition 7.7.1, we have that

$$\left[\Phi, x_\sharp\Psi\right]_r + \left[\Psi, x_\sharp\Phi\right]_r \leqslant 0 \quad \forall \Psi \in \mathbf{G}. \tag{9.4.55}$$

By Theorem 7.2.1,

$$\begin{aligned}
\frac{1}{2}\frac{d}{ds}\bigg|_{s=t} W_2^2(\mu_s, x_\sharp\Phi) &= \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp\mu_t, x_\sharp\Phi\right]_r \\
&= \left[(S_tX_0, \mathbf{F}^\circ[S_tX_0])_\sharp\mathbb{P}, x_\sharp\Phi\right]_r \\
&= \left[\mathbf{F}^\circ[\mu_t], x_\sharp\Phi\right]_r \\
&\leqslant -\left[\Phi, \mu_t\right]_r,
\end{aligned}$$

where, thanks to (9.4.55), we applied Theorem 9.4.16 to get the last inequality. This proves that $(\mu_s)_{s\geqslant 0}$ is a 0-EVI solution (cf. Definition 8.1.1) for $\mathbf{F}$ starting from $\mu_0 \in D(\mathbf{F})$.

(2) $(\mu_s)_{s\geqslant 0}$ *is the unique absolutely continuous 0-EVI solution for* $\mathbf{F}$ *starting from* μ_0. Let $(\nu_s)_{s\geqslant 0}$ be a locally absolutely continuous 0-EVI solution for $\mathbf{F}$ starting from $\nu_0 \in D(\mathbf{F})$ and let $\boldsymbol{w}$ be its Wasserstein velocity field coming from Theorem 2.4.6. Let $t \in A((\mu_s)_{s\geqslant 0}, (\nu_s)_{s\geqslant 0})$ be fixed, where $A((\mu_s)_{s\geqslant 0}, (\nu_s)_{s\geqslant 0})$ is the subset of $A((\mu_s)_{s\geqslant 0}) \cap A((\nu_s)_{s\geqslant 0})$ coming from in Theorem 7.2.3. By definition of 0-EVI solution for $\mathbf{F}$ we have that

$$\frac{1}{2}\frac{d}{ds}\bigg|_{s=t} W_2^2(\nu_s, x_\sharp\Psi') = \left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{w}_t)_\sharp\nu_t, x_\sharp\Psi'\right]_r \leqslant -\left[\Psi', \nu_t\right]_r \quad \forall \Psi' \in \mathbf{F}. \tag{9.4.56}$$

If $\Psi \in \mathbf{G}$ (cf. (9.4.40)), we can take a sequence $(\Psi'_n)_n \subset \mathbf{F}$ converging to Ψ in $\mathcal{P}_2^{sw}(\mathsf{T}\mathbb{H})$ obtaining by Lemma 7.3.1 that (9.4.56) holds also for Ψ so that

$$\left[(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{w}_t)_\sharp\nu_t, x_\sharp\Psi\right]_r \leqslant -\left[\Psi, \nu_t\right]_r \quad \forall \Psi \in \mathbf{G}. \tag{9.4.57}$$

We thus have by Theorem 7.2.3 that

$$\begin{aligned}\frac{1}{2}\frac{d}{ds}\Big|_{s=t} W_2^2(\mu_s,\nu_s) &= \left[(\boldsymbol{i}_{\mathbb{H}},\boldsymbol{v}_t)_\sharp\mu_t,(\boldsymbol{i}_{\mathbb{H}},\boldsymbol{w}_t)_\sharp\nu_t\right]_r\\ &= \left[(S_tX_0,\mathbf{F}^\circ[S_tX_0])_\sharp\mathbb{P},(\boldsymbol{i}_{\mathbb{H}},\boldsymbol{w}_t)_\sharp\nu_t\right]_r\\ &= \left[\mathbf{F}^\circ[\mu_t],(\boldsymbol{i}_{\mathbb{H}},\boldsymbol{w}_t)_\sharp\nu_t\right]_r\\ &\leqslant 0,\end{aligned}$$

where we were allowed to use (9.4.44) with $\Phi = (\boldsymbol{i}_{\mathbb{H}},\boldsymbol{w}_t)_\sharp\nu_t$ thanks to (9.4.57). Thus

$$W_2(\mu_t,\nu_t) \leqslant W_2(\mu_0,\nu_0) \quad \forall t\in[0,+\infty)$$

so that we obtain the sought uniqueness. □

9.4.2 *JKO scheme and Hilbertian resolvent*

The aim of this section is to show that, under suitable assumptions, the JKO/Minimizing movement scheme for $\phi:\mathcal{P}_2(\mathbb{H})\to(-\infty,+\infty]$ and the (law of the) Implicit Euler Hilbertian scheme constructed starting from the MPVF $-\boldsymbol{\partial}\phi$ coincide. We will work under the following assumptions.

Hypothesis 9.4.19. We assume that:

(a) $\mathfrak{N}$ is a fixed unbounded directed subset of the integers w.r.t. the order relation $\prec$ as in (6.2.4).

(b) $(\Omega,\mathcal{B},\mathbb{P},(\mathfrak{P}_N)_{N\in\mathfrak{N}})$, with $\mathfrak{P}_N=\{\Omega_{N,n}\}_{n\in I_N}$ and $I_N:=\{0,\dots,N-1\}$, is a fixed $\mathfrak{N}$-refined standard Borel probability space as in Definition 6.2.9.

(c) $\phi:\mathcal{P}_2(\mathbb{H})\to(-\infty,+\infty]$ is a proper, lower semicontinuous and geodesically convex functional.

(d) $C\subset D(\phi)\cap\mathcal{P}_{\mathfrak{N}}(\mathbb{H})$ is a fixed nonempty set such that $C\cap\mathcal{P}_N(\mathbb{H})$ is W_∞-relatively open in $\mathcal{P}_N(\mathbb{H})$ and convex along couplings in $\mathcal{P}_N(\mathbb{H}\times\mathbb{H})$ for every $N\in\mathfrak{N}$, where $\mathcal{P}_N(\mathbb{H})$ and $\mathcal{P}_{\mathfrak{N}}(\mathbb{H})$ are as in (9.4.1) and (9.4.2) respectively.

(e) C is *dense in energy*, meaning that for every $\mu\in D(\phi)$ there exists $(\mu_n)_n\subset C$ such that

$$\mu_n\to\mu \quad\text{and}\quad \phi(\mu_n)\to\phi(\mu).$$

We consider the MPVF $\mathbf{F}:=-\boldsymbol{\partial}\phi\subset\mathcal{P}_2(\mathsf{T}\mathbb{H})$, the (opposite of the) Wasserstein subdifferential of ϕ defined as

$$\Psi\in-\boldsymbol{\partial}\phi(\mu) \quad\text{iff}\quad \mu\in D(\phi),\quad \phi(\nu)-\phi(\mu)\geqslant[\Psi,\nu]_r \quad \text{for every } \nu\in D(\phi). \tag{9.4.58}$$

It is easy to check that $\mathbf{F}$ is a dissipative MPVF. We will use the notations $\mathcal{H}, \mathcal{H}_N, \iota_\cdot, \mathcal{D}_N, \mathbf{F}_N, \hat{\mathbf{F}}_N, \mathbf{F}_\infty, \mathbf{F}$ precisely as in the first part of Section 9.4 with $\mathbf{F} = -\boldsymbol{\partial}\phi$. Finally we denote by $\psi : \mathcal{H} \to (-\infty, +\infty]$ the functional on $\mathcal{H}$ defined as

$$\psi(X) := \phi(\iota_X) \quad X \in \mathcal{H}.$$

Proposition 9.4.20. *The functional ψ is proper, convex and lower semicontinuous. In particular, ϕ is convex along any coupling and the opposite of the subdifferential of ψ, $-\partial\psi \subset \mathcal{H} \times \mathcal{H}$, is a maximal dissipative operator.*

Proof. We only check the convexity of ψ since the other properties are trivially inherited by the functional ϕ. We proceed in two steps.
(1) ϕ *is convex in* C, *meaning that for every* $\mu_0, \mu_1 \in C$, *every* $\boldsymbol{\mu} \in \Gamma(\mu_0, \mu_1)$ *and every* $t \in [0, 1]$ *it holds*

$$\phi(x^t_\sharp \boldsymbol{\mu}) \leqslant (1 - t)\phi(\mu_0) + t\phi(\mu_1).$$

The thesis follows if we prove that $f(t) := \phi(\mu_t)$, $t \in [0, 1]$ is convex, where $\mu_t := x^t_\sharp \boldsymbol{\mu}$. By Theorem 6.5.2 there exists a natural number $K \in \mathbb{N}$ and times $0 = t_0 < t_1 < \cdots < t_{K-1} < t_K = 1$ such that $x^{t_{j-1},t_j}_\sharp \boldsymbol{\mu}$ is optimal for every $j = 1, \dots, K$. Setting $s_{2j} := t_j$, $j = 0, \dots, K$ and $s_{2j+1} := (t_j + t_{j+1})/2$, $j = 0, \dots, K-1$, we get a partition $0 = s_0 < s_1 < \cdots < s_{N-1} < s_N = 1$ with $N = 2K$ of $[0, 1]$ such that $x^{s_{j-1},s_j}_\sharp \boldsymbol{\mu}$, $j = 1, \dots, N$, is the unique optimal transport plan connecting its marginals. Hence, f is convex in each interval $[s_{j-1}, s_j]$ for $j = 1, \dots, N$ by definition of geodesic convexity. In particular, f is continuous in $[0, 1]$, the right (resp. left) derivative of f, denoted by f'_r (resp. f'_l), exists at every point of $[0, 1)$ (resp. $(0, 1]$) and it is increasing in each subinterval $[s_{j-1}, s_j)$ (resp. $(s_{j-1}, s_j]$); moreover f'_r and f'_l coincide in a dense set $I \subset \cup_{j=1}^K (s_{j-1}, s_j)$. Let now $t \in [0, 1)$ and let $j \in 1, \dots, N$ be such that $t \in [s_{j-1}, s_j)$; let $h > 0$ be such that $t + h < s_j$. By definition of subdifferential at μ_t and using the fact that $x^{t,t+h}_\sharp \boldsymbol{\mu}$ is the unique element of $\Gamma_o(\mu_t, \mu_{t+h})$, we get that

$$f(t + h) - f(t) \geqslant [\Phi_t, \mu_{t+h}]_r = [\Phi_t, x^{t,t+h}_\sharp \boldsymbol{\mu}]_{r,0} = h[\Phi_t, \boldsymbol{\mu}]_{r,t}$$

for every $\Phi_t \in -\boldsymbol{\partial}\phi(\mu_t)$. Dividing by h and passing to the limit as $h \downarrow 0$ shows that $f'_r(t) \geqslant [-\boldsymbol{\partial}\phi, \boldsymbol{\mu}]_{r,t}$. The same argument shows that $f'_l(t) \leqslant [-\boldsymbol{\partial}\phi, \boldsymbol{\mu}]_{l,t}$ for every $t \in (0, 1]$. This, together with Theorem 7.6.2 also implies that, if $u \in I$, then

$$f'_r(u) = f'_l(u) = [-\boldsymbol{\partial}\phi, \boldsymbol{\mu}]_{r,u} = [-\boldsymbol{\partial}\phi, \boldsymbol{\mu}]_{l,u}.$$

To conclude that f is convex it is enough to show that $f'_l(s_j) \leqslant f'_r(s_j)$ for every $j = 1, \dots, N-1$. Let us fix $j \in \{1, \dots, N-1\}$ and let us consider any $s \in (s_{j-1}, s_j)$ and $t \in (s_j, s_{j+1})$. Then we can find two points $u_1 \in (s, s_j) \cap I$ and $u_2 \in (s_j, t) \cap I$ so that

$$f'_l(s) \leqslant f'_l(u_1) = [-\boldsymbol{\partial}\phi, \boldsymbol{\mu}]_{r,u_1} \leqslant [-\boldsymbol{\partial}\phi, \boldsymbol{\mu}]_{l,u_2} = f'_r(u_2) \leqslant f'_r(t),$$

where we have used (7.8.7) for the second inequality. Passing to the limit as $s \uparrow s_j$ and $t \downarrow s_j$ and using the right (resp. left) continuity of the right (resp. left)

derivative (see e.g. [102, Theorem 24.1]), we conclude. This proves the convexity of f and concludes the first step.
(2) ψ *is convex.*

Let $X, Y \in D(\psi)$ and let $t \in [0,1]$; then $\iota_X, \iota_Y \in D(\phi)$ and we can thus find by Hypothesis 9.4.19(e) sequences $(\mu_n)_n, (\nu_n) \subset C$ such that $W_2(\mu_n, \iota_X) \to 0$, $W_2(\nu_n, \iota_Y) \to 0$, $\phi(\mu_n) \to \phi(\iota_X)$ and $\phi(\nu_n) \to \phi(\iota_Y)$ as $n \to +\infty$. By Proposition 6.2.18 we can find sequences $(X_n)_n, (Y_n)_n \subset \mathcal{H}$ such that $\iota_{X_n} = \mu_n$, $\iota_{Y_n} = \nu_n$, $X_n \to X$ and $Y_n \to Y$. In particular by step (1) with $\boldsymbol{\mu} = \iota^2_{X_n,Y_n}$ we have that

$$\begin{aligned}\psi((1-t)X_n + tY_n) &= \phi(x^t_\sharp \iota^2_{X_n,Y_n}) \\ &\leqslant (1-t)\phi(\mu_n) + t\phi(\nu_n) \\ &= (1-t)\psi(X_n) + t\psi(Y_n).\end{aligned}$$

Passing to the $\liminf_n$ and using the lower semicontinuity of ψ yield the sought convexity. □

Proposition 9.4.21. *Under Hypothesis 9.4.19, it holds*

$$F = -\partial\psi.$$

Proof. Since both operators are maximal dissipative it is enough to prove that $F \subset -\partial\psi$. We do it in several steps.
(1) *If* $N \in \mathfrak{N}$, $(X, V) \in F_N$, $Y \in \mathcal{D}_N$, *then* $\psi(Y) - \psi(X) \geqslant \langle V, X - Y\rangle_{\mathcal{H}}$.

By the very definition of F_N, there exists $W \in \mathcal{H}$ such that $V = \Pi_N W$ and $(X, W)_\sharp \mathbb{P} \in -\boldsymbol{\partial}\phi$ i.e.

$$\psi(Y) - \psi(X) = \phi(\iota_Y) - \phi(\iota_X) \geqslant \left[(X, W)_\sharp \mathbb{P}, \iota_Y\right]_r = \langle V, X - Y\rangle_{\mathcal{H}},$$

where we used (9.4.14) for the last equality.
(2) *If* $N \in \mathfrak{N}$, $X, Y \in \mathcal{D}_N$, $V \in \hat{F}_N[X]$, *then* $\psi(Y) - \psi(X) \geqslant \langle V, X - Y\rangle_{\mathcal{H}}$.

This follows by step (1) arguing as in the proof of (9.4.20), first approximating $V \in \bar{F}_N[X]$ with elements of F_N and then by using the equality $\hat{F}_N[X] = \overline{\mathrm{co}}\left(\bar{F}_N[X]\right)$.
(3) *If* $N \in \mathfrak{N}$, $(X, V) \in \hat{F}_N$, $Y \in D(\hat{F}_N)$, *then* $\psi(Y) - \psi(X) \geqslant \langle V, X - Y\rangle_{\mathcal{H}}$.

This follows by step (2) arguing as in the proof of Lemma 9.4.5 replacing $[\Psi, \iota^2_{Y,X}]_{r,0}$ with $\psi(X) - \psi(Y)$.
(4) *If* $(X, V) \in F_\infty$, $Y \in D(\psi)$, *then* $\psi(Y) - \psi(X) \geqslant \langle V, X - Y\rangle_{\mathcal{H}}$.

If $(X, V) \in F_\infty$ there exists $M \in \mathfrak{N}$ such that $(X, V) \in F_N$ for every $N \in \mathfrak{N}$ such that $M \prec N$. By Hypotesis 9.4.19(e) we can find a sequence $(\nu_n) \subset C$ such that $W_2(\nu_n, \iota_Y) \to 0$ and $\phi(\nu_n) \to \phi(\iota_Y)$ as $n \to +\infty$. By Proposition 6.2.18, we can find a sequence $(Y_n)_n \subset \mathcal{H}$ such that $\iota_{Y_n} = \nu_n$ and $Y_n \to Y$. We can thus find a sequence $(N_n) \subset \mathfrak{N}$ such that $(X, V) \in F_{N_n}$ and $Y_n \in \mathcal{D}_{N_n} \subset D(\hat{F}_{N_n})$ for every $n \in \mathbb{N}$. By step (3) written for (X, V) and Y_n and passing to the limit as $n \to +\infty$, we conclude.
(5) *If* $(X, V) \in F$, *then* $(X, V) \in -\partial\psi$.

By (9.4.35), we can find a sequence $(X_n, V_n) \subset F_\infty$ such that $X_n \to X$ and $V_n \to V$. Let $Y \in D(\psi)$; by step (4) written for (X_n, V_n) and Y and passing to the limit as $n \to +\infty$, we conclude. □

Proposition 9.4.22. *Under Hypothesis 9.4.19 the the Implicit Euler scheme provides a step of JKO scheme. In particular, let $\mu \in \mathcal{P}_2(\mathbb{H})$ and let $\tau > 0$. If $X \in \mathcal{H}$ is such that $\iota_X = \mu$, then the law μ_τ of $X_\tau := \mathbf{J}_\tau X$ satisfies*

$$\mu_\tau \in \underset{\nu \in \mathcal{P}_2(X)}{\arg\min} \left\{ \frac{1}{2\tau} W_2^2(\mu, \nu) + \phi(\nu) \right\}, \tag{9.4.59}$$

i.e. μ_τ is a step of JKO for ϕ starting from μ.

Proof. By Proposition 9.4.9 and Theorem 9.3.3, we have that μ_τ doesn't depend on the choice of $X \in \mathcal{H}$ such that $\iota_X = \mu$; if $\nu \in \mathcal{P}_2(\mathbb{H})$, we can thus find $(X', Y) \in \mathcal{H}^2$ such that $\iota^2_{X',Y} \in \Gamma_o(\mu, \nu)$. By the properties of the resolvent operator $\mathbf{J}_\tau$, we have that

$$\begin{aligned} \phi(\mu_\tau) + \frac{1}{2\tau} W_2^2(\mu_\tau, \mu) &\leqslant \psi(\mathbf{J}_\tau X') + \frac{1}{2\tau} |\mathbf{J}_\tau X' - X'|^2_{\mathcal{H}} \\ &\leqslant \psi(Y) + \frac{1}{2\tau} |Y - X'|^2_{\mathcal{H}} \\ &= \phi(\nu) + \frac{1}{2\tau} W_2^2(\mu, \nu), \end{aligned}$$

which yields the conclusion. □

In this last lemma we remark that the hypotheses we assumed on ϕ in this section are implied by the one assumed on $-\boldsymbol{\partial}\phi$ in the previous section.

Lemma 9.4.23. *Hypothesis 9.4.19(e) for ϕ is weaker than Hypothesis 9.4.11 for $-\boldsymbol{\partial}\phi$.*

Proof. In particular we show that

$$\begin{gathered} \text{there exists } C \subset D(\boldsymbol{\partial}\phi) \text{ s.t. for every } \mu \in D(\boldsymbol{\partial}\phi) \\ \text{there exists } (\mu_n) \subset C,\ \Phi_n \in -\boldsymbol{\partial}\phi(\mu_n) \text{ s.t. } \mu_n \to \mu \text{ and } \sup_n |\Phi_n|_2 < +\infty \end{gathered} \tag{9.4.60}$$

implies

$$\begin{gathered} \text{there exists } D \subset D(\phi) \text{ s.t. for every } \mu \in D(\phi) \text{ there exists } (\mu_n)_n \subset D \text{ s.t.} \\ \mu_n \to \mu \text{ and } \phi(\mu_n) \to \phi(\mu). \end{gathered} \tag{9.4.61}$$

If we take $D := C$, we see that $D \subset D(\boldsymbol{\partial}\phi) \subset D(\phi)$. If $\mu \in D(\boldsymbol{\partial}\phi)$ and $(\mu_n) \subset D$, $\Phi_n \in -\boldsymbol{\partial}\phi(\mu_n)$ are such that $\mu_n \to \mu$ and $\sup_n |\Phi_n|_2 < +\infty$, by lower semicontinuity of ϕ we have that $\liminf_n \phi(\mu_n) \geqslant \phi(\mu)$. Moreover

$$\phi(\mu) - \phi(\mu_n) \geqslant [\Phi_n, \mu]_r \geqslant -|\Phi_n|_2 W_2(\mu_n, \mu)$$

so that

$$\phi(\mu_n) \leqslant \phi(\mu) + |\Phi_n|_2 W_2(\mu_n, \mu)$$

and passing to the $\limsup_n$ shows that $\phi(\mu_n) \to \phi(\mu)$. This shows that D is dense in energy in $D(\boldsymbol{\partial}\phi)$. Since this is dense in energy in $D(\phi)$ we can conclude by a diagonal argument. □

9.5 EXAMPLES OF DISSIPATIVE MPVFS AND FLOWS

In this section we give some examples of MPVFs that fit our framework. The first subsection 9.5.1 is devoted to subdifferentials of functionals. In Subsection 9.5.4, we give some examples of MPVFs generating λ-flows with particular properties. We then conclude with Subsection 9.5.5, where we compare our framework with that developed in [94].

9.5.1 *Subdifferentials of λ-convex functionals*

Recall that a functional $\phi : \mathcal{P}_2(\mathbb{H}) \to (-\infty, +\infty]$ is λ-(geodesically) convex on $\mathcal{P}_2(\mathbb{H})$ (see [5, Definition 9.1.1]) if for any μ_0, μ_1 in the proper domain $D(\phi) := \{\mu \in \mathcal{P}_2(\mathbb{H}) \mid \phi(\mu) < +\infty\}$ there exists $\boldsymbol{\mu} \in \Gamma_o(\mu_0, \mu_1)$ such that

$$\phi(\mu_t) \leqslant (1-t)\phi(\mu_0) + t\phi(\mu_1) - \frac{\lambda}{2}t(1-t)W_2^2(\mu_0, \mu_1)$$

for every $t \in [0, 1]$, where $\mu : [0, 1] \to \mathcal{P}_2(\mathbb{H})$ is the constant speed geodesic induced by $\boldsymbol{\mu}$, i.e. $\mu_t = x^t_\sharp \boldsymbol{\mu}$.

The *Fréchet subdifferential* $\boldsymbol{\partial}\phi$ of ϕ [5, Definition 10.3.1] is a MPVF which can be characterized [5, Theorem 10.3.6] by

$$\Phi \in \boldsymbol{\partial}\phi[\mu] \quad \Leftrightarrow \quad \mu \in D(\phi),\ \phi(\nu) - \phi(\mu) \geqslant -[\Phi, \nu]_l + \frac{\lambda}{2}W_2^2(\mu, \nu).$$

for every $\nu \in D(\phi)$. According to the notation introduced in (7.1.10), we set

$$-\boldsymbol{\partial}\phi[\mu] = J_\sharp \boldsymbol{\partial}\phi[\mu], \quad \text{with } J(x, v) := (x, -v), \tag{9.5.1}$$

and we have the following result.

Theorem 9.5.1. *If $\phi : \mathcal{P}_2(\mathbb{H}) \to (-\infty, +\infty]$ is a proper, lower semicontinuous and λ-convex functional, then $-\boldsymbol{\partial}\phi$ is a $(-\lambda)$*-dissipative *MPVF according to (7.5.1).*

In the following proposition, we prove a correspondence between gradient flows for ϕ and $(-\lambda)$-EVI solutions for the MPVF $-\boldsymbol{\partial}\phi$. We refer respectively to (7.6.2), (7.6.7) and Definition 7.6.4 for the definitions of $I(\mu|\mathbf{F})$, $\Gamma^0_o(\cdot, \cdot|\mathbf{F})$ and $[\mathbf{F}, \mu]_{0+}$.

Proposition 9.5.2. *Let $\phi : \mathcal{P}_2(\mathbb{H}) \to (-\infty, +\infty]$ be a proper, lower semicontinuous and λ-convex functional and let $\mu : \mathcal{I} \to D(\boldsymbol{\partial}\phi)$ be a locally absolutely continuous curve, with $\mathcal{I}$ a (bounded or unbounded) interval in $\mathbb{R}$. Then*

1. *if μ is a Gradient Flow for ϕ i.e.*

$$(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t \in -\boldsymbol{\partial}\phi(\mu_t) \quad \text{a.e. } t \in \mathcal{I},$$

then μ is a $(-\lambda)$-EVI solution of (8.1.1) for the MPVF $-\boldsymbol{\partial}\phi$ as in (9.5.1);

2. *if μ is a $(-\lambda)$-EVI solution of (8.1.1) for the MPVF $-\boldsymbol{\partial}\phi$ and the domain of $\boldsymbol{\partial}\phi$ satisfies*

$$\text{for a.e. } t \in \mathcal{I},\ \Gamma^0_o(\mu_t, \nu|\boldsymbol{\partial}\phi) \neq \emptyset \quad \text{for every } \nu \in D(\boldsymbol{\partial}\phi),$$

then μ is a Gradient Flow for ϕ.

Proof. The first assertion is a consequence Theorem 8.1.4(1). We prove the second claim; by (8.1.5b) we have that for a.e. $t \in \mathcal{I}$ it holds

$$\left[(\mathbf{i}_{\mathbb{H}}, \boldsymbol{\nu}_t)_\sharp \mu_t, \nu\right]_r \leqslant [(\mathbf{i}_{\mathbb{H}}, \boldsymbol{\nu}_t)_\sharp \mu_t, \mu_t]_{r,0} \leqslant [-\boldsymbol{\partial}\phi, \mu_t]_{0+}$$

for every $\nu \in D(\phi)$ and $\boldsymbol{\mu}_t \in \Gamma^0_o(\mu_t, \nu|\boldsymbol{\partial}\phi)$. We show that for every $\nu_0, \nu_1 \in D(\boldsymbol{\partial}\phi)$ and every $\boldsymbol{\nu} \in \Gamma^0_o(\nu_0, \nu_1|\mathbf{F})$

$$[-\boldsymbol{\partial}\phi, \boldsymbol{\nu}]_{0+} \leqslant \phi(\nu_1) - \phi(\nu_0) - \frac{\lambda}{2} W_2^2(\nu_0, \nu_1). \tag{9.5.2}$$

To prove that, we take $s \in I(\boldsymbol{\nu}|\boldsymbol{\partial}\phi) \cap (0,1)$ and $\Phi_s \in -\boldsymbol{\partial}\phi(\nu_s)$, where we have set $\nu_s := x^s_\sharp \boldsymbol{\nu}$. By definition of subdifferential we have

$$[\Phi_s, \nu_1]_r \leqslant \phi(\nu_1) - \phi(\nu_s) - \frac{\lambda}{2} W_2^2(\nu_s, \nu_1).$$

Dividing by $(1-s)$, using (7.4.6) and passing to the infimum w.r.t. $\Phi_s \in -\boldsymbol{\partial}\phi(\nu_s)$ we obtain

$$[-\boldsymbol{\partial}\phi, \boldsymbol{\nu}]_{r,s} \leqslant \frac{1}{1-s}\left(\phi(\nu_1) - \phi(\nu_s)\right) - \frac{\lambda(1-s)}{2} W_2^2(\nu_0, \nu_1).$$

Passing to the limit as $s \downarrow 0$ and using the lower semicontinuity of ϕ lead to the result. Once that (9.5.2) is established we have that for a.e. $t \in \mathcal{I}$ it holds

$$\left[(\mathbf{i}_{\mathbb{H}}, \boldsymbol{\nu}_t)_\sharp \mu_t, \nu\right]_r \leqslant \phi(\nu) - \phi(\mu_t) - \frac{\lambda}{2} W_2^2(\mu_t, \nu) \quad \text{for every } \nu \in D(\boldsymbol{\partial}\phi). \tag{9.5.3}$$

To conclude it is enough to use the lower semicontinuity of the l.h.s. (see Lemma 7.3.1) and the fact that $D(\boldsymbol{\partial}\phi)$ is dense in $D(\phi)$ in energy: indeed we can apply [86, Corollary 4.5] and [5, Lemma 3.1.2] to the proper, lower semicontinuous and convex functional $\phi^\lambda : \mathcal{P}_2(\mathbb{H}) \to (-\infty, +\infty]$ defined as

$$\phi^\lambda(\nu) = \phi(\nu) - \frac{\lambda}{2} m_2^2(\nu)$$

to get the existence, for every $\nu \in D(\phi)$, of a family $(\nu^\tau)_{\tau>0} \subset D(\phi^\lambda) = D(\phi)$ s.t.

$$\nu^\tau \to \nu, \quad \phi^\lambda(\nu^\tau) \to \phi^\lambda(\nu) \quad \text{as } \tau \downarrow 0.$$

Of course $\phi(\nu^\tau) \to \phi(\nu)$ as $\tau \downarrow 0$ and, applying [5, Lemma 10.3.4], we see that $\nu^\tau \in D(\boldsymbol{\partial}\phi^\lambda)$. However $\boldsymbol{\partial}\phi^\lambda = L^\lambda_\sharp \boldsymbol{\partial}\phi$ (see (7.5.5)) so that $\nu^\tau \in D(\boldsymbol{\partial}\phi)$. We can thus write (9.5.3) for ν^τ in place of ν and pass to the limit as $\tau \downarrow 0$, obtaining that, by definition of subdifferential, $(\mathbf{i}_{\mathbb{H}}, \boldsymbol{\nu}_t)_\sharp \mu_t \in -\boldsymbol{\partial}\phi(\mu_t)$ for a.e. $t \in \mathcal{I}$. □

Referring to [5], here we list interesting and explicit examples of $(-\lambda)$-dissipative MPVFs, according to (7.5.1), induced by proper, lower semicontinuous and λ-convex functionals, focusing on the cases when $D(\boldsymbol{\partial}\phi) = \mathcal{P}_2(\mathbb{H})$.

1. *Potential energy.* Let $P : \mathbb{H} \to \mathbb{R}$ be a l.s.c. and λ-convex functional satisfying

$$|\partial^o P(x)| \leqslant C(1 + |x|) \quad \text{for every } x \in \mathbb{H},$$

for some constant $C > 0$, where $\partial^o P(x)$ is the element of minimal norm in $\partial P(x)$. By [5, Proposition 10.4.2] the PVF

$$\mathbf{F}[\mu] := (\boldsymbol{i}_{\mathbb{H}}, -\partial^o P)_\sharp \mu, \quad \mu \in \mathcal{P}_2(\mathbb{H}),$$

is a $(-\lambda)$-dissipative selection of $-\boldsymbol{\partial}\mathcal{F}_P$ for the potential energy functional

$$\mathcal{F}_P(\mu) := \int_{\mathbb{H}} P \, d\mu, \quad \mu \in \mathcal{P}_2(\mathbb{H}).$$

2. *Interaction energy.* If $W : \mathbb{H} \to [0, +\infty)$ is an even, differentiable, and λ-convex function for some $\lambda \in \mathbb{R}$, whose differential has a linear growth, then, by [5, Theorem 10.4.11], the PVF

$$\mathbf{F}[\mu] := (\boldsymbol{i}_{\mathbb{H}}, (-\nabla W * \mu))_\sharp \, \mu, \quad \mu \in \mathcal{P}_2(\mathbb{H}),$$

is a $(-\lambda)$-dissipative selection of $-\boldsymbol{\partial}\mathcal{F}_W$, the opposite of the Wasserstein subdifferential of the interaction energy functional

$$\phi_W(\mu) := \frac{1}{2} \int_{\mathbb{H}^2} W(x-y) \, d(\mu \otimes \mu)(x,y), \quad \mu \in \mathcal{P}_2(\mathbb{H}).$$

3. *Opposite Wasserstein distance.* Let $\bar{\mu} \in \mathcal{P}_2(\mathbb{H})$ be fixed and consider the functional $\phi_{\text{Wass}} : \mathcal{P}_2(\mathbb{H}) \to \mathbb{R}$ defined as

$$\phi_{\text{Wass}}(\mu) := -\frac{1}{2} W_2^2(\mu, \bar{\mu}), \quad \mu \in \mathcal{P}_2(\mathbb{H}),$$

which is geodesically (-1)-convex [5, Proposition 9.3.12]. Setting

$$\mathbf{b}(\mu) := \arg\min \left\{ \int_{\mathbb{H}} |\mathbf{b}(x) - x|^2 \, d\mu \, : \, \mathbf{b} = \mathbf{b}_{\boldsymbol{\gamma}} \in L^2_\mu(\mathbb{H}; \mathbb{H}), \; \boldsymbol{\gamma} \in \Gamma_o(\mu, \bar{\mu}) \right\},$$

the PVF

$$\mathbf{F}[\mu] := (\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{i}_{\mathbb{H}} - \mathbf{b}(\mu))_\# \mu, \quad \mu \in \mathcal{P}_2(\mathbb{H})$$

is a selection of $-\boldsymbol{\partial}\phi_{\text{Wass}}(\mu)$ and it is therefore 1-dissipative according to (7.5.1).

9.5.2 *MPVF concentrated on the graph of a multifunction*

The previous example of Section 9.5.1 has a natural generalization in terms of dissipative graphs in $\mathbb{H} \times \mathbb{H}$ [12, 13, 26]. We consider a (non-empty) λ-dissipative set $F \subset \mathbb{H} \times \mathbb{H}$, i.e. satisfying

$$\langle v_0 - v_1, x_0 - x_1 \rangle \leqslant \lambda |x_0 - x_1|^2 \quad \text{for every } (x_0, v_0), (x_1, v_1) \in F.$$

The corresponding MPVF defined as

$$\mathbf{F} := \left\{ \Phi \in \mathcal{P}_2(\mathsf{T}\mathbb{H}) \mid \Phi \text{ is concentrated on } F \right\}$$

is λ-dissipative as well, according to (7.5.1). In fact, if $\Phi_0, \Phi_1 \in \mathbf{F}$ with $\nu_i = x_\sharp \Phi_i$, $i = 0, 1$, and $\boldsymbol{\Theta} \in \Lambda(\Phi_0, \Phi_1)$ then $(x_0, v_0, x_1, v_1) \in F \times F$ $\boldsymbol{\Theta}$-a.e., so that

$$\int_{\mathsf{TH}\times\mathsf{TH}} \langle v_0 - v_1, x_0 - x_1\rangle \, d\boldsymbol{\Theta}(x_0, v_0, x_1, v_1) \leqslant \lambda \int_{\mathsf{TH}\times\mathsf{TH}} |x_0 - x_1|^2 \, d\boldsymbol{\Theta} \\ = \lambda W_2^2(\nu_0, \nu_1).$$

since $(x^0, x^1)_\sharp \boldsymbol{\Theta} \in \Gamma_o(\nu_0, \nu_1)$. Taking the supremum w.r.t. $\boldsymbol{\Theta} \in \Lambda(\Phi_0, \Phi_1)$ we obtain $[\Phi_0, \Phi_1]_l \leqslant \lambda W_2^2(\nu_0, \nu_1)$ which is even stronger than λ-dissipativity. If $D(F) = \mathbb{H}$ then $D(\mathbf{F})$ contains $\mathcal{P}_c(\mathbb{H})$, the set of Borel probability measures with compact support. If F has also a linear growth, then it is easy to check that $D(\mathbf{F}) = \mathcal{P}_2(\mathbb{H})$ as well.

Despite the analogy just shown with dissipative operators in Hilbert spaces, there are important differences with the Wasserstein framework, as highlighted in the following examples. In particular, in Subsection 6.2 we showed how dissipativity allows to deduce relevant properties when the MPVF $\mathbf{F}$ is tested against optimal directions. On the contrary, whenever $v_\sharp \mathbf{F}[\mu]$ is orthogonal to $\mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{H})$, we are not able to deduce information through the dissipativity assumption, as shown in Example 9.5.3 and Example 9.5.4.

Example 9.5.3. Let $\mathbb{H} = \mathbb{R}^2$, let $B := \{x \in \mathbb{R}^2 \mid |x| \leqslant 1\}$ be the closed unit ball, let $\mathcal{L}_B$ be the (normalized) Lebesgue measure on B, and let $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{r}(x_1, x_2) = (x_2, -x_1)$ be the anti-clockwise rotation of $\pi/2$ degrees. We define the MPVF

$$\mathbf{F}[\nu] = \begin{cases} (\mathbf{i}_{\mathbb{R}^2}, 0)_\sharp \nu, & \text{if } \nu \in \mathcal{P}_2(\mathbb{R}^2) \setminus \{\mathcal{L}_B\}, \\ \left\{(\mathbf{i}_{\mathbb{R}^2}, a\mathbf{r})_\sharp \mathcal{L}_B \mid a \in \mathbb{R}\right\}, & \text{if } \nu = \mathcal{L}_B. \end{cases}$$

Observe that $D(\mathbf{F}) = \mathcal{P}_2(\mathbb{R}^2)$ and $\mathbf{F}$ is obviously unbounded at $\nu = \mathcal{L}_B$, i.e.

$$\sup\{|\Phi|_2 \,:\, \Phi \in \mathbf{F}[\mathcal{L}_B]\} = +\infty.$$

The MPVF $\mathbf{F}$ is also dissipative with $\lambda = 0$ according to (7.5.1): indeed, thanks to Remark 7.1.5 it is enough to check that

$$\left[(\mathbf{i}_{\mathbb{R}^2}, a\mathbf{r})_\sharp \mathcal{L}_B, \nu\right]_r = 0 \quad \text{for every } \nu \in \mathcal{P}_2(\mathbb{R}^2),\ a \in \mathbb{R}. \tag{9.5.4}$$

To prove (9.5.4), we notice that the optimal transport plan from $\mathcal{L}_B$ to ν is concentrated on a map which belongs to the tangent space $\mathrm{Tan}_{\mathcal{L}_B} \mathcal{P}_2(\mathbb{R}^2)$ [5, Prop. 8.5.2]; by Remark 7.4.2 we have just to check that

$$\int_{\mathbb{R}^2} \langle \mathbf{r}(x), \nabla \varphi(x)\rangle \, d\mathcal{L}_B(x) = 0 \quad \text{for every } \varphi \in C_c^\infty(\mathbb{R}^2),$$

that is a consequence of the Divergence Theorem on B. This example is in contrast with the Hilbertian theory of dissipative operators according to which an everywhere defined dissipative operator is locally bounded (see [26, Proposition 2.9]).

Example 9.5.4. In the same setting of the previous example, let us define the MPVF

$$\mathbf{F}[\nu] = (\mathbf{i}_{\mathbb{R}^2}, \mathbf{r})_\sharp \nu, \quad \mathbf{r}(x_1, x_2) = (x_2, -x_1), \quad \nu \in \mathcal{P}_2(\mathbb{R}^2).$$

It is easy to check that $\mathbf{F}$ is dissipative according to (7.5.1) and Lipschitz continuous (as a map from $\mathcal{P}_2(\mathbb{R}^2)$ to $\mathcal{P}_2(\mathsf{T}\mathbb{R}^2)$). Moreover, arguing as in Example 9.5.3, we can show that $(\boldsymbol{i}_{\mathbb{R}^d}, 0)_\sharp \mathcal{L}_B \in \hat{\mathbf{F}}[\mathcal{L}_B]$, where $\hat{\mathbf{F}}$ is defined in (7.7.4). This is again in contrast with the Hilbertian theory of dissipative operators, stating that a single valued, everywhere defined, and continuous dissipative operator coincides with its maximal extension (see [26, Proposition 2.4]).

9.5.3 *Interaction field induced by a dissipative map*

Let us consider the Hilbert space $\mathsf{Y} = \mathbb{H}^n$, $n \in \mathbb{N}$, endowed with the scalar product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \frac{1}{n}\sum_{i=1}^n \langle x_i, y_i \rangle$, for every $\boldsymbol{x} = (x_i)_{i=1}^n$, $\boldsymbol{y} = (y_i)_{i=1}^n \in \mathbb{H}^n$. We identify TY with $(\mathsf{T}\mathbb{H})^n$ and we denote by x^i, v^i the i-th coordinate maps. Every permutation $\sigma : \{1, \cdots, n\} \to \{1, \cdots, n\}$ in $\mathrm{Sym}(n)$ operates on Y by the obvious formula $\sigma(\boldsymbol{x})_i = x_{\sigma(i)}$, $i = 1, \cdots, n$, $\boldsymbol{x} \in \mathsf{Y}$.

Let $G : \mathsf{Y} \to \mathsf{Y}$ be a Borel λ-dissipative map bounded on bounded sets (this property is always true if Y has finite dimension) and satisfying

$$\boldsymbol{x} \in D(G) \quad \Rightarrow \quad \sigma(\boldsymbol{x}) \in D(G),\ G(\sigma(\boldsymbol{x})) = \sigma(G(\boldsymbol{x})) \quad \text{for every permutation } \sigma. \tag{9.5.5}$$

Denoting by $(G^1, \cdots, G^n)$ the components of G, by x^i the projections from Y to $\mathbb{H}$ and by $\mu^{\otimes n} = \bigotimes_{i=1}^n \mu$, we have that the MPVF

$$\mathbf{F}[\mu] := (x^1, G^1)_\sharp \mu^{\otimes n} \quad \text{with domain } D(\mathbf{F}) := \mathcal{P}_b(\mathbb{H})$$

is λ-dissipative as well according to (7.5.1). Indeed, let $\mu, \nu \in D(\mathbf{F})$, $\boldsymbol{\gamma} \in \Gamma_o(\mu, \nu)$ and let

$$\Phi = (x^1, G^1)_\sharp \mu^{\otimes n} \quad \text{and} \quad \Psi = (x^1, G^1)_\sharp \nu^{\otimes n}.$$

We can consider the plan $\boldsymbol{\beta} := P_\sharp \boldsymbol{\gamma}^{\otimes n} \in \Gamma(\mu^{\otimes n}, \nu^{\otimes n})$, where

$$P((x_1, y_1), \cdots, (x_n, y_n)) := ((x_1, \cdots, x_n), (y_1, \cdots, y_n)).$$

Considering the map $H^1(\boldsymbol{x}, \boldsymbol{y}) := (x_1, G^1(\boldsymbol{x}), y_1, G^1(\boldsymbol{y}))$ we have $\boldsymbol{\Theta} := H^1_\sharp \boldsymbol{\beta} \in \Lambda(\Phi, \Psi)$, so that

$$\begin{aligned}
[\Phi, \Psi]_r &\leqslant \int \langle v_1 - w_1, x_1 - y_1 \rangle \, \mathrm{d}\boldsymbol{\Theta}(x_1, v_1, y_1, w_1) \\
&= \int \langle G^1(\boldsymbol{x}) - G^1(\boldsymbol{y}), x_1 - y_1 \rangle \, \mathrm{d}\boldsymbol{\beta}(\boldsymbol{x}, \boldsymbol{y}) \\
&= \frac{1}{n} \sum_{k=1}^n \int \langle G^k(\boldsymbol{x}) - G^k(\boldsymbol{y}), x_k - y_k \rangle \, \mathrm{d}\boldsymbol{\beta}(\boldsymbol{x}, \boldsymbol{y}) \\
&= \int \langle G(\boldsymbol{x}) - G(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \, \mathrm{d}\boldsymbol{\beta}(\boldsymbol{x}, \boldsymbol{y}),
\end{aligned}$$

where we used (9.5.5) and the invariance of $\boldsymbol{\beta}$ with respect to permutations. The λ-dissipativity of G then yields

$$\begin{aligned}\int \langle G(\boldsymbol{x}) - G(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \, d\boldsymbol{\beta}(\boldsymbol{x}, \boldsymbol{y}) &\leqslant \lambda \int |\boldsymbol{x} - \boldsymbol{y}|_{\mathsf{Y}}^2 \, d\boldsymbol{\beta}(\boldsymbol{x}, \boldsymbol{y}) \\ &= \lambda \frac{1}{n} \sum_{k=1}^{n} \int |x_k - y_k|_{\mathsf{Y}}^2 \, d\boldsymbol{\beta}(\boldsymbol{x}, \boldsymbol{y}) \\ &= \lambda \frac{1}{n} \sum_{k=1}^{n} \int |x_k - y_k|_{\mathsf{Y}}^2 \, d\boldsymbol{\gamma}(x_k, y_k) \\ &= \lambda W_2^2(\mu, \nu).\end{aligned}$$

A typical example when $n = 2$ is provided by

$$G(x_1, x_2) := (A(x_1 - x_2), A(x_2 - x_1))$$

where $A : \mathbb{H} \to \mathbb{H}$ is a Borel, locally bounded, dissipative and antisymmetric map satisfying $A(-z) = -A(z)$. We easily get

$$\begin{aligned}&\langle G(\boldsymbol{x}) - G(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \\ &= \frac{1}{2}\Big(\langle A(x_1 - x_2) - A(y_1 - y_2), x_1 - y_1 \rangle - \langle A(x_1 - x_2) - A(y_1 - y_2), x_2 - y_2 \rangle \Big) \\ &= \frac{1}{2} \langle A(x_1 - x_2) - A(y_1 - y_2), x_1 - x_2 - (y_1 - y_2) \rangle \leqslant 0.\end{aligned}$$

In this case

$$\mathbf{F}[\mu] = (\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{a}[\mu])_\sharp \mu, \quad \boldsymbol{a}[\mu](x) = \int_{\mathbb{H}} A(x - y) \, d\mu(y) \quad \text{for every } x \in \mathbb{H}.$$

9.5.4 *A few borderline examples*

In this subsection, we collect a few examples which reveal the importance of some of the technical tools we developed in Section 8.1. First of all we exhibit an example of dissipative MPVF generating a 0-flow, for which solutions starting from given initial data are merely continuous. In particular, the nice regularizing effect of gradient flows (see [25] for the Hilbert case and [5, Theorem 4.0.4, Theorem 11.2.1] for the general metric and Wasserstein settings), according to which a solution belongs to the domain of the functional for any $t > 0$ even if the initial datum merely belongs to its closure, does not hold for general dissipative evolutions. This also clarifies the interest in a definition of continuous, not necessarily absolutely continuous, solution given in Definition 8.1.1.

Example 9.5.5 (Lifting of dissipative evolutions and lack of regularizing effect). Let us consider the situation of Corollary 9.2.15, choosing the Hilbert space $\mathbb{H} = \ell^2(\mathbb{N})$. Following [106, Example 3] we can easily find a maximal linear dissipative operator $A : D(A) \subset \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ whose semigroup does not provide a regularizing effect. We define A as

$$A(x_1, x_2, \ldots, x_{2k-1}, x_{2k}, \ldots) = (-x_2, x_1, \ldots, -kx_{2k}, kx_{2k-1}, \ldots), \quad x \in D(A),$$

with domain

$$D(A) := \left\{ x \in \ell^2(\mathbb{N}) : \sum_{k=1}^{\infty} k^2 |x_k|^2 < \infty \right\},$$

so that there is no regularizing effect for the semigroup $(R_t)_{t\geqslant 0}$ generated by (the graph of) A: evolutions starting outside the domain $D(A)$ stay outside the domain and do not give raise to locally Lipschitz or a.e. differentiable curves. Corollary 9.2.15 shows that the 0-flow $(S_t)_{t\geqslant 0}$ generated by $\mathbf{F}$ on $\mathcal{P}_2(X)$ is given by

$$S_t[\mu_0] = (R_t)_\sharp \mu_0 \qquad \text{for every } \mu_0 \in \overline{D(\mathbf{F})} = \mathcal{P}_2(\mathbb{H})$$

so that there is the same lack of regularizing effect on probability measures.

In the next example we show that a constant MPVF generates a barycentric solution.

Example 9.5.6 (Constant PVF and barycentric evolutions). Given $\theta \in \mathcal{P}_2(\mathbb{H})$, we consider the constant PVF

$$\mathbf{F}[\mu] := \mu \otimes \theta.$$

$\mathbf{F}$ is dissipative according to (7.5.1): in fact, if $\Phi_i = \mu_i \otimes \theta$, $i = 0, 1$, $\boldsymbol{\mu} \in \Gamma_o(\mu_0, \mu_1)$, and $\mathbf{r} : \mathbb{H} \times \mathbb{H} \times X \to \mathsf{T}\mathbb{H} \times \mathsf{T}\mathbb{H}$ is defined by $\mathbf{r}(x_0, x_1, v) := (x_0, v; x_1, v)$, then

$$\Theta = \mathbf{r}_\sharp(\boldsymbol{\mu} \otimes \theta) \in \Lambda(\Phi_0, \Phi_1)$$

so that (7.1.12) yields

$$[\Phi_0, \Phi_1]_r \leqslant \int \langle x_0 - x_1, v - v \rangle \, d(\boldsymbol{\mu} \otimes \theta)(x_0, x_1, v) = 0.$$

Applying Proposition 9.2.11 and Theorem 9.2.9 we immediately see that $\mathbf{F}$ generates a 0-flow $(S_t)_{t\geqslant 0}$ in $\mathcal{P}_2(\mathbb{H})$, obtained as a limit of the Explicit Euler scheme. It is also straightforward to notice that we can apply Theorem 8.3.4 to $\mathbf{F}$ so that for every $\mu_0 \in \mathcal{P}_2(\mathbb{H})$ the unique EVI solution $\mu_t = S_t \mu_0$ satisfies the continuity equation

$$\partial_t \mu_t + \nabla \cdot (\mathbf{b} \mu_t) = 0, \quad \mathbf{b} = \int_{\mathbb{H}} v \, d\theta(v).$$

Since $\mathbf{b}$ is constant, we deduce that S_t acts as a translation with constant velocity $\mathbf{b}$, i.e.

$$\mu_t = (\mathbf{i}_{\mathbb{H}} + t\mathbf{b})_\sharp \mu_0,$$

so that S_t coincides with the semigroup generated by the PVF $\mathbf{F}'[\mu] := (\mathbf{i}_{\mathbb{H}}, \mathbf{b})_\sharp \mu$.

We conclude this subsection with a 1-dimensional example of a curve which satisfies the barycentric property but it is not an EVI solution.

Example 9.5.7. Let $\mathbb{H} = \mathbb{R}$. It is well known (see e.g. [88]) that $\mathcal{P}_2(\mathbb{R})$ is isometric to the closed convex subset $\mathcal{K} \subset L^2(0,1)$ of the (essentially) increasing maps under the action of the isometry $\mathcal{J} : \mathcal{P}_2(\mathbb{R}) \to \mathcal{K}$ which maps each measure $\mu \in \mathcal{P}_2(\mathbb{R})$ into the pseudo inverse of its cumulative distribution function.

It follows that for every $\bar{\nu} \in \mathcal{P}_2(\mathbb{R})$ the functional $\phi : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ defined as

$$\phi(\mu) := \frac{1}{2} W_2^2(\mu, \bar{\nu})$$

is 1-convex, since it satisfies $\phi(\mu) = \mathcal{G}(\mathcal{J}(\mu))$ where $\mathcal{G} : L^2(0,1) \to \mathbb{R}$ is defined as

$$\mathcal{G}(u) := \begin{cases} \frac{1}{2}\|u - \mathcal{J}(\bar{\nu})\|^2 & \text{if } u \in \mathcal{K}, \\ +\infty & \text{otherwise.} \end{cases}$$

Thus ϕ generates a gradient flow $(S_t)_{t \geqslant 0}$ which is a semigroup of contractions in $\mathcal{P}_2(\mathbb{R})$; for every $\mu_0 \in \mathcal{P}_2(\mathbb{R})$, the map $S_t[\mu_0]$ is the unique (-1)-EVI solution for the MPVF $-\boldsymbol{\partial}\phi$ starting from $\mu_0 \in \mathcal{P}_2(\mathbb{R})$ (see Proposition 9.5.2). Since the notion of gradient flow is purely metric, the gradient flow of $\mathcal{G}$ starting from $\mathcal{J}(\mu_0)$ is just the image through $\mathcal{J}$ of the gradient flow of ϕ starting from $\mu_0 \in \mathcal{P}_2(\mathbb{R})$. Indeed: let μ be the gradient flow for ϕ starting from $\mu_0 \in \mathcal{P}_2(\mathbb{R})$, then by e.g. [5, Theorem 11.1.4] we have that μ satisfies

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t, \nu) \leqslant \phi(\nu) - \phi(\mu_t) - \frac{1}{2}W_2^2(\mu_t, \nu) \quad \text{for a.e. } t > 0, \text{ for } \nu \in \mathcal{P}_2(\mathbb{R}),$$

so that we get

$$\frac{d}{dt}\frac{1}{2}\|\mathcal{J}(\mu_t) - \mathcal{J}(\nu)\|^2 \leqslant \mathcal{G}(\mathcal{J}(\nu)) - \mathcal{G}(\mathcal{J}(\mu_t)) - \frac{1}{2}\|\mathcal{J}(\mu_t) - \mathcal{J}(\nu)\|^2,$$

which, recalling the characterization of gradient flows in Hilbert spaces, gives that $u(t) := \mathcal{J}(\mu_t)$ is the gradient flow of $\mathcal{G}$ starting from $\mathcal{J}(\mu_0)$.

It is easy to check that

$$u(t) := e^{-t}\mathcal{J}(\mu_0) + (1 - e^{-t})\mathcal{J}(\bar{\nu})$$

is the gradient flow of $\mathcal{G}$ starting from $u_0 = \mathcal{J}(\mu_0)$. Note that $u(t)$ is the $L^2(0,1)$ geodesic from $\mathcal{J}(\bar{\nu})$ to $\mathcal{J}(\mu_0)$ evaluated at the rescaled time e^{-t}, so that $S_t[\mu_0]$ must coincide with the evaluation at time e^{-t} of the (unique) geodesic connecting $\bar{\nu}$ to μ_0 i.e.

$$S_t[\mu_0] = x^s_\sharp \boldsymbol{\gamma}, \quad s = e^{-t} \in (0,1],$$

where $\boldsymbol{\gamma} \in \Gamma_o(\bar{\nu}, \mu_0)$.

Let us now consider the particular case $\bar{\nu} = \frac{1}{2}\delta_{-a} + \frac{1}{2}\delta_a$, where $a > 0$ is a fixed parameter and $\mu_0 = \delta_0$. It is straightforward to see that

$$\mu_t = S_t[\delta_0] = \frac{1}{2}\delta_{a(1-e^{-t})} + \frac{1}{2}\delta_{a(e^{-t}-1)}, \quad t \geqslant 0$$

so that

$$(\boldsymbol{i}_{\mathbb{H}}, \boldsymbol{v}_t)_\sharp \mu_t = \frac{1}{2}\delta_{((1-e^{-t})a, e^{-t}a)} + \frac{1}{2}\delta_{((e^{-t}-1)a, -e^{-t}a)} \in -\boldsymbol{\partial}\phi(\mu_t), \quad \text{a.e. } t > 0,$$

where $\boldsymbol{v}$ is the Wasserstein velocity field of μ_t. On the other hand, [5, Lemma 10.3.8] shows that

$$\delta_0 \otimes \left(\frac{1}{2}\delta_{-a} + \frac{1}{2}\delta_a\right) \in -\boldsymbol{\partial}\phi(\delta_0)$$

so that the constant curve $\bar{\mu}_t := \delta_0$ for $t \geqslant 0$ has the barycentric property for the MPVF $-\boldsymbol{\partial}\phi$ but it is not a EVI solution for $-\boldsymbol{\partial}\phi$, being different from $\mu_t = S_t[\delta_0]$.

9.5.5 *Comparison with [94]*

In this section, we provide a brief comparison between the assumptions we required in order to develop a strong concept of solution to (8.1.1) and the hypotheses assumed in [94]. We remind that the relation between our solution and the weaker notion studied in [94] was exploited in Section 8.3. Here, we conclude with a further remark coming from the connections between our approximating scheme proposed in (EE) and the schemes proposed in [31] and [94].

We consider a finite time horizon $[0, T]$ with $T > 0$, the space $\mathbb{H} = \mathbb{R}^d$ and we deal with measures in $\mathcal{P}_b(\mathbb{R}^d)$ and in $\mathcal{P}_b(\mathsf{T}\mathbb{R}^d)$, i.e. compactly supported. We also deal with *single-valued* probability vector fields (PVF) for simplicity, which can be considered as everywhere defined maps $\mathbf{F} : \mathcal{P}_b(\mathbb{R}^d) \to \mathcal{P}_b(\mathsf{T}\mathbb{R}^d)$ such that $x_\sharp \mathbf{F}[\nu] = \nu$. This is indeed the framework examined in [94].

We start by recalling the assumptions required in [94] for a PVF $\mathbf{F} : \mathcal{P}_b(\mathbb{R}^d) \to \mathcal{P}_b(\mathsf{T}\mathbb{R}^d)$.

(H1) there exists a constant $M > 0$ such that for all $\nu \in \mathcal{P}_b(\mathbb{R}^d)$,

$$\sup_{(x,v)\in \operatorname{supp}(\mathbf{F}[\nu])} |v| \leqslant M\left(1 + \sup_{x\in\operatorname{supp}(\nu)} |x|\right);$$

(H2) $\mathbf{F}$ satisfies the following Lipschitz condition: there exists a constant $L \geqslant 0$ such that for every $\Phi = \mathbf{F}[\nu]$, $\Phi' = \mathbf{F}[\nu']$ there exists $\boldsymbol{\Theta} \in \Lambda(\Phi, \Phi')$ satisfying

$$\int_{\mathsf{T}\mathbb{R}^d\times\mathsf{T}\mathbb{R}^d} |v_0 - v_1|^2 \,d\boldsymbol{\Theta}(x_0, v_0, x_1, v_1) \leqslant L^2 W_2^2(\nu, \nu'),$$

with $\Lambda(\cdot, \cdot)$ as in Definition 7.1.7.

Remark 9.5.8. Condition (H1) is (H:bound) in [94], while (H1) corresponds to (H:lip) in [94] in case $p = 2$ (see also Remark 5 in [94]).
We stress that actually in [94] condition (H2) is local, meaning that L is allowed to depend on the radius R of a ball centered at 0 and containing the supports of ν and ν'. Thanks to assumption (H1), it is easy to show that for every final time T all the discrete solutions of the Explicit Euler scheme and of the scheme of [94] starting from an initial measure with support in $B(0, R)$ are supported in a ball $B(0, R')$ where R' solely depends on R and T. We can thus restrict the PVF $\mathbf{F}$ to the (geodesically convex) set of measures with support in $B(0, R')$ and act as L does not depend on the support of the measures.

Proposition 9.5.9. *If* $\mathbf{F} : \mathcal{P}_b(\mathbb{R}^d) \to \mathcal{P}_b(\mathsf{T}\mathbb{R}^d)$ *is a PVF satisfying (H2), then* $\mathbf{F}$ *is* λ*-dissipative according to* (7.5.1) *for* $\lambda = \frac{L^2+1}{2}$*, the Explicit Euler scheme is globally solvable in* $D(\mathbf{F})$*, and* $\mathbf{F}$ *generates a* λ*-flow, whose trajectories are the limit of the Explicit Euler scheme in each finite interval* $[0,T]$.

Proof. The λ-dissipativity comes from Lemma 7.5.7. We prove that (9.2.14) holds. Let $\nu \in D(\mathbf{F})$ and take $\boldsymbol{\Theta} \in \Lambda(\mathbf{F}[\nu], \mathbf{F}[\delta_0])$ such that

$$\int_{\mathsf{T}\mathbb{R}^d \times \mathsf{T}\mathbb{R}^d} |v' - v''|^2 \, d\boldsymbol{\Theta} \leqslant L^2 W_2^2(\nu, \delta_0) = L^2 m_2^2(\nu).$$

Since $\mathbf{F}[\delta_0] \in \mathcal{P}_c(\mathsf{T}\mathbb{R}^d)$ by assumption, there exists $D > 0$ such that $\mathrm{supp}(v_\sharp \mathbf{F}[\delta_0]) \subset B_D(0)$. Hence, we have

$$\begin{aligned} L^2 m_2^2(\nu) &\geqslant \int_{\mathsf{T}\mathbb{R}^d \times \mathsf{T}\mathbb{R}^d} |v' - v''|^2 \, d\boldsymbol{\Theta} \\ &\geqslant \int_{\mathsf{T}\mathbb{R}^d \times \mathsf{T}\mathbb{R}^d} [|v'| - D]_+^2 \, d\boldsymbol{\Theta} \\ &\geqslant \int_{\mathsf{T}\mathbb{R}^d} |v'|^2 \, d\mathbf{F}[\nu] - 2D \int_{\mathsf{T}\mathbb{R}^d} |v'| \, d\mathbf{F}[\nu], \end{aligned}$$

where $[.]_+$ denotes the positive part. By the trivial estimate $|v'| \leqslant D + \frac{|v'|^2}{4D}$, we conclude

$$|\mathbf{F}[\nu]|_2^2 \leqslant 2\left(2D^2 + L^2 m_2^2(\nu)\right).$$

Hence (9.2.14) and thus the global solvability of the Explicit Euler scheme in $D(\mathbf{F})$ by Proposition 9.2.11. To conclude it is enough to apply Theorem 9.2.13(a) and Theorem 9.1.8. □

It is immediate to notice that the semi-discrete Lagrangian scheme proposed in [31] coincides with the Explicit Euler Scheme given in Definition 9.1.1. In particular, we can state the following comparison between the limit obtained by the Explicit Euler scheme (EE) (leading to the λ-EVI solution of (8.1.1)) and that of the approximating LASs scheme proposed in [94] (leading to a barycentric solution to (8.1.1) in the sense of Definition 8.3.1).

Corollary 9.5.10. *Let* $\mathbf{F}$ *be a PVF satisfying (H1)-(H2),* $\mu_0 \in \mathcal{P}_b(\mathbb{R}^d)$ *and let* $T \in (0, +\infty)$. *Let* $(n_k)_{k\in\mathbb{N}}$ *be a sequence such that the LASs scheme* $(\mu^{n_k})_{k\in\mathbb{N}}$ *of [94, Definition 3.1] converges uniformly-in-time and let* $(M_{\tau_k})_{k\in\mathbb{N}}$ *be the affine interpolants of the Explicit Euler Scheme defined in* (9.1.2)*, with* $\tau_k = \frac{T}{n_k}$*. Then* $(\mu^{n_k})_{k\in\mathbb{N}}$ *and* $(M_{\tau_k})_{k\in\mathbb{N}}$ *converge to the same limit curve* $\mu : [0,T] \to \mathcal{P}_b(\mathbb{R}^d)$*, which is the unique* λ*-EVI solution of* (8.1.1) *in* $[0,T]$.

Proof. By Proposition 9.5.9, $\mathbf{F}$ is a $\left(\frac{L^2+1}{2}\right)$-dissipative MPVF according to (7.5.1) s.t. $M(\mu_0, \tau, T, \tilde{L}) \neq \emptyset$ for every $\tau > 0$, where $\tilde{L} > 0$ is a suitable constant depending on μ_0 and $\mathbf{F}$. Thus by Theorem 9.1.8, $(M_{\tau_k})_{k\in\mathbb{N}}$ uniformly converges to a λ-EVI solution $\mu : [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ which is unique since $\mathbf{F}$ generates a $\left(\frac{L^2+1}{2}\right)$-flow. Since we start from a compactly supported μ_0, the semi-discrete Lagrangian scheme of [31] and our Euler Scheme actually coincide. To conclude we apply [31, Theorem 4.1] obtaining that μ is also the limit of the LASs scheme. □

We conclude that among the possibly not-unique (see [31]) barycentric solutions to (8.1.1) - i.e. the solutions in the sense of [94]/Definition 8.3.1 - we are *selecting* only one (the λ-EVI solution), which turns out to be the one associated with the LASs approximating scheme.

In light of this observation, we revisit an interesting example studied in [94, Section 7.1] and [31, Section 6].

Example 9.5.11 (Splitting particle). For every $\nu \in \mathcal{P}_b(\mathbb{R})$ define:

$$B(\nu) := \sup\left\{x : \nu(]-\infty, x]) \leqslant \frac{1}{2}\right\}, \qquad \eta(\nu) := \nu(]-\infty, B(\nu)]) - \frac{1}{2},$$

so that $\nu(\{B(\nu)\}) = \eta(\nu) + \frac{1}{2} - \nu(]-\infty, B(\nu)[)$. We define the PVF

$$\mathbf{F}[\nu] := \int \mathbf{F}_x[\nu]\, d\nu(x),$$

by

$$\mathbf{F}_x[\nu] := \begin{cases} \delta_{-1} & \text{if } x < B(\nu) \\ \delta_1 & \text{if } x > B(\nu) \\ \frac{\left(\eta\delta_1 + \left(\frac{1}{2} - \nu(]-\infty, B(\nu)[)\right)\delta_{-1}\right)}{\nu(\{B(\nu)\})} & \text{if } x = B(\nu), \nu(\{B(\nu)\}) > 0. \end{cases}$$

By [94, Proposition 7.2], $\mathbf{F}$ satisfies assumptions (H1)-(H2) with $L = 0$ and the LASs scheme admits a unique limit. Moreover, the solution $\mu : [0, T] \to \mathcal{P}_b(\mathbb{R})$ obtained as limit of LASs, is given by

$$\begin{aligned} \mu_t(A) =& \mu_0((A \cap]-\infty, B(\mu_0) - t[) + t) + \mu_0((A \cap]B(\mu_0) + t, +\infty[) - t) \\ &+ \frac{1}{\mu_0(\{B(\mu_0)\})}\left(\eta\delta_{B(\mu_0)+t}(A) + (\frac{1}{2} - \mu_0(]-\infty, B(\mu_0)[))\delta_{B(\mu_0)-t}(A)\right). \end{aligned} \tag{9.5.6}$$

By Corollary 9.5.10, (9.5.6) is the (unique) λ-EVI solution of (8.1.1). In particular:

i) if $\mu_0 = \frac{1}{b-a}\mathcal{L}\llcorner_{[a,b]}$, i.e. the normalized Lebesgue measure restricted to $[a, b]$, we get $\mu_t = \frac{1}{b-a}\mathcal{L}\llcorner_{[a-t, \frac{a+b}{2}-t]} + \frac{1}{b-a}\mathcal{L}\llcorner_{[\frac{a+b}{2}+t, b+t]}$;

ii) if $\mu_0 = \delta_{x_0}$, we get $\mu_t = \frac{1}{2}\delta_{x_0+t} + \frac{1}{2}\delta_{x_0-t}$.

Notice that, in case (i), since $\mu_t \ll \mathcal{L}$ for all $t \in (0, T)$, i.e. $\mu_t \in \mathcal{P}_2^r(\mathbb{R})$, we can also apply Theorem 8.3.7 to conclude that μ is the λ-EVI solution of (8.1.1) with $\mu_0 = \frac{1}{b-a}\mathcal{L}\llcorner_{[a,b]}$. Moreover, take $\varepsilon > 0$, and consider case (i) where we denote by μ_0^ε the initial datum and by μ^ε the corresponding λ-EVI solution to (8.1.1) with $a = x_0 - \varepsilon$, $b = x_0 + \varepsilon$. We can apply (9.2.17) with $\mu_0 = \mu_0^\varepsilon$ and $\mu_1 = \delta_{x_0}$ in order to give another proof that, for all $t \in [0, T]$, the W_2-limit of $S_t[\mu_0^\varepsilon]$ as $\varepsilon \downarrow 0$, that is $S_t[\delta_{x_0}] = \frac{1}{2}\delta_{x_0+t} + \frac{1}{2}\delta_{x_0-t}$, is a λ-EVI solution starting from δ_{x_0}. Thus we end up with (ii).

Dealing with case (ii), we recall that, if $\mu_0 = \delta_{x_0}$ then also the stationary curve $\bar{\mu}_t = \delta_{x_0}$, for all $t \in [0, T]$, satisfies the barycentric property of Definition 8.3.1 (see [31, Example 6.1]), thus it is a solution in the sense of [94]. However, $\bar{\mu}$ is not a λ-EVI solution since it does not coincide with the curve given by (ii). This fact can also be checked by a direct calculation as follows: we find $\nu \in \mathcal{P}_b(\mathbb{R})$ such that

$$\frac{d}{dt}\frac{1}{2}W_2^2(\bar{\mu}_t, \nu) > \lambda W_2^2(\bar{\mu}_t, \nu) - [\mathbf{F}[\nu], \bar{\mu}_t]_r \qquad t \in (0, T), \tag{9.5.7}$$

where $\lambda = \frac{1}{2}$ is the dissipativity constant of the PVF $\mathbf{F}$ coming from the proof of Proposition 9.5.9. Notice that the l.h.s. of (9.5.7) is always zero since $t \mapsto \bar{\mu}_t = \delta_0$ is constant. Take $\nu = \mathcal{L}\llcorner_{[0,1]}$ so that we get $\mathbf{F}[\nu] = \int \mathbf{F}_x[\nu] \, d\nu(x)$, with $\mathbf{F}_x[\nu] = \delta_1$ if $x > \frac{1}{2}$, $\mathbf{F}_x[\nu] = \delta_{-1}$ if $x < \frac{1}{2}$. Noting that $\Lambda(\mathbf{F}[\nu], \delta_0) = \{\mathbf{F}[\nu] \otimes \delta_0\}$, by using the characterization in Theorem 7.1.8 we compute

$$\begin{aligned}
[\mathbf{F}[\nu], \delta_0]_r &= \int_{\mathsf{TH}} \langle x, v \rangle \, d\mathbf{F}[\nu] \\
&= \int_0^{1/2} \langle x, v \rangle \, d\mathbf{F}_x[\nu](v) \, dx + \int_{1/2}^1 \langle x, v \rangle \, d\mathbf{F}_x[\nu](v) \, dx \\
&= \frac{1}{4}.
\end{aligned}$$

Since $W_2^2(\delta_0, \nu) = m_2^2(\nu) = \frac{1}{3}$, we have

$$\lambda W_2^2(\bar{\mu}_t, \nu) - [\mathbf{F}[\nu], \bar{\mu}_t]_r = \frac{1}{6} - \frac{1}{4} < 0,$$

and thus we obtain the desired inequality (9.5.7) with $\nu = \mathcal{L}\llcorner_{[0,1]}$.

Part III

KANTOROVICH-WASSERSTEIN-SOBOLEV SPACES

We prove a general criterium for the density in energy of subalgebras of Lipschitz functions in the metric-Sobolev space $H^{1,p}(X, d, \mathfrak{m})$ associated with a positive Borel measure $\mathfrak{m}$ in a separable and complete metric space (X, d).

We then provide a relevant application to the case of the algebra of cylindrical functions in the space $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,d_{\mathbb{M}}}, \mathfrak{m})$ arising from a positive measure $\mathfrak{m}$ on the Kantorivich-Rubinstein-Wasserstein space $(\mathcal{P}_2(\mathbb{M}), W_{2,d_{\mathbb{M}}})$ of probability measures in a complete Riemannian manifold or a separable Hilbert space $\mathbb{M}$. We will show that such a Sobolev space is always Hilbertian, independently of the choice of the reference measure $\mathfrak{m}$ so that the resulting Cheeger energy is a Dirichlet form.

We will eventually provide an explicit characterization for the corresponding notion of $\mathfrak{m}$-Wasserstein gradient, showing useful calculus rules and its consistency with the tangent bundle and the Γ-calculus inherited from the Dirichlet form.

10

DENSITY OF SUBALGEBRAS OF LIPSCHITZ FUNCTIONS IN METRIC SOBOLEV SPACES

In this Chapter, we treat the case of general metric Sobolev spaces and we show that sufficiently rich subalgebras of Lipschitz functions characterize the space. In Section 10.1 we recap the construction of metric Sobolev spaces with the relaxation approach of Cheeger; in Section 10.2 we present our main density result, while in Section 10.3 we briefly treat the case of intrinsic distances.

This Chapter is the result of a collaboration with Massimo Fornasier and Giuseppe Savaré.

10.1 SOBOLEV FUNCTIONS AND MINIMAL RELAXED GRADIENTS

In this section we will briefly recap the construction of metric Sobolev spaces adapting the relaxation viewpoint of the Cheeger energy to the presence of a distinguished algebra of Lipschitz functions [6, 7, 108]. Let (X, d) be a complete and separable metric space. We will denote by $\mathrm{Lip}_b(X, d)$ the space of bounded and Lipschitz real functions $f : X \to \mathbb{R}$. The asymptotic Lipschitz constant of $f \in \mathrm{Lip}_b(X, d)$ is defined as

$$\mathrm{lip}_d f(x) := \lim_{r \downarrow 0} \mathrm{Lip}(f, B(x, r), d) = \limsup_{y,z \to x,\ y \neq z} \frac{|f(y) - f(z)|}{d(y, z)}, \tag{10.1.1}$$

where $B(x, r)$ denotes the open ball centered at x with radious r and, for $A \subset X$, the quantity $\mathrm{Lip}(f, A, d)$ is defined as

$$\mathrm{Lip}(f, A, d) := \sup_{x,y \in A,\, x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}. \tag{10.1.2}$$

We will simply write $\mathrm{Lip}_b(X), \mathrm{lip} f, \mathrm{Lip}(f, A)$, omitting to explicitly mention d, when the choice of the metric d is clear from the context.

We will also deal with a unital algebra $\mathscr{A} \subset \mathrm{Lip}_b(X)$ separating the points of X, i.e.

$$1 \in \mathscr{A}, \quad \text{for every } x_0, x_1 \in X \text{ there exists } f \in \mathscr{A}\colon \quad f(x_0) \neq f(x_1). \tag{10.1.3}$$

The initial Haudorff topology $\tau_{\mathscr{A}}$ induced on X by $\mathscr{A}$ is clearly coarser than the metric topology of X. Let $\mathfrak{m}$ be a finite and positive Borel measure on X (being X a Polish space, $\mathfrak{m}$ is also a Radon measure). We will denote by $\mathcal{L}^0(X, \mathfrak{m})$ the set of $\mathfrak{m}$-measurable real functions defined in X; $L^0(X, \mathfrak{m})$ is the usual quotient of $\mathcal{L}^0(X, \mathfrak{m})$ obtained by identifying two functions which coincide $\mathfrak{m}$-a.e. in X. In a similar way, $\mathcal{L}^p(X, \mathfrak{m})$ and $L^p(X, \mathfrak{m})$ are the usual Lebesgue spaces of p-summable $\mathfrak{m}$-measurable (equivalence classes of) real functions, $p \in [1, +\infty]$. It is worth noticing that by [108, Lemma 2.1.27] we have that

$$\mathscr{A} \text{ is dense in } L^p(X, \mathfrak{m}) \text{ for every } p \in [1, \infty). \tag{10.1.4}$$

We will endow $L^0(X, \mathfrak{m})$ with the topology of the convergence in measure, which is induced by the metric

$$d_{L^0}(f_1, f_2) := \int_X \vartheta(|f_1 - f_2|)\, d\mathfrak{m} \tag{10.1.5}$$

where $\vartheta : [0, +\infty) \to [0, +\infty)$ is any increasing, concave, bounded function with $\vartheta(0) = \lim_{r\downarrow 0} \vartheta(r) = 0$.

In the following we fix an exponent $p \in (1, +\infty)$.

Definition 10.1.1 ($(p, \mathscr{A})$-relaxed gradients)**.** We say that $G \in L^p(X, \mathfrak{m})$ is a $(p, \mathscr{A})$-relaxed gradient of a $\mathfrak{m}$-measurable function $f \in L^0(X, \mathfrak{m})$ if there exists a sequence $(f_n)_{n\in\mathbb{N}} \in \mathscr{A}$ such that:

1. $f_n \to f$ in $\mathfrak{m}$-measure and $\mathrm{lip} f_n \to \tilde{G}$ weakly in $L^p(X, \mathfrak{m})$;
2. $\tilde{G} \leqslant G$ $\mathfrak{m}$-a.e. in X.

The minimal $(p, \mathscr{A})$-relaxed gradient of f (denoted by $|Df|_{\star,\mathscr{A}}$) is the element of minimal L^p-norm among all the $(p, \mathscr{A})$-relaxed gradients of f. We will just write $|Df|_\star$ if $\mathscr{A} = \mathrm{Lip}_b(X)$.

We collect in the following Theorem the main properties of $|Df|_{\star,\mathscr{A}}$ we will extensively use.

Theorem 10.1.2.

(1) The set

$$S := \Big\{(f, G) \in L^0(X, \mathfrak{m}) \times L^p(X, \mathfrak{m}) : G \text{ is a } (p, \mathscr{A})\text{-relaxed gradient of } f\Big\}$$

is convex and it is closed with respect to to the product topology of the convergence in $\mathfrak{m}$-measure and the weak convergence in $L^p(X, \mathfrak{m})$. In particular, the restriction $S_q := S \cap L^q(X, \mathfrak{m}) \times L^p(X, \mathfrak{m})$ is weakly closed in $L^q(X, \mathfrak{m}) \times L^p(X, \mathfrak{m})$ for every $q \in (1, +\infty)$.

(2) (Strong approximation) If $f \in L^0(X, \mathfrak{m})$ has a $(p, \mathscr{A})$ relaxed gradient then $|Df|_{\star,\mathscr{A}}$ is well defined. If f takes values in a closed (possibly unbounded) interval $I \subset \mathbb{R}$ then there exists a sequence $f_n \in \mathscr{A}$ with values in I such that

$$f_n \to f \ \mathfrak{m}\text{-a.e. in } X, \quad \mathrm{lip} f_n \to |Df|_{\star,\mathscr{A}} \text{ strongly in } L^p(X, \mathfrak{m}). \tag{10.1.6}$$

If moreover $f \in L^q(X, \mathfrak{m})$ for some $q \in [1, +\infty)$ then we can also find a sequence as in (10.1.6) converging strongly to f in $L^q(X, \mathfrak{m})$.

(3) (Pointwise minimality) If G is a $(p, \mathscr{A})$-relaxed gradient of $f \in L^0(X, \mathfrak{m})$ then $|Df|_{\star,\mathscr{A}} \leqslant G$ $\mathfrak{m}$-a.e. in X.

(4) (Leibniz rule) If $f, g \in L^\infty(X, \mathfrak{m})$ have $(p, \mathscr{A})$-relaxed gradient, then $h := fg$ has $(p, \mathscr{A})$-relaxed gradient and

$$|D(fg)|_{\star,\mathscr{A}} \leqslant |f|\,|Dg|_{\star,\mathscr{A}} + |g|\,|Df|_{\star,\mathscr{A}} \quad \mathfrak{m}\text{-a.e. in } X. \tag{10.1.7}$$

(5) *(Sublinearity) If* $f, g \in L^0(X, \mathfrak{m})$ *have* $(p, \mathscr{A})$*-relaxed gradient then*

$$|D(\alpha f + \beta g)|_{\star,\mathscr{A}} \leqslant |\alpha|\, |Df|_{\star,\mathscr{A}} + |\beta|\, |Dg|_{\star,\mathscr{A}} \quad \mathfrak{m}\text{-a.e. in } X. \tag{10.1.8}$$

(6) *(Locality) If* $f \in L^0(X, \mathfrak{m})$ *has a* $(p, \mathscr{A})$ *relaxed gradient, then for any* $\mathscr{L}^1$*-negligible Borel subset* $N \subset \mathbb{R}$ *we have*

$$|Df|_{\star,\mathscr{A}} = 0 \quad \mathfrak{m}\text{-a.e. on } f^{-1}(N). \tag{10.1.9}$$

(7) *(Chain rule) If* $f \in L^0(X, \mathfrak{m})$ *has a* $(p, \mathscr{A})$ *relaxed gradient and* $\phi \in \mathrm{Lip}(\mathbb{R})$ *then* $\phi \circ f$ *has* $(p, \mathscr{A})$*-relaxed gradient and*

$$|D(\phi \circ f)|_{\star,\mathscr{A}} \leqslant |\phi'(f)|\, |Df|_{\star,\mathscr{A}} \quad \mathfrak{m}\text{-a.e. in } X, \tag{10.1.10}$$

and equality holds in (10.1.10) *if* ϕ *is monotone or* C^1.

(8) *(Truncations) If* $f_j \in L^0(X, \mathfrak{m})$ *has* $(p, \mathscr{A})$ *relaxed gradient,* $j = 1, \cdots, J$*, then also the functions* $f_+ := \max(f_1, \cdots, f_J)$ *and* $f_- := \min(f_1, \cdots, f_J)$ *have* $(p, \mathscr{A})$ *relaxed gradient and*

$$|Df_+|_{\star,\mathscr{A}} = |Df_j|_{\star,\mathscr{A}} \quad \mathfrak{m}\text{-a.e. on } \{x \in X : f_+ = f_j\}, \tag{10.1.11}$$

$$|Df_-|_{\star,\mathscr{A}} = |Df_j|_{\star,\mathscr{A}} \quad \mathfrak{m}\text{-a.e. on } \{x \in X : f_- = f_j\}. \tag{10.1.12}$$

Proof. We give a few references for the proofs. The case when $p = 2$, $\mathscr{A} = \mathrm{Lip}_b(X)$ and the local slope of f is used to define relaxed gradients have been considered in [7, Sec. 4], whose proof generalizes easily to the case $p \in (1, \infty)$ and the asymptotic Lipschitz constant (10.1.1), see also [6].

The definition and the properties involving a general unital subalgebra $\mathscr{A}$ have been discussed in [108, Sec. 3]: points (1,2) correspond to Lemma 3.1.6 and Corollary 3.1.9, (3) has been stated in Lemma 3.1.11, (4) refers to Corollary 3.1.10, (5,6,7,8) are proved in Theorem 3.1.12 and its Corollary 3.1.13.

Let us make three further technical comments:

- both [7, 108] involve an auxiliary topology τ: in the present case, being X complete and separable and d a canonical metric (thus d only take finite values), we can select τ as the (Polish) topology induced by d.
- In order to deal with *extended* distances, in [108] has also been assumed that the unital algebra $\mathscr{A}$ satisfies the *stronger* compatibility condition

$$d(x, y) = \sup\Big\{ f(x) - f(y) : f \in \mathscr{A},\ \mathrm{Lip}(f, X) \leqslant 1 \Big\}, \tag{10.1.13}$$

  which clearly implies that $\mathscr{A}$ separates the points of X as in (10.1.3). However, such a property is not needed in the construction and the proofs of Section 3.1.1 of [108]. The only point where (10.1.13) explicitely occurs is in the proof of Locality [108, Lemma 3.1.11], to ensure that the restriction of $\mathscr{A}$ to each compact set $K \subset X$ is uniformly dense in $C(K)$, a property which is guaranteed by (10.1.3) thanks to Stone-Weierstrass Theorem.

- The standard approach of [7, 108] considers first functions f belonging to $L^p(X, \mathfrak{m})$ instead of general $\mathfrak{m}$-measurable functions. However, the compatibility with truncations showing that for every $k > 0$

$$|\mathrm{D}\, T_k(f)|_{\star,\mathscr{A}}(x) = \begin{cases} |\mathrm{D}f|_{\star,\mathscr{A}}(x) & \text{if } |f(x)| < k, \\ 0 & \text{if } |f(x)| \geqslant k, \end{cases} \qquad T_k(f) := -k \vee f \wedge k, \tag{10.1.14}$$

and the possibility to find strong approximations of $T_k(f)$ satisfying (10.1.6) and taking values in $[-k, k]$ (see [108, Cor. 3.1.9]) allow for a standard extension of the theory from $L^p(X, \mathfrak{m})$ to $L^0(X, \mathfrak{m})$, see also the discussion related to (4.16) of [7]. □

Starting from Definition 10.1.1 and using the properties of Theorem 10.1.2 it is natural to introduce the following notions.

Definition 10.1.3 (Cheeger energy and Sobolev space)**.** We call $D^{1,p}(X, d, \mathfrak{m}; \mathscr{A})$ the set of functions in $L^0(X, \mathfrak{m})$ with a $(p, \mathscr{A})$-relaxed gradient and we set

$$\mathsf{CE}_{p,\mathscr{A}}(f) := \int_X |\mathrm{D}f|^p_{\star,\mathscr{A}}(x)\, \mathrm{d}\mathfrak{m}(x) \quad \text{for every } f \in D^{1,p}(X, d, \mathfrak{m}; \mathscr{A}), \tag{10.1.15}$$

with $\mathsf{CE}_{p,\mathscr{A}}(f) := +\infty$ if $f \notin D^{1,p}(X, d, \mathfrak{m}; \mathscr{A})$. The Sobolev space $H^{1,p}(X, d, \mathfrak{m}; \mathscr{A})$ is defined as $L^p(X, \mathfrak{m}) \cap D^{1,p}(X, d, \mathfrak{m}; \mathscr{A})$ and it is a Banach space with the norm $\|f\|^p_{H^{1,p}(X,d,\mathfrak{m};\mathscr{A})} := \|f\|^p_{L^p} + \mathsf{CE}_{p,\mathscr{A}}(f)$. As usual, we will write $D^{1,p}(X, d, \mathfrak{m})$, $\mathsf{CE}_p(f)$, $H^{1,p}(X, d, \mathfrak{m})$ and $\|f\|_{H^{1,p}}$ when $\mathscr{A} = \mathrm{Lip}_b(X)$.

Remark 10.1.4 (Cheeger energy as relaxation of the pre-Cheeger energy)*.* We can equivalently define the Cheeger energy $\mathsf{CE}_{p,\mathscr{A}}$ as the L^0-lower semicontinuous relaxation of the restriction to $\mathscr{A}$ of the *pre-Cheeger energy* pCE_p, the latter being defined as

$$\mathsf{pCE}_p(f) := \int_X (\mathrm{lip} f)^p\, \mathrm{d}\mathfrak{m}, \quad f \in \mathrm{Lip}_b(X). \tag{10.1.16}$$

In other words, for every $f \in L^0(X, \mathfrak{m})$ it holds ([108, Corollary 3.1.7])

$$\mathsf{CE}_{p,\mathscr{A}}(f) = \inf \left\{ \liminf_{n\to+\infty} \mathsf{pCE}_p(f_n) : f_n \in \mathscr{A},\ f_n \to f \text{ in } L^0(X, \mathfrak{m}) \right\}. \tag{10.1.17}$$

In particular the functional $\mathsf{CE}_{p,\mathscr{A}}$ is lower semicontinuous in $L^0(X, \mathfrak{m})$. Here the choice of the L^0-topology does not play a crucial role, since, by Theorem 10.1.2(2), the restriction of $\mathsf{CE}_{p,\mathscr{A}}$ to $L^q(X, \mathfrak{m})$, $q \in [1, \infty)$, can be equivalently obtained as L^q-relaxation:

$$\mathsf{CE}_{p,\mathscr{A}}(f) = \inf \left\{ \liminf_{n\to+\infty} \mathsf{pCE}_p(f_n) : f_n \in \mathscr{A},\ f_n \to f \text{ in } L^q(X, \mathfrak{m}) \right\}, \quad f \in L^q(X, \mathfrak{m}). \tag{10.1.18}$$

It is clear that we have the obvious implication for $f \in L^0(X, \mathfrak{m})$:

$$f \text{ has a } (p, \mathscr{A})\text{-relaxed gradient} \quad \Rightarrow \quad \begin{cases} f \text{ has a } (p, \mathrm{Lip}_b(X))\text{-relaxed gradient and} \\ |Df|_\star \leqslant |Df|_{\star,\mathscr{A}} \ \mathfrak{m}\text{-a.e. in } X. \end{cases} \tag{10.1.19}$$

The converse implication together with the identity $|Df|_\star = |Df|_{\star,\mathscr{A}}$ is an important density property for an algebra $\mathscr{A}$: by Theorem 10.1.2(2), it is equivalent to the following property.

Definition 10.1.5 (Density in energy of a subalgebra of Lipschitz functions)**.** We say that a subalgebra $\mathscr{A} \subset \mathrm{Lip}_b(X)$ is *dense in* p*-energy* if for every $f \in L^0(X, \mathfrak{m})$ with a p-relaxed gradient there exists a sequence $(f_n)_{n\in\mathbb{N}}$ satisfying

$$f_n \in \mathscr{A}, \quad f_n \to f \ \mathfrak{m}\text{-a.e. in } X, \quad \mathrm{lip} f_n \to |Df|_\star \text{ strongly in } L^p(X, \mathfrak{m}). \tag{10.1.20}$$

When $\mathscr{A}$ is unital and separating, this is equivalent to the fact that f has a $(p, \mathscr{A})$-relaxed gradient and

$$|Df|_{\star,\mathscr{A}} = |Df|_\star \quad \mathfrak{m}\text{-a.e. in } X. \tag{10.1.21}$$

In particular $D^{1,p}(X, d, \mathfrak{m}; \mathscr{A}) = D^{1,p}(X, d, \mathfrak{m})$.

Remark 10.1.6. As we already mentioned in Remark 10.1.4, the choice of arbitrary measurable maps $f \in L^0(X, \mathfrak{m})$ in Definition 10.1.5 and of the pointwise $\mathfrak{m}$-a.e. convergence in (10.1.20) is not restrictive: a simple truncation argument (which can be implemented by using odd polynomials, see [108, Corollary 2.1.24]) shows that $\mathscr{A}$ is dense in p-energy if and only if for every $f \in L^p(X, \mathfrak{m})$ with a p-relaxed gradient there exists a sequence $(f_n)_{n\in\mathbb{N}}$ satisfying

$$f_n \in \mathscr{A}, \quad f_n \to f \text{ in } L^p(X, \mathfrak{m}), \quad \mathrm{lip} f_n \to |Df|_\star \text{ strongly in } L^p(X, \mathfrak{m}). \tag{10.1.22}$$

If $\mathscr{A}$ is unital and separating this is equivalent to $H^{1,p}(X, d, \mathfrak{m}; \mathscr{A}) = H^{1,p}(X, d, \mathfrak{m})$ with equal norms.

A first sufficient condition, in the more general framework of extended topological metric measure spaces, is provided by the compatibility condition (10.1.13) [108, Theorems 3.2.7, 5.3.1].

In the present Polish setting, we notice that (10.1.20) (and, a fortiori, (10.1.13)) implies the weaker condition

$$\text{for every } y \in X \text{ the function } d_y : x \mapsto d(x, y) \text{ has } (p, \mathscr{A})\text{-relaxed gradient } 1 \tag{10.1.23}$$

which is equivalent, thanks to Theorem 10.1.2(2), to

$$|Dd_y|_{\star,\mathscr{A}} \leqslant 1 \quad \mathfrak{m}\text{-a.e. in } X. \tag{10.1.24}$$

In fact, using the truncations (10.1.14), each function d_y can be approximated by the increasing sequence $f_k := T_k d_y$ of 1-Lipschitz maps, so that

$$|Dd_y|_\star \leqslant 1 \quad \mathfrak{m}\text{-a.e. in } X \text{ for every } y \in X, \tag{10.1.25}$$

and therefore (10.1.20) yields (10.1.24).

Remark 10.1.7 (The effect of truncations). The $(p, \mathscr{A})$-relaxed gradient is not affected by truncations of the distance functions, in particular it is not restrictive to assume d bounded above by a constant, e.g. 1. In fact, if we introduce a parameter $a > 0$ and the truncated distance

$$d_a(x_1, x_2) := d(x_1, x_2) \wedge a \quad \text{for every } x_1, x_2 \in X, \tag{10.1.26}$$

(X, d_a) is still a complete and separable metric space, the sets $\mathrm{Lip}_b(X, d)$ and $\mathrm{Lip}_b(X, d_a)$ coincide, and it is easy to check that

$$\mathrm{lip}_d f = \mathrm{lip}_{d_a} f \quad \text{for every bounded and Lipschitz function } f. \tag{10.1.27}$$

We deduce that d and d_a induce the same $(p, \mathscr{A})$-relaxed gradient. Notice moreover that using (10.1.26) we can also easily cover the case of extended distances (i.e. possibly assuming the value $+\infty$), *provided* (X, d_a) *is a separable metric space*. The case when (X, d_a) is not separable requires a more refined setting involving an auxiliary topology τ [108].

Remark 10.1.8. Using the above truncation argument, one can easily see that if d, d' are two metrics satisfying

$$d(x_1, x_2) = d'(x_1, x_2) \quad \text{whenever } d(x_1, x_2) \wedge d'(x_1, x_2) < a \tag{10.1.28}$$

then the sets $\mathrm{Lip}_b(X, d)$ and $\mathrm{Lip}_b(X, d')$ coincide,

$$\mathrm{lip}_d f = \mathrm{lip}_{d'} f \quad \text{for every bounded and Lipschitz function } f. \tag{10.1.29}$$

and d and d' induce the same $(p, \mathscr{A})$-relaxed gradient.

It is possible to express (10.1.24) in a more flexible way, by using suitable nonlinear functions of d_y. We state a general result.

Lemma 10.1.9. *Let* $I = (a, b)$ *be an interval (possibly unbounded) of* $\mathbb{R}$ *and let* $\zeta : \mathbb{R} \to \mathbb{R}$ *be a Lipschitz and nondecreasing map satisfying*

$$\textit{the restriction of } \zeta \textit{ to } I \textit{ is of class } C^1 \textit{ with } \zeta'(s) > 0 \textit{ if } s \in I. \tag{10.1.30}$$

If $f : X \to \bar{I}$ *is a Borel function, then the condition*

$$f \in D^{1,2}(X, \mathscr{A}), \quad |Df|_{\star,\mathscr{A}} \leqslant 1 \tag{10.1.31}$$

is equivalent to

$$\zeta \circ f \in D^{1,2}(X, \mathscr{A}), \quad \big|D(\zeta \circ f)\big|_{\star,\mathscr{A}}(x) \leqslant \zeta'(f(x)) \quad \textit{for } \mathfrak{m}\textit{-a.e. } x \in X. \tag{10.1.32}$$

Proof. It is clear that if $|Df|_{\star,\mathscr{A}} \leqslant 1$ then (10.1.32) holds, thanks to (10.1.10). In order to prove the converse implication, we consider a strictly decreasing sequence $a_n \downarrow a$, a strictly increasing sequence $b_n \uparrow b$ and nondecreasing and bounded Lipschitz functions $\psi_n : \mathbb{R} \to \mathbb{R}$ such that

$$\psi_n(z) = a_n \text{ if } z < \zeta(a_n), \quad \psi_n(\zeta(s)) = s \text{ for every } s \in [a_n, b_n],$$
$$\psi_n(z) = b_n \text{ if } z > \zeta(b_n).$$

The restriction of ψ_n to the interval $[\zeta(a_n), \zeta(b_n)]$ is of class C^1.

Setting $h(x) := \zeta(f(x))$, the Chain rule (10.1.10) yields

$$|D(\psi_n \circ h)|_{\star,\mathscr{A}}(x) \leqslant (\psi_n' \circ h)\,|Dh|_{\star,\mathscr{A}}(x) \leqslant (\psi_n' \circ \zeta(f(x)))\zeta'(f(x)).$$

Since $\psi_n(h(x)) = a_n \vee f(x) \wedge b_n$ the locality property (10.1.9), the truncation property 10.1.2(8), and the fact that $\psi'(\zeta(s))\zeta'(s) = 1$ if $s \in [a_n, b_n]$ yield

$$|D(\psi_n \circ h)|_{\star,\mathscr{A}} \leqslant 1 \quad \mathfrak{m}\text{-a.e.} \tag{10.1.33}$$

Since $\psi_n \circ h \to f$ pointwise in X as $n \to \infty$, passing to the limit in (10.1.33) we get $|Df|_{\star,\mathscr{A}} \leqslant 1$. □

Remark 10.1.10. Thanks to Lemma 10.1.9, if d is a bounded metric and $q > 1$, (10.1.24) is equivalent to

$$|Dd_y^q|_{\star,\mathscr{A}}(x) \leqslant q\, d_y^{q-1}(x) \quad \text{for } \mathfrak{m}\text{-a.e. } x \in X. \tag{10.1.34}$$

In particular, if (10.1.34) holds for some $q \geqslant 1$, it holds for any $q \geqslant 1$.

10.2 A DENSITY RESULT

We have seen that in the present setting of Polish spaces, condition (10.1.24) (or, equivalently, (10.1.32) for some admissible truncation satisfying (10.1.30)) is a necessary condition for the validity of the approximation property (10.1.20) and of the identification $|Df|_\star = |Df|_{\star,\mathscr{A}}$. We want to show that (10.1.24) or (10.1.32) are also *sufficient* conditions.

Theorem 10.2.1. *Let $(X, d, \mathfrak{m})$ be a Polish metric measure space, let $Y \subset X$ be a dense subset, and let $\mathscr{A}$ be a unital separating subalgebra of $\mathrm{Lip}_b(X)$ as in (10.1.3). If*

$$\textit{for every } y \in Y \textit{ it holds} \quad d_y \in D^{1,2}(X, \mathscr{A}), \quad \big|Dd_y\big|_{\star,\mathscr{A}} \leqslant 1 \tag{10.2.1}$$

then $\mathscr{A}$ is dense in p-energy according to Definition 10.1.5.

Proof. We split the proof in various steps. Notice that by (10.1.19) it is sufficient to prove that

$$|Df|_{\star,\mathscr{A}} \leqslant |Df|_\star \quad \mathfrak{m}\text{-a.e. in } X. \tag{10.2.2}$$

(1) *It is not restrictive to assume d bounded above by 1*: see Remark 10.1.7.

By Remark 10.1.10 we know that (10.1.34) holds for every $y \in Y$ and every $q \geqslant 1$.

(2) *It is sufficient to prove that*

$$\mathsf{CE}_{p,\mathscr{A}}(f) \leqslant \int_X (\mathrm{lip} f)^p\, d\mathfrak{m} = p\mathsf{CE}_p(f) \quad \textit{for every } f \in \mathrm{Lip}_b(X). \tag{10.2.3}$$

In fact, if f has $(p, \mathrm{Lip}_b(X))$-relaxed gradient, by (10.1.6) we can find a sequence $f_n \in \mathrm{Lip}_b(X)$ such that $f_n \to f$ $\mathfrak{m}$-a.e. and $\mathrm{lip} f_n \to |Df|_\star$ strongly in $L^p(X, \mathfrak{m})$ as

$n \to \infty$. By the L^0-lower semicontinuity of the $\mathrm{CE}_{p,\mathscr{A}}$-energy, passing to the limit in (10.2.3) written for f_n we get

$$\mathrm{CE}_{p,\mathscr{A}}(f) = \int_X |Df|^p_{\star,\mathscr{A}}\,dm \leqslant \int_X |Df|^p_\star\,dm = \mathrm{CE}_p(f) < \infty.$$

We deduce that f has a $(p,\mathscr{A})$ relaxed gradient and that (10.1.21) holds, since $|Df|_\star \leqslant |Df|_{\star,\mathscr{A}}$ m-a.e.

(3) For every $f \in \mathrm{Lip}_b(X)$ and $t > 0$ we introduce the Hopf-Lax regularization $Q_t f : X \to \mathbb{R}$ defined by

$$Q_t f(x) := \inf_{y \in X} \frac{1}{2t} d^2(x,y) + f(y), \quad x \in X. \tag{10.2.4}$$

It is clear that $Q_t f$ is bounded (it takes values in the interval $[\inf_X f, \sup_X f]$ and Lipschitz, being the infimum of a family of t^{-1}-Lipschitz functions. We consider the upper semicontinuous function [7, (3.4) and Prop. 3.2]

$$D_t^+ f(x) := \sup_{(y_n)} \limsup_{n\to\infty} d(x,y_n), \tag{10.2.5}$$

where the $(y_n)_n$'s vary among all the minimizing sequences of (10.2.4). $D_t^+ f$ is also uniformly bounded and satisfies (see e.g. [108, Lemma 3.2.1])

$$\frac{D_t f^+(x)}{t} \leqslant 2\mathrm{Lip}(f,X). \tag{10.2.6}$$

In fact, if y_n is a minimizing sequence of (10.2.4), for every $\varepsilon > 0$ we eventually have

$$\frac{1}{2t} d^2(x,y_n) + f(y_n) \leqslant Q_t f(x) + \varepsilon \leqslant f(x) + \varepsilon$$

i.e., setting $L := \mathrm{Lip}(f,X)$,

$$\frac{1}{2t} d^2(x,y_n) \leqslant \varepsilon + f(x) - f(y_n) \leqslant \varepsilon + L d(x,y_n) \leqslant \varepsilon + \frac{1}{4t} d(x,y_n)^2 + tL^2.$$

We thus get

$$\limsup_{n\to\infty} \frac{1}{4t} d^2(x,y_n) \leqslant \varepsilon + tL^2$$

which yields (10.2.6) since $\varepsilon > 0$ is arbitrary.

(4) *For every* $f \in \mathrm{Lip}_b(X)$ *and for every* $t > 0$

$$|DQ_t f|_{\star,\mathscr{A}}(x) \leqslant t^{-1} D_t^+ f(x) \quad \textit{for } m\textit{-a.e. } x \in X. \tag{10.2.7}$$

Let $Y' = \{y_n\}_{n\in\mathbb{N}}$ be a countable set dense in Y; since $f \in \mathrm{Lip}_b(X)$ it is easy to check that

$$\begin{aligned} Q_t f(x) &= \inf_{y\in Y} \frac{1}{2t} d^2(x,y) + f(y) = \lim_{n\to\infty} Q_t^n f(x), \\ Q_t^n f(x) &:= \min_{1\leqslant k\leqslant n} \frac{1}{2t} d^2(x,y_k) + f(y_k). \end{aligned} \tag{10.2.8}$$

We can now use (10.1.34) with $q = 2$ and, also using Theorem 10.1.2 (8), we obtain that the upper semi-continuous (see[108, Lemma 3.2.2 (b)]) function

$$D^n_t(x) := \frac{1}{t}\max\Big\{d(x,y_k) : 1 \leqslant k \leqslant n,\ Q^n_t(x) = \frac{1}{2t}d^2(x,y_k) + f(y_k)\Big\}, \quad (10.2.9)$$

is a $(p,\mathscr{A})$-relaxed gradient of $Q^n_t f$. It is then clear that for every x there exists a sequence $n \mapsto y'(n;x)$ with $y'(n;x) \in \{y_1, \cdots, y_n\}$ such that $D^n_t(x) = \frac{1}{t}d(x,y'(n;x))$ and $Q^n_t f(x) = \frac{1}{2t}d^2(x,y'(n;x)) + f(y'(n;x)) \to Q_t f(x)$ as $n \to \infty$, i.e. $y'(n;x)$ is a minimizing sequence of (10.2.4). We deduce that

$$\limsup_{n\to\infty} D^n_t(x) = \limsup_{n\to\infty} d(x,y'(n;x)) \leqslant D^+_t f(x) \text{ for every } x \in X. \quad (10.2.10)$$

Since D^n_t are uniformly bounded, up to extracting a suitable subsequence we can suppose that $t^{-1}D^n_t \rightharpoonup G$ in $L^p(X,m)$, G is a $(p,\mathscr{A})$-relaxed gradient of $Q_t f$, and $G \leqslant t^{-1}D^+_t$ thanks to Fatou's Lemma.

(5) *For every $x \in X$, $t > 0$, and $f \in \mathrm{Lip}_b(X)$ we have*

$$\frac{f(x) - Q_t f(x)}{t} = \frac{1}{2}\int_0^1 \Big(\frac{D^+_{rt} f(x)}{rt}\Big)^2 dr, \quad (10.2.11)$$

$$\limsup_{t\downarrow 0} \frac{f(x) - Q_t f(x)}{t} \leqslant \frac{1}{2}\big(\mathrm{lip} f(x)\big)^2. \quad (10.2.12)$$

This follows by [108, Thm. 3.2.4] (see also [5, Thm. 3.1.4, Lemma 3.1.5]).

(5) *Conclusion.* We argue as [108, Theorem 3.2.7]: (10.2.11) and (10.2.6) yield the uniform bound

$$\frac{f(x) - Q_t f(x)}{t} \leqslant 2\big(\mathrm{Lip}(f,X)\big)^2 \quad \text{for every } x \in X,\ t > 0. \quad (10.2.13)$$

Integrating (10.2.12) in X and applying Fatou's Lemma we get

$$\limsup_{t\downarrow 0} \int_X \frac{f(x) - Q_t f(x)}{t}\, dm(x) \leqslant \frac{1}{2}\int_X \big(\mathrm{lip} f(x)\big)^2 dm(x). \quad (10.2.14)$$

On the other hand, (10.2.11) and Fubini's Theorem yield

$$\int_X \frac{f(x) - Q_t f(x)}{t}\, dm(x) = \frac{1}{2}\int_0^1\int_X \Big(\frac{D^+_{rt} f(x)}{rt}\Big)^2 dm(x)\, dr. \quad (10.2.15)$$

A further application of Fatou's Lemma yields

$$\liminf_{t\downarrow 0} \int_X \frac{f(x) - Q_t f(x)}{t}\, dm(x) \geqslant \frac{1}{2}\liminf_{t\downarrow 0}\int_X \Big(\frac{D^+_t f(x)}{t}\Big)^2 dm(x). \quad (10.2.16)$$

Using the fact that $t^{-1}D^+_t f$ is uniformly bounded by (10.2.6), we can find a decreasing and vanishing sequence $n \mapsto t(n)$ and a limit function $G \in L^\infty(X,m)$ such that

$$t(n)^{-1}D^+_{t(n)} f \rightharpoonup^* G \quad \text{weakly}^* \text{ in } L^\infty(X,m) \quad \text{as } n \to \infty,$$

$$\lim_{n\to\infty}\int_X \Big(\frac{D^+_{t(n)} f(x)}{t(n)}\Big)^2 dm(x) = \liminf_{t\downarrow 0}\int_X \Big(\frac{D^+_t f(x)}{t}\Big)^2 dm(x). \quad (10.2.17)$$

Since $t^{-1}D_t^+ f$ it is a $(p,\mathscr{A})$-relaxed gradient of $Q_t f$ by claim (4) and $Q_t f \to f$ pointwise everywhere, using Theorem 10.1.2(1) we get that G is a $(p,\mathscr{A})$-relaxed gradient of f.
Using the lower semicontinuity of the L^2-norm w.r.t. the weak* $L^\infty(X,\mathfrak{m})$ convergence, we get that

$$\lim_{n\to\infty}\int_X \Big(\frac{D^+_{t(n)}f(x)}{t(n)}\Big)^2 \,d\mathfrak{m}(x) \geqslant \int_X G^2\,d\mathfrak{m}(x) \geqslant \int_X |Df|^2_{\star,\mathscr{A}}(x)\,d\mathfrak{m}(x), \quad (10.2.18)$$

where we also used the pointwise minimality of $|Df|^2_{\star,\mathscr{A}}(x)$ given by Theorem 10.1.2(3). Combining (10.2.18), (10.2.17), (10.2.16) and (10.2.14) we deduce that

$$\int_X |Df|^2_{\star,\mathscr{A}}(x)\,d\mathfrak{m}(x) \leqslant \int_X \big(\mathrm{lip} f(x)\big)^2\,d\mathfrak{m}(x)$$

so that (10.2.3) holds. □

Corollary 10.2.2 (Density in energy of $\mathscr{A}$). *If $\mathscr{A}$ is a separating unital subalgebra of* $\mathrm{Lip}_b(X)$ *satisfying* (10.2.1) *then*

$$\mathrm{CE}_{p,\mathscr{A}}(f) = \mathrm{CE}_p(f) \quad \textit{for every } \mathfrak{m}\textit{-measurable function } f : X \to \mathbb{R}. \quad (10.2.19)$$

In particular, $H^{1,p}(X,d,\mathfrak{m}) = H^{1,p}(X,d,\mathfrak{m};\mathscr{A})$.

As we have already said, (10.2.19) can be interpreted as a density result in $H^{1,p}(X,d,\mathfrak{m})$: for every $f \in H^{1,p}(X,d,\mathfrak{m})$ there exists a sequence $f_n \in \mathscr{A}$, $n \in \mathbb{N}$, such that

$$f_n \to f,\ \mathrm{lip} f_n \to |Df|_* \quad \text{strongly in } L^p(X,\mathfrak{m}). \quad (10.2.20)$$

The next result shows that we can always consider the algebra generated by (suitable compositions/truncations of) the distance functions.

Corollary 10.2.3. *Let* Y *be a dense subset of* X *and let* $\zeta : [0,+\infty) \to [0,+\infty)$ *be a nondecreasing and bounded Lipschitz function for which there exists a constant* $a > 0$ *such that the restriction of* ζ *to the interval* $[0,a]$ *is of class* C^1 *with* $\zeta'(s) > 0$ *for every* $s \in [0,a]$.
Then the unital algebra $\mathscr{A}$ *generated by the functions* $x \mapsto \zeta(d(x,y))$ *satisfies* (10.1.21).

Proof. It is not difficult to check that $\mathscr{A}$ separates the points of X, so that in order to apply Theorem 10.2.1, it is enough to check that (10.2.1) holds. Recalling Remark 10.1.7, it will be sufficient to prove that

$$|D(d_y \wedge a)|_{\star,\mathscr{A}} \leqslant 1 \quad \mathfrak{m}\text{-a.e., for every } y \in Y. \quad (10.2.21)$$

Let $\psi : [0,+\infty) \to [0,+\infty)$ a nondecreasing and bounded Lipschitz function such that $\psi(\zeta(r)) = r$ for every $r \in [0,a]$. The restriction of ψ to the interval $[0,\zeta(a)]$ is of class C^1.

If $y \in Y$ and $f(x) := \zeta(d_y(x))$ (recall the notation (10.1.23)), a direct computation shows that

$$\mathrm{lip} f(x) \leqslant \zeta'(d_y(x)) \quad \text{for every } x \in B(y,a).$$

Since $f \in \mathscr{A}$, the Chain rule (10.1.10) yields

$$|D(\psi \circ f)|_{\star,\mathscr{A}} \leqslant (\psi' \circ f)\, |Df|_{\star,\mathscr{A}} \leqslant (\psi' \circ f)\, \mathrm{lip} f.$$

Since $\psi(f(x)) = d_y(x)$ in $B(y, a)$ and $\psi(f(x)) \geqslant \psi(f(a)) = a$ if $x \in X \setminus B(y, a)$, the locality property (10.1.9) and the truncation property 10.1.2(8) yield

$$\big|D(d_y \wedge a)\big|_{\star,\mathscr{A}} = \big|D\big((\psi \circ f) \wedge a\big)\big|_{\star,\mathscr{A}} \leqslant \big(\psi' \circ f\big)\, \mathrm{lip} f\, \chi_{B(y,a)}. \tag{10.2.22}$$

On the other hand, since $\psi'(\zeta(r))\zeta'(r) = 1$ if $r \in [0, a)$,

$$\psi'(f(x))\, \mathrm{lip} f(x) \leqslant \psi'(\zeta(d_y(x))\zeta'(d_y(x)) = 1 \quad \text{for every } x \in B(y, a)$$

so that we obtain (10.2.21). □

We conclude this section with a simple application to the case when $p = 2$ and $\mathrm{lip} f$ has good properties for functions of $\mathscr{A}$.

Theorem 10.2.4 (An Hilbertianity condition). *Let* $p = 2$ *and let* $\mathscr{A}$ *be a separating unital subalgebra of* $\mathrm{Lip}_b(X)$ *satisfying* (10.2.1). *If for every* $f, g \in \mathscr{A}$

$$\int_X \Big(|\mathrm{lip}(f+g)|^2 + |\mathrm{lip}(f-g)|^2\Big)\, dm = 2\int_X \Big(|\mathrm{lip} f|^2 + |\mathrm{lip} g|^2\Big)\, dm, \tag{10.2.23}$$

then $H^{1,2}(X, d, m)$ *is an Hilbert space,* CE_2 *is a quadratic form, and* $\mathscr{A}$ *is strongly dense.*

Proof. It is sufficient to prove that the Cheeger energy is a quadratic form in its domain. Thanks to [40, Prop. 11.9] and the 2-homogeneity of CE_2, this property is equivalent to

$$\mathsf{CE}_2(f+g) + \mathsf{CE}_2(f-g) \leqslant 2\mathsf{CE}_2(f) + 2\mathsf{CE}_2(g) \quad \text{for every } f, g \in H^{1,p}(X, d, m). \tag{10.2.24}$$

We can find two sequences $f_n, g_n \in \mathscr{A}$ such that $f_n \to f$, $g_n \to g$ in m-measure as $n \to \infty$ and $\mathrm{lip} f_n \to |Df|_\star$, $\mathrm{lip} g_n \to |Dg|_\star$ in $L^2(X, m)$. Clearly we have $f_n + g_n \to f + g$, $f_n - g_n \to f - g$ in m-measure and (10.2.23) shows that $\mathrm{lip}(f_n + g_n)$ and $\mathrm{lip}(f_n - g_n)$ are uniformly bounded in $L^2(X, m)$. Up to extracting a suitable sequence, it is not restrictive to assume that $\mathrm{lip}(f_n + g_n) \rightharpoonup G_+ \geqslant |D(f+g)|_\star$ and $\mathrm{lip}(f_n - g_n) \rightharpoonup G_- \geqslant |D(f-g)|_\star$ m-a.e. in X. (10.2.23) then yields

$$\begin{aligned}
\mathsf{CE}_2(f+g) + \mathsf{CE}_2(f-g) &= \int_X |D(f+g)|_\star^2\, dm + \int_X |D(f-g)|_\star^2\, dm \\
&\leqslant \liminf_{n\to\infty} \int_X |\mathrm{lip}(f_n + g_n)|_\star^2\, dm \\
&\quad + \int_X |\mathrm{lip}(f_n - g_n)|_\star^2\, dm \\
&= \liminf_{n\to\infty} 2\int_X |\mathrm{lip} f_n|_\star^2\, dm \\
&\quad + 2\int_X |\mathrm{lip} g_n|_\star^2\, dm \\
&= 2\mathsf{CE}_2(f) + \mathsf{CE}_2(g).
\end{aligned}$$

Since $H^{1,2}(X, d, m)$ is Banach space, we deduce that $H^{1,2}(X, d, m)$ is an Hilbert space, so it is reflexive. This also shows that $\mathscr{A}$ is strongly dense. □

Remark 10.2.5. In the framework of Theorem 10.2.4, there exists a scalar product $\langle\cdot,\cdot\rangle_{H^{1,2}}$ on $H^{1,2}(X,d,m)$ inducing the norm $\|\cdot\|_{H^{1,2}}$ and satisfying

$$\langle f,g\rangle_{H^{1,2}} = \int_X fg\,dm + CE_2(f,g) \text{ for every } f,g \in H^{1,p}(X,d,m), \tag{10.2.25}$$

where $CE_2(\cdot,\cdot)$ denotes the quadratic form associated to $CE_2(\cdot)$.

10.3 INTRINSIC DISTANCES

By using the general properties of metric Sobolev spaces and the equivalence with the Newtonian viewpoint based on the notion of upper gradient [19, 62], it is possible to improve considerably the density result of Corollary 10.2.2. Let us first recall the notion of *metric velocity*

$$|\dot\gamma|_d(t) := \limsup_{h\to 0}\frac{d(\gamma(t+h),\gamma(t))}{|h|} \tag{10.3.1}$$

and *length*

$$\begin{aligned}\ell_d(\gamma,[\alpha,\beta]) &:= \sup\Big\{\sum_{n=1}^N d(\gamma(t_{n-1}),\gamma(t_n)) : t_0=\alpha<t_1<\cdots<t_{N-1}<t_N=\beta\Big\}\\ &= \int_\alpha^\beta |\dot\gamma|_d(t)\,dt\end{aligned} \tag{10.3.2}$$

of a d-Lipschitz curve $\gamma : [a,b] \to X$; here $[\alpha,\beta] \subset [a,b]$ and we just write $\ell_d(\gamma)$ for $\ell_d(\gamma,[a,b])$.

If $Y \subset X$ is a given set, we can introduce the *length (or intrinsic) extended distance* $d_{Y,\ell}$ induced by d on Y, as the infimum of the length of Y-valued Lipschitz curves connecting two given points $y_0, y_1 \in Y$:

$$d_{Y,\ell}(y_0,y_1) := \inf\Big\{\ell_d(\gamma) : \gamma \in \mathrm{Lip}([0,1];(Y,d)),\ \gamma(0)=y_0,\ \gamma(1)=y_1\Big\} \tag{10.3.3}$$

$$= \inf\left\{\begin{array}{l}\ell>0 : \gamma\in \mathrm{Lip}([0,\ell];(Y,d)) \text{ s.t.}\\ \gamma(0)=y_0,\ \gamma(\ell)=y_1,\ |\dot\gamma|_d\leqslant 1 \text{ a.e.}\end{array}\right\}. \tag{10.3.4}$$

Clearly we have

$$d(y_0,y_1) \leqslant d_{X,\ell}(y_0,y_1) \leqslant d_{Y,\ell}(y_0,y_1) \quad \text{for every } y_0,y_1 \in Y. \tag{10.3.5}$$

If $g : X \to [0,+\infty]$ is a Borel function, the integral of g along γ is defined by

$$\int_\gamma g := \int_a^b g(\gamma(t))|\dot\gamma|_d(t)\,dt. \tag{10.3.6}$$

It is well known that length and integral are invariant with respect to arc-length reparametrization of γ and it is always possible to find a 1-Lipschitz curve $R_\gamma : [0, \ell_d(\gamma)] \to X$ such that

$$R_\gamma(\ell_d(\gamma, [a,t]) = \gamma(t) \text{ for } t \in [a,b], \quad |\dot{R}_\gamma|(s) = 1 \text{ a.e. in } [0, \ell_d(\gamma)], \\ \int_{R_\gamma} g = \int_\gamma g \tag{10.3.7}$$

for every nonnegative Borel function g (see e.g. [108, Section 3.3]). A Borel function $g : X \to [0, +\infty]$ is an upper gradient of $f : X \to \mathbb{R}$ if

$$|f(\gamma(b)) - f(\gamma(a))| \leqslant \int_\gamma g \quad \text{for every } \gamma \in \mathrm{Lip}([a,b]; (X,d)) \tag{10.3.8}$$

Functions in $\mathcal{L}^p(X, m)$ wich admits an upper gradient in $\mathcal{L}^p(X, m)$ characterize the Newtonian-Sobolev space $N^{1,p}(X, d, m)$ [19, 62]. We state here a useful consequence of the main equivalence results [7, Theorem 6.2] [6, Theorem 7.4].

Theorem 10.3.1. *Let Y be a Borel subset of X of full m measure (i.e. $m(X \setminus Y) = 0$) satisfying*

$$\gamma \in \mathrm{Lip}([a,b]; (X,d)), \quad R_\gamma(s) \in Y \text{ for } \mathscr{L}^1\text{-a.e. } s \in [0, \ell_d(\gamma)] \quad \Rightarrow \quad \gamma([a,b]) \subset Y, \tag{10.3.9}$$

let $f : X \to \mathbb{R}$ be a m measurable function and let $g : Y \to [0, +\infty]$ be a Borel function satisfying

$$|f(\gamma(b)) - f(\gamma(a))| \leqslant \int_\gamma g \quad \text{for every } \gamma \in \mathrm{Lip}([a,b]; (Y,d)). \tag{10.3.10}$$

If $\int_Y |g|^p \, dm < \infty$ then f has a p-relaxed gradient and

$$|Df|_\star \leqslant g \quad m\text{-a.e. in } Y. \tag{10.3.11}$$

Notice that condition (10.3.10) is weaker than (10.3.8), since the upper gradient condition is imposed only along curves taking values in Y; however, starting from any function $g \in \mathcal{L}^p(Y, m)$ satisfying (10.3.10) we can define a new Borel function $\tilde{g} : X \to [0, +\infty]$ whose restriction to Y coincides with g such that $\tilde{g}|_{X \setminus Y} \equiv +\infty$. Clearly

$$\int_X \tilde{g}^p \, dm = \int_Y g^p \, dm < +\infty \quad \text{since } m(X \setminus Y) = 0.$$

Moreover $\tilde{g}$ is an upper gradient for f according to (10.3.8): in fact it is sufficient to check (10.3.8) for those curves γ with $\gamma = R_\gamma$ and $\int_\gamma \tilde{g} < +\infty$; since $\tilde{g}(\gamma(s)) = +\infty$ if $\gamma(s) \notin Y$, we deduce that $\gamma(s) \in Y$ for m-a.e. $s \in [0, \ell_d(\gamma)]$ so that $\gamma \in \mathrm{Lip}([0, \ell_d(\gamma)]; (Y, d))$ by (10.3.9), and (10.3.8) then follows by (10.3.10).

It is also immediate to check that (10.3.9) holds if Y is closed.

We consider the situation where

(A) $Y \subset X$ is a Borel set with full m-measure satisfying (10.3.9);

(B) a metric $\delta : Y \times Y \to [0, +\infty)$ is given on Y such that (Y, δ) is complete and separable and (recall Remark 10.1.7)

$$d_1(y_1, y_2) \leqslant \delta(y_1, y_2) \leqslant d_{Y,\ell}(y_1, y_2) \quad \text{for every } y_1, y_2 \in Y. \tag{10.3.12}$$

Remark 10.3.2 (Y-intrinsic distance). δ is intrinsically equivalent to d on Y, i.e. every d-Lipschitz curve $\gamma : [0,1] \to Y$ is also δ-Lipschitz, its δ-length coincides with the corresponding d-length, and integration along γ does not depend on the choice of the distance. In particular condition (10.3.10) can be equivalently stated in terms of δ.
To see that these conditions are implied by (10.3.12), let us fix a d-Lipschitz curve $\gamma : [0,1] \to Y$ with Lipschitz constant bounded by $L \geqslant 0$; then

$$d_{Y,\ell}(\gamma(s), \gamma(t)) \leqslant \ell_d\left(\gamma|_{[s,t]}\right) = \int_s^t |\dot\gamma|_d(r)\,dr \leqslant L|t-s| \quad 0 \leqslant s \leqslant t \leqslant 1,$$

so that γ is $d_{Y,\ell}$-Lipschitz continuous and thus, by (10.3.12), also δ-Lipschitz continuous. To see that the δ and the d-lengths of γ coincide, it is enough to show that $\ell_\delta(\gamma) \leqslant \ell_d(\gamma)$, since (10.3.12) and the trivial equality $\ell_{d_1}(\gamma) = \ell_d(\gamma)$ already give the other inequality; by (10.3.12) we immediately have $\ell_\delta(\gamma) \leqslant \ell_{d_{Y,\ell}}(\gamma)$ and by the very definition of $d_{Y,\ell}$ we see that $\ell_{d_{Y,\ell}}(\gamma) \leqslant \ell_d(\gamma)$. Finally, to see that the integral along γ does not depend on the choice of the distance, it is enough to see that $|\dot\gamma|_d = |\dot\gamma|_\delta$ a.e. in $[0,1]$. The $\leqslant$ inequality is an immediate consequence of (10.3.12) and (10.3.1), while the $\geqslant$ follows by

$$\frac{\delta(\gamma(s), \gamma(t))}{t-s} \leqslant \frac{\ell_\delta(\gamma|_{[s,t]})}{t-s} = \frac{\ell_d(\gamma|_{[s,t]})}{t-s} = \frac{1}{t-s}\int_s^t |\dot\gamma|_d(r)\,dr \quad 0 \leqslant s < t \leqslant 1,$$

and passing to the limit as $s \to t$ for every Lebesgue point t of $|\dot\gamma|_d$.
Since $\mathfrak{m}(X \setminus Y) = 0$ we can identify $L^p(Y, \mathfrak{m})$ with $L^p(X, \mathfrak{m})$. In general, the topology induced by δ is finer than the d topology on Y, and they coincide if δ is continuous w.r.t. d. It is also clear from property (B) that the restriction to Y of every bounded d-Lipschitz function $f : X \to \mathbb{R}$ is also δ-Lipschitz. Thanks to (10.3.12) (which in particular implies that δ-balls of radious $r < 1$ centered at some point $y \in Y$ are included in d-balls of the same radious and with the same center) it is also clear that

$$\text{lip}_\delta f(y) \leqslant \text{lip}_d f(y) \quad \text{for every } y \in Y,\ f \in \text{Lip}_b(X, d). \tag{10.3.13}$$

Since $\text{lip}_\delta f$ is bounded and δ-u.s.c. in Y, it is $\mathfrak{m}$-measurable and we can define the δ pre-Cheeger energy

$$\text{pCE}_{p,\delta}(f) := \int_Y |\text{lip}_\delta f(y)|^p \,d\mathfrak{m}(y) \tag{10.3.14}$$

and we can still consider its l.s.c. envelope in $L^0(Y, \mathfrak{m})$

$$\text{CE}_{p,\delta,\mathscr{A}}(f) := \inf\left\{ \liminf_{n\to\infty} \text{pCE}_{p,\delta}(f_n) : f_n \in \mathscr{A},\ f_n \to f \text{ in } L^0(X, \mathfrak{m}) \right\}. \tag{10.3.15}$$

When $\mathscr{A} = \text{Lip}_b(X, d)$ we simply write $\text{CE}_{p,\delta}(f)$.

Theorem 10.3.3. *If* $\mathscr{A}$ *is a separating unital subalgebra of* $\mathrm{Lip}_b(X, d)$ *satisfying* (10.2.1) *and* (Y, δ) *satisfies the conditions* (A), (B) *above, we have*

$$\mathsf{CE}_{p,\delta}(f) = \mathsf{CE}_{p,\delta,\mathscr{A}}(f) = \mathsf{CE}_{p,\mathscr{A}}(f) = \mathsf{CE}_p(f) \quad \textit{for every } f \in L^0(X, \mathfrak{m}). \tag{10.3.16}$$

In particular, the minimal p*-relaxed gradients of* $f \in L^0(X, \mathfrak{m})$ *computed w.r.t.* $(\delta, \mathscr{A})$*,* $(\delta, \mathrm{Lip}_b(Y))$*,* $(d, \mathscr{A})$ *or* $(d, \mathrm{Lip}_b(X))$ *coincide and we have*

$$D^{1,p}(Y, \delta, \mathfrak{m}) = D^{1,p}(Y, \delta, \mathfrak{m}; \mathscr{A}) = D^{1,p}(X, d, \mathfrak{m}) = D^{1,p}(X, d, \mathfrak{m}; \mathscr{A}).$$

Proof. Since $p\mathsf{CE}_{p,\delta}(f) \leqslant \int_X \left(\mathrm{lip}_d f(x)\right)^p \, d\mathfrak{m}$ for every $f \in \mathrm{Lip}_b(X, d)$, we clearly have

$$\mathsf{CE}_{p,\delta}(f) \leqslant \mathsf{CE}_{p,\delta,\mathscr{A}}(f) \leqslant \mathsf{CE}_{p,\mathscr{A}}(f) = \mathsf{CE}_p(f) \quad \text{for every } f \in L^0(X, \mathfrak{m}),$$

where the last equality follows from Corollary 10.2.2. It is then sufficient to prove that $\mathsf{CE}_{p,\delta}(f) \geqslant \mathsf{CE}_p(f)$ in order to get (10.3.16). Using (10.3.15) and the $L^0(X, \mathfrak{m})$-lower semicontinuity of CE_p (see Remark 10.1.4), the latter inequality will be a consequence of

$$\int_Y |\mathrm{lip}_\delta f(y)|^p \, d\mathfrak{m}(y) \geqslant \mathsf{CE}_p(f) \quad \text{for every } f \in \mathrm{Lip}_b(X, d). \tag{10.3.17}$$

In order to prove (10.3.17) it is sufficient to apply Theorem 10.3.1 and prove that the Borel function $g := \mathrm{lip}_\delta f$ satisfies (10.3.10). Now we use the fact that the restriction to Y of a function $f \in \mathrm{Lip}_b(X)$ belongs to $\mathrm{Lip}_b(Y, \delta)$ and every d-Lipschitz curve γ with values in Y is also δ-Lipschitz, the respective length coincide and therefore also the arc-length reparametrizations are the same. Since lip_δ is an upper gradient we thus obtain

$$|f(\gamma(b)) - f(\gamma(a))| \leqslant \int_\gamma \mathrm{lip}_\delta f. \qquad \square$$

11 THE WASSERSTEIN-SOBOLEV SPACE

In this chapter, we apply the results of Chapter 10 to the Sobolev-Wasserstein space $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,d_{\mathbb{M}}}, \mathfrak{m})$ arising from a measure $\mathfrak{m}$ on the Wasserstein space $(\mathcal{P}_2(\mathbb{M}), W_{2,d_{\mathbb{M}}})$ of probability measures in a complete Riemannian manifold or a separable Hilbert space $\mathbb{M}$. In particular in Section 11.1 we give the main definitions, we study the space of cylindrical functions and we prove the main density result; in Section 11.2 we provide an explicit characterization for the notion of $\mathfrak{m}$-Wasserstein gradient, showing useful calculus rules; finally in Section 11.3 we treat the case where the underlying space is not $\mathbb{R}^d$ but it is a complete Riemannian manifold or a separable Hilbert space.

This Chapter is the result of a collaboration with Massimo Fornasier and Giuseppe Savaré.

11.1 CONSTRUCTION AND PROPERTIES OF THE WASSERSTEIN-SOBOLEV SPACE

In this section we consider the metric space $\mathcal{P}_2(\mathbb{R}^d)$, endowed with the L^2-Wasserstein distance $d = W_2$ and a finite positive Borel measure $\mathfrak{m}$. We will denote by $\mathbb{W}_2 = \mathbb{W}_2(\mathbb{R}^d, \mathfrak{m})$ the metric-measure space $(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ and we want to study the Wasserstein-Sobolev space $H^{1,2}(\mathbb{W}_2)$.

Let us start with the following useful results about optimal potentials for the optimal transport problem in $\mathbb{R}^d$.

Theorem 11.1.1. *Let $\mu, \nu \in \mathcal{P}_2^r(\mathbb{R}^d)$ with $\operatorname{supp}\nu = \overline{B(0,R)}$ for some $R > 0$. Then there exists a unique pair of lower semicontinuous and convex functions*

$$\varphi : B(0,R) \to (-\infty, +\infty], \quad \varphi^* : \mathbb{R}^d \to (-\infty, +\infty]$$

such that

(i) $\varphi^*(y) = \sup_{x \in B(0,R)} \{\langle x, y\rangle - \varphi(x)\}$ *for every* $y \in \mathbb{R}^d$,

(ii) $\varphi^*(0) = 0$,

(iii) $\displaystyle\int_{B(0,R)} \varphi \, d\nu + \int_{\mathbb{R}^d} \varphi^* \, d\mu = \frac{1}{2} m_2^2(\nu) + \frac{1}{2} m_2^2(\mu) - \frac{1}{2} W_2^2(\nu, \mu).$

Moreover the pair (φ, φ^*) satisfies

1. *φ and φ^* are real valued,*

2. $\inf_{B(0,R)} \varphi = 0$,

3. $\displaystyle W_2^2(\mu, \nu) = \int_{B(0,R)} |x - \nabla\varphi(x)|^2 \, d\nu(x) = \int_{\mathbb{R}^d} |y - \nabla\varphi^*(y)|^2 \, d\mu(y),$

4. φ^* *is* R*-Lipschitz continuous.*

Proof. Let $c : \mathbb{R}^d \times \mathbb{R}^d \to [0, +\infty)$ be defined as

$$c(x, y) := \frac{1}{2}|x - y|^2 \quad x, y \in \mathbb{R}^d.$$

Since the cost function c is continuous, finite and it is in $L^1(\mathbb{R}^d \times \mathbb{R}^d, \boldsymbol{\gamma})$ for every $\boldsymbol{\gamma} \in \Gamma(\nu, \mu)$, it is a classical result (see e.g. [5, Theorem 6.4.1]) that, given $\boldsymbol{\gamma}_0 \in \Gamma_o(\nu, \mu)$ we can find a c-concave (see [5, Definition 6.1.2]) function $\phi : \mathbb{R}^d \to [-\infty, +\infty)$ such that $\phi \in L^1(\mathbb{R}^d, \nu)$, $\phi^c \in L^1(\mathbb{R}^d, \mu)$ and

$$\phi(x) + \phi^c(y) = c(x, y) \quad \text{for every } (x, y) \in \operatorname{supp} \boldsymbol{\gamma}_0, \tag{11.1.1}$$

where $\phi^c : \mathbb{R}^d \to [-\infty, +\infty)$ is the c-transform of ϕ defined as

$$\phi^c(y) := \inf_{x \in \mathbb{R}^d} \{c(x, y) - \phi(x)\}, \quad y \in \mathbb{R}^d.$$

We thus get that the function $x \mapsto \frac{1}{2}|x|^2 - \phi(x)$ is lower semicontinuous, convex and it is not identically equal to $+\infty$ on the ball $B(0, R)$ since

$$\int_{B(0,R)} |\phi| d\nu = \int_{\overline{B(0,R)}} |\phi| d\nu = \int_{\mathbb{R}^d} |\phi| d\nu = \|\phi\|_{L^1(\mathbb{R}^d, \nu)} < +\infty,$$

due to the fact that $\nu(\partial B(0, R)) = 0$, being ν absolutely continuous w.r.t. the d-dimensional Lebesgue measure on $\mathbb{R}^d$.
Thus $m := \inf_{B(0,R)} \phi \in \mathbb{R}$; let us define $\tilde{\varphi} : \mathbb{R}^d \to (-\infty, +\infty]$ as

$$\tilde{\varphi}(x) := \frac{1}{2}|x|^2 - \phi(x) - m, \quad x \in \mathbb{R}^d;$$

it is clear that $\tilde{\varphi}$ is convex, lower semicontinuous and $\inf_{B(0,R)} \tilde{\varphi} = 0$. Let us now prove that $\tilde{\varphi}(x) \in \mathbb{R}$ for every $x \in B(0, R)$. First of all we observe that

$$D := \left\{ x \in \overline{B(0, R)} \mid \tilde{\varphi}(x) \in \mathbb{R} \right\}$$

is a convex nonempty set with $\nu(D) = 1$ and $\nu(\partial D) = 0$, so that $\nu(\operatorname{int}(D)) = 1$ and in particular $\operatorname{int}(D) \neq \emptyset$. Since the first projection P of $\operatorname{supp} \boldsymbol{\gamma}$ on $\mathbb{R}^d$ is dense in $\operatorname{supp} \nu = \overline{B(0, R)}$ and $\tilde{\varphi}$ is finite on P, we get that $\overline{D} = \overline{B(0, R)}$. Then $D \supset \operatorname{int}(D) = \operatorname{int}(\overline{D}) = \operatorname{int}\left(\overline{B(0, R)}\right) = B(0, R)$ so that $B(0, R) \subset D$.
Let us define $\tilde{\varphi}^*, \varphi^* : \mathbb{R}^d \to (-\infty, +\infty]$ as

$$\tilde{\varphi}^*(y) := \sup_{x \in \mathbb{R}^d} \{\langle x, y \rangle - \tilde{\varphi}(x)\}, \qquad y \in \mathbb{R}^d,$$
$$\varphi^*(y) := \sup_{x \in B(0,R)} \{\langle x, y \rangle - \tilde{\varphi}(x)\}, \quad y \in \mathbb{R}^d.$$

It is clear that $\tilde{\varphi}^*$ and φ^* are convex and lower semicontinuous functions with $\varphi^*(0) = -\inf_{B(0,R)} \varphi = 0$. Moreover from (11.1.1) it follows that

$$\tilde{\varphi}(x) + \tilde{\varphi}^*(y) = \langle x, y \rangle \quad \text{for every } (x, y) \in \operatorname{supp} \boldsymbol{\gamma}_0. \tag{11.1.2}$$

By their definition, it is clear that $\tilde{\varphi}^*(y) \geqslant \varphi^*(y)$ for every $y \in \mathbb{R}^d$. Since $\operatorname{supp}\nu = \overline{B(0,R)}$ and $\nu(\partial B(0,R)) = 0$, we get by (11.1.2) that for μ-a.e. $y \in \mathbb{R}^d$, there exists $x \in B(0,R)$ such that $\tilde{\varphi}(x) + \tilde{\varphi}^*(y) = \langle x, y\rangle$. This gives that $\varphi^* = \tilde{\varphi}^*$ μ-a.e. so that, still by (11.1.2) one gets

$$\int_{\mathbb{R}^d} \tilde{\varphi} d\nu + \int_{\mathbb{R}^d} \tilde{\varphi}^* d\mu = \int_{B(0,R)} \varphi d\nu + \int_{\mathbb{R}^d} \varphi^* d\mu = \frac{1}{2}m_2^2(\nu) + \frac{1}{2}m_2^2(\mu) - \frac{1}{2}W_2^2(\nu,\mu),$$

where we have defined φ as the restriction of $\tilde{\varphi}$ to $B(0,R)$. We have already noticed that $\inf_{B(0,R)} \varphi = 0$ while the R-Lipschitz continuity of φ^* immediately follows by its definition. The equality

$$W_2^2(\mu,\nu) = \int_{B(0,R)} |x - \nabla\varphi(x)|^2 \, d\nu(x) = \int_{\mathbb{R}^d} |y - \nabla\varphi^*(y)|^2 \, d\mu(y)$$

is classical and can be obtained starting from (iii) with a standard argument (see e.g. the third step in the proof of Theorem 2.12 in [118]). This shows the existence of a pair (φ, φ^*) as in the statement satisfying points (i) – (iii) and (1) – (4).
Let us show that points (i) – (iii) are also sufficient to get uniqueness. If (φ_0, φ_0^*) is another pair as in the statement satisfying points (i) – (iii), then, again arguing e.g. as in the third step in the proof of Theorem 2.12 in [118], one gets by (iii) that both $\nabla\varphi$ and $\nabla\varphi_0$ are optimal transport maps from ν to μ, implying that $\nabla\varphi_0 = \nabla\varphi$ $\mathcal{L}^d$-a.e. in $B(0,R)$ by the a.e. uniqueness of the optimal transport map (see e.g. [118, Theorem 2.12]). Since $\inf_{B(0,R)} \varphi = \inf_{B(0,R)} \varphi_0 = 0$ by (ii), we get that $\varphi = \varphi_0$ in $B(0,R)$ which gives by (i) that also $\phi^* = \phi_0^*$ in $\mathbb{R}^d$. □

11.1.1 *The algebra of cylindrical functions*

We denote by $C_b^1(\mathbb{R}^d)$ the space of bounded and Lipschitz C^1 functions $\phi : \mathbb{R}^d \to \mathbb{R}$. Every $\phi \in C_b^1(\mathbb{R}^d)$ induces the function L_ϕ on $\mathcal{P}(\mathbb{R}^d)$

$$L_\phi : \mu \to \int_{\mathbb{R}^d} \phi \, d\mu \tag{11.1.3}$$

which clearly belongs to $\mathrm{Lip}_b(\mathcal{P}_2(\mathbb{R}^d), W_2)$ thanks to (2.4.5). More generally, if $\boldsymbol{\phi} = (\phi_1, \cdots, \phi_N) \in \big(C_b^1(\mathbb{R}^d)\big)^N$, we denote by

$$L_{\boldsymbol{\phi}} := (L_{\phi_1}, \cdots, L_{\phi_N}) \tag{11.1.4}$$

the corresponding map from $\mathcal{P}_2(\mathbb{R}^d)$ to $\mathbb{R}^N$.

Our construction is based on the algebra of C^1 cylindrical functions generated by (11.1.3) and it is quite similar to the one of [44, Section 2]. Working in the flat space $\mathbb{R}^d$ allows for a further simplification in the structure of the tangent bundle and of corresponding vector fields.

Definition 11.1.2 (C^1-Cylindrical functions)**.** We say that a function $F : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is a C^1*-cylindrical function* if there exist $N \in \mathbb{N}$, $\psi \in C_b^1(\mathbb{R}^N)$ and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_N) \in (C_b^1(\mathbb{R}^d))^N$ such that

$$F(\mu) = \psi(L_{\boldsymbol{\phi}}(\mu)) = \psi\big(L_{\phi_1}(\mu), \cdots, L_{\phi_N}(\mu)\big) \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.1.5}$$

We denote the set of such functions by $\mathrm{Cyl}\big(\mathcal{P}_2(\mathbb{R}^d)\big)$.

Remark 11.1.3. Notice that $\mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right) \subset \mathrm{Lip}_b(\mathcal{P}_2(\mathbb{R}^d), W_2)$ is a unital subalgebra.

Remark 11.1.4. Since for every $\boldsymbol{\phi} \in \left(C_b^1(\mathbb{R}^d)\right)^N$ the range of $L_{\boldsymbol{\phi}}$ is always contained in the bounded set $[-M, M]^N$ where $M := \max_{i=1,\dots,d} \|\phi_i\|_\infty$, also functions $F = \psi \circ L_{\boldsymbol{\phi}}$ with $\psi \in C^1(\mathbb{R}^N)$ belong to $\mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$. Indeed it is enough to consider a function $\tilde{\psi} \in C_b^1(\mathbb{R}^N)$ coinciding with ψ on $[-M, M]^N$ and equal to 0 outside $[-M-1, M+1]^N$ so that $F = \tilde{\psi} \circ L_{\boldsymbol{\phi}}$. In particular every function of the form L_ϕ, $\phi \in C_b^1(\mathbb{R}^d)$, belongs to $\mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$.

Let us consider the set

$$\mathcal{D} := \left\{(\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d : x \in \mathrm{supp}(\mu)\right\}. \tag{11.1.6}$$

The set $\mathcal{D}$ is a Borel set (in fact it is a G_δ): if $(r_n)_n = \mathbb{Q} \cap (0, +\infty)$ we have that $\mathcal{D} = \cap_n \mathcal{D}_n$, where

$$\mathcal{D}_n := \left\{(\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d : \mu(B(x, r_n)) > 0\right\},$$

and each $\mathcal{D}_n$ is open in $\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$, being the inverse image of $(0, +\infty)$ through the lower semicontinuous map $(\mu, x) \mapsto \mu(B(x, r_n))$.

Definition 11.1.5. If $F = \psi \circ L_{\boldsymbol{\phi}} \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ as in (11.1.5) for some $N \in \mathbb{N}$, $\psi \in C_b^1(\mathbb{R}^N)$ and $\boldsymbol{\phi} \in (C_b^1(\mathbb{R}^d))^N$, then the Wasserstein differential of F, $DF : \overline{\mathcal{D}} \to \mathbb{R}^d$, is defined by

$$DF(\mu, x) := \sum_{n=1}^{N} \partial_n \psi\left(L_{\boldsymbol{\phi}}(\mu)\right) \nabla \phi_n(x), \quad (\mu, x) \in \overline{\mathcal{D}}. \tag{11.1.7}$$

We will also denote by $DF[\mu]$ the function $x \mapsto DF(\mu, x)$ and we will set

$$\|DF[\mu]\|_\mu^2 := \int_{\mathbb{R}^d} |DF[\mu](x)|^2 d\mu(x), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.1.8}$$

Remark 11.1.6. It is not difficult to check that

$$DF \text{ is continuous in } \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \tag{11.1.9}$$

with respect to the natural product (narrow and euclidean) topology.

In principle DF (and thus $\|DF[\mu]\|_\mu$) may depend on the choice of $N \in \mathbb{N}$, $\psi \in C_b^1(\mathbb{R}^N)$ and $\boldsymbol{\phi} \in (C_b^1(\mathbb{R}^d))^N$ used to represent F. In Proposition 11.1.10 we show that for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ the function $DF[\mu]$ is uniquely characterized in $\mathrm{supp}(\mu)$ and $\|DF[\mu]\|_\mu$ is well defined, so that DF is uniquely characterized by F in $\mathcal{D}$. By (11.1.9), DF is also uniquely characterized by F on $\overline{\mathcal{D}}$.

We have seen that the Wasserstein differential DF can be considered as a map from $\overline{\mathcal{D}}$ with values in $\mathbb{R}^d$. It is natural to introduce the measure $\boldsymbol{m} = \int \delta_\mu \otimes \mu \, d\mathfrak{m}(\mu) \in \mathcal{P}(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d)$ obtained integrating the measures μ w.r.t. $\mathfrak{m}$: for every bounded Borel function $H : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ we have

$$\int H(\mu, x) \, d\boldsymbol{m}(\mu, x) = \int_{\mathcal{P}_2(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} H(\mu, x) \, d\mu(x) \right) d\mathfrak{m}(\mu). \tag{11.1.10}$$

Since $\operatorname{supp}(\mathfrak{m}) \subset \overline{\mathcal{D}}$, it is then clear that DF belongs to $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$ and

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \|DF[\mu]\|_\mu^2 \, dm(\mu) = \int_{\overline{\mathcal{D}}} |DF(\mu, x)|^2 \, d\mathfrak{m}(\mu, x). \tag{11.1.11}$$

Lemma 11.1.7. *Let Y be a Polish space and let $G : \mathcal{P}(Y) \times Y \to [0, +\infty)$ be a bounded and continuous function. If $(\mu_n)_{n \in \mathbb{N}}$ is a sequence in $\mathcal{P}(Y)$ narrowly converging to μ as $n \to +\infty$, then*

$$\lim_{n \to \infty} \int_Y G(\mu_n, y) d\mu_n(y) = \int_Y G(\mu, y) d\mu(y).$$

Proof. We set $g_n(x) := G(\mu_n, x)$, $g(x) := G(\mu, x)$. Since G is continuous, g_n converge uniformly to g on compact subsets of Y as $n \to \infty$. Thanks to [5, Lemma 5.2.1] $(g_n)_\sharp \mu_n$ converge narrowly to $g_\sharp \mu$ in $\mathcal{P}(\mathbb{R})$. On the other hand, the support of $(g_n)_\sharp \mu$ is uniformly bounded so that

$$\begin{aligned}
\lim_{n \to \infty} \int_Y G(\mu_n, y) d\mu_n(y) &= \lim_{n \to \infty} \int_{\mathbb{R}} r \, d((g_n)_\sharp \mu_n)(r) \\
&= \int_{\mathbb{R}} r \, d(g_\sharp \mu)(r) \\
&= \int_Y G(\mu, y) \, d\mu(y).
\end{aligned}$$

□

Lemma 11.1.8. *Let $F = \psi \circ L_\Phi \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ as in (11.1.5) and let $(\mu_t)_{t \in [0,1]}$ be an absolutely continuous curve in $\mathcal{P}_2(\mathbb{R}^d)$. Then*

$$F(\mu_1) - F(\mu_0) = \int_0^1 \int_{\mathbb{R}^d} \langle DF[\mu_t](x), v_t(x) \rangle d\mu_t(x) \, dt, \tag{11.1.12}$$

where $v_t \in L^2(\mathbb{R}^d, \mu_t; \mathbb{R}^d)$ is the Wasserstein velocity field (cf. Theorem 2.4.6) of $(\mu_t)_{t \in [0,1]}$ at time t and DF is as in (11.1.7).

Proof. Observe that, since F is Lipschitz continuous and $t \mapsto \mu_t$ is absolutely continuous, the map $t \mapsto F(\mu_t)$ is absolutely continuous and thus it holds

$$F(\mu_1) - F(\mu_0) = \int_0^1 \frac{d}{dt} F(\mu_t) dt.$$

It is then enough to prove that

$$\frac{d}{dt} F(\mu_t) = \int_{\mathbb{R}^d} \langle DF(\mu_t, x), v_t(x) \rangle \, d\mu_t(x) \quad \text{for a.e. } t \in (0, 1). \tag{11.1.13}$$

We have, for every $t \in A((\mu_t)_{t \in [0,1]}) \subset (0, 1)$ (cf. Theorem 2.4.6), that

$$\begin{aligned}
\frac{d}{dt} F(\mu_t) &= \sum_{i=1}^N \partial_i \psi(L_\Phi(\mu_t)) \frac{d}{dt} \int_{\mathbb{R}^d} \phi_i \, d\mu_t \\
&= \sum_{i=1}^N \partial_i \psi(L_\Phi(\mu_t)) \int_{\mathbb{R}^d} \langle \nabla \phi_i, v_t(x) \rangle \, d\mu_t(x) \\
&= \int_{\mathbb{R}^d} \langle DF(\mu_t, x), v_t(x) \rangle \, d\mu_t(x),
\end{aligned}$$

where we used Theorem 2.4.6. □

Remark 11.1.9. In case the curve $(\mu_t)_{t\in[0,1]}$ has the simple form

$$\mu_t := (\mathbf{i}_{\mathbb{R}^d} + tu)_\sharp\mu, \quad t \in [0,1]$$

for some map $u \in L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$, (11.1.13) holds for every $t \in [0,1]$ and takes the simpler form

$$\frac{d}{dt}F(\mu_t) = \int_{\mathbb{R}^d} \langle DF(\mu_t, x), u(x)\rangle \, d\mu_t(x)$$

and, in particular, we get

$$\lim_{t\downarrow 0} \frac{F(\mu_t) - F(\mu)}{t} = \int_{\mathbb{R}^d} \langle DF(\mu, x), u(x))\rangle \, d\mu(x). \tag{11.1.14}$$

Proposition 11.1.10. *Let* $F = \psi \circ L_{\boldsymbol{\phi}} \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ *as in* (11.1.5). *Then*

$$\|DF[\mu]\|_\mu = \mathrm{lip}F(\mu) \quad \textit{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

In particular $\|DF[\mu]\|_\mu$ *does not depend on the choice of the representation of* F *and* DF *just depends on* F *on* $\overline{\mathcal{D}}$.

Proof. Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and let $(\mu_n', \mu_n'') \in \mathcal{P}_2(\mathbb{R}^d)^2$ with $\mu_n' \neq \mu_n''$ be such that $(\mu_n', \mu_n'') \to (\mu, \mu)$ in W_2 and

$$\lim_n \frac{|F(\mu_n') - F(\mu_n'')|}{W_2(\mu_n', \mu_n'')} = \mathrm{lip}F(\mu).$$

Let us define, for every $t \in [0,1]$, the map $x^t : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ as

$$x^t(x_0, x_1) := (1-t)x_0 + tx_1, \quad (x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d.$$

Using (11.1.12) along $\mu_n^t := x^t_\sharp \boldsymbol{\mu}_n$ for plans $\boldsymbol{\mu}_n \in \Gamma_o(\mu_n', \mu_n'')$ (it is easy to check that $(\mu_t)_{t\in[0,1]}$ is Lipschitz continuous), we get

$$\begin{aligned}
\left|F(\mu_n') - F(\mu_n'')\right| &\leqslant \left|\int_0^1 \int_{\mathbb{R}^d} \langle DF(\mu_n^t, x), v_t(x)\rangle \, d\mu_n^t(x) \, dt\right| \\
&\leqslant \left(\int_0^1 \int_{\mathbb{R}^d} \left|DF(\mu_n^t, x),\right|^2 \, d\mu_n^t(x) \, dt\right)^{\frac{1}{2}} \\
&\quad \left(\int_0^1 \int_{\mathbb{R}^d} |v_t^n(x)|^2 \, d\mu_n^t(x) \, dt\right)^{\frac{1}{2}} \\
&= W_2(\mu_n', \mu_n'') \left(\int_0^1 \int_{\mathbb{R}^d} \left|DF(\mu_n^t, x)\right|^2 \, d\mu_n^t(x) \, dt\right)^{\frac{1}{2}},
\end{aligned}$$

where we used Theorem 2.4.6. Dividing both sides by $W_2(\mu_n', \mu_n'')$, we obtain

$$\frac{|F(\mu_n') - F(\mu_n'')|}{W_2(\mu_n', \mu_n'')} \leqslant \left(\int_0^1 \int_{\mathbb{R}^d} \left|DF(\mu_n^t, x)\right|^2 \, d\mu_n^t(x) \, dt\right)^{\frac{1}{2}}.$$

Observe that $\boldsymbol{\mu}_n \to (\boldsymbol{i}_{\mathbb{R}^d}, \boldsymbol{i}_{\mathbb{R}^d})_\sharp \mu$ in $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ so that $\mu_n^t \to \mu$ in $\mathcal{P}(\mathbb{R}^d)$ for every $t \in [0,1]$. We can pass to the limit as $n \to +\infty$ the above inequality using the dominated convergence Theorem and Lemma 11.1.7 with

$$G(\mu, x) := |DF(\mu, x)|^2, \quad \mu \in \mathcal{P}(\mathbb{R}^d),\, x \in \mathbb{R}^d.$$

We hence get

$$\operatorname{lip}F(\mu) \leqslant \left(\int_0^1 \int_{\mathbb{R}^d \times \mathbb{R}^d} |DF(\mu, x)|^2 \, d\mu(x) \, dt \right)^{\frac{1}{2}} = \|DF[\mu]\|_\mu .$$

This proves one inequality. Let us now consider the map $T : \mathbb{R}^d \to \mathbb{R}^d$ defined as

$$T(x) := DF[\mu](x), \quad x \in \mathbb{R}^d.$$

By definition of $\operatorname{Tan}_\mu(\mathcal{P}_2(\mathbb{R}^d))$, we have that $T \in \operatorname{Tan}_\mu(\mathcal{P}_2(\mathbb{R}^d))$ so that, by Propsition 8.5.6 in [5], we have

$$\lim_{\varepsilon \downarrow 0} \frac{W_2(\mu, (\boldsymbol{i}_{\mathbb{R}^d} + \varepsilon T)_\sharp \mu)}{\varepsilon} = \|T\|_{L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)} = \|DF[\mu]\|_\mu .$$

Moreover, if we apply (11.1.14) to the curve $\mu_\varepsilon := (\boldsymbol{i}_{\mathbb{R}^d} + \varepsilon T)_\sharp \mu$, $\varepsilon \in [0,1]$, we get

$$\lim_{\varepsilon \downarrow 0} \frac{F(\mu_\varepsilon) - F(\mu)}{\varepsilon} = \int_{\mathbb{R}^d} \langle DF(\mu, x), T(x) \rangle \, d\mu(x) = \|DF[\mu]\|_\mu^2 ,$$

thus

$$\operatorname{lip}F(\mu) \geqslant \lim_{\varepsilon \downarrow 0} \frac{F(\mu_\varepsilon) - F(\mu)}{W_2(\mu_\varepsilon, \mu)} = \|DF[\mu]\|_\mu .$$

This shows the other inequality and concludes the proof. □

11.1.2 *The density result*

Recall that for a bounded Lipschitz function $F : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ the pre-Cheeger energy (cf. (10.1.16)) associated to $\mathfrak{m}$ is defined by

$$\mathsf{pCE}_2(F) = \int_{\mathcal{P}_2(\mathbb{R}^d)} \left(\operatorname{lip}F(\mu)\right)^2 d\mathfrak{m}(\mu). \tag{11.1.15}$$

Thanks to Proposition 11.1.10, if F is a cylindrical function in $\operatorname{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ we have a nice equivalent expression

$$\mathsf{pCE}_2(F) = \int_{\mathcal{P}_2(\mathbb{R}^d)} \|DF[\mu]\|_\mu^2 \, d\mathfrak{m}(\mu) = \int |DF(\mu, x)|^2 \, d\boldsymbol{\mathfrak{m}}(\mu, x), \tag{11.1.16}$$

which shows that the restriction of pCE_2 to $\operatorname{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ is a quadratic form (thus satisfying (10.2.23)) induced by the bilinear form

$$\mathsf{pCE}_2(F, G) := \int DF(\mu, x) \cdot DG(\mu, x) \, d\boldsymbol{\mathfrak{m}}(\mu, x), \quad F, G \in \operatorname{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right). \tag{11.1.17}$$

It is therefore important to prove that $\operatorname{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ is dense in energy and therefore $H^{1,2}(\mathbb{W}_2)$ is an Hilbert space: this is precisely the object of our main result.

Theorem 11.1.11. *The algebra* $\mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ *is dense in energy in* $D^{1,2}(\mathbb{W}_2)$*: for every* $F \in D^{1,2}(\mathbb{W}_2)$ *there exists a sequence* $F_n \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$*,* $n \in \mathbb{N}$*, such that*

$$F_n \to F \ \ \mathfrak{m}\text{-a.e.,} \quad \mathrm{lip}(F_n) \to |DF|_\star \text{ in } L^2(X, \mathfrak{m}); \tag{11.1.18}$$

if moreover $F \in L^p(X, \mathfrak{m})$*,* $p \in [1, +\infty)$*, then we can find a sequence* $F_n \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ *as in* (11.1.18) *and converging to* F *in* $L^p(X, \mathfrak{m})$*.*

Corollary 11.1.12. $H^{1,2}(\mathbb{W}_2)$ *is a separable Hilbert space and* $\mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ *is strongly dense in* $H^{1,2}(\mathbb{W}_2)$*.*

According to the terminology introduced in [54] (see also [7]) we can say that $(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ is infinitesimally Hilbertian for every Borel probability measure $\mathfrak{m}$.
We devote the remaining part of this subsection to the proof of the Theorem above. We start with a Lemma from convex analysis which will be useful in the proof of Proposition 11.1.19.

Lemma 11.1.13. *Let* $R > 0$ *and let* $(\varphi_n)_n$ *and* $(\psi_n)_n$ *be sequences of functions such that*

(a) $\varphi_n : \mathbb{R}^d \to \mathbb{R}$ *is convex and* R*-Lipschitz continuous with* $\varphi_n(0) = 0$ *for every* $n \in \mathbb{N}$*;*

(b) $\psi_n : B(0, R) \to \mathbb{R}$ *is convex and lower semicontinuous for every* $n \in \mathbb{N}$*;*

(c) it holds

$$\varphi_n(x) = \sup_{y \in B(0,R)} \{\langle x, y\rangle - \psi_n(y)\} \quad \text{for every } x \in \mathbb{R}^d,$$

for every $n \in \mathbb{N}$*.*

Then there exist a subsequence $j \mapsto n(j)$ *and two convex and lower semicontinuous functions* $\varphi : \mathbb{R}^d \to \mathbb{R}$ *and* $\psi : B(0, R) \to (-\infty, +\infty]$ *such that*

(i) $\varphi_{n(j)} \to \varphi$ *locally uniformly on* $\mathbb{R}^d$*;*

(ii) $\liminf_j \psi_{n(j)}(y) \geqslant \psi(y) \geqslant 0$ *for every* $y \in B(0, R)$*;*

(iii) it holds

$$\varphi(x) = \sup_{y \in B(0,R)} \{\langle x, y\rangle - \psi(y)\} \quad \text{for every } x \in \mathbb{R}^d;$$

(iv) $\nabla\varphi_{n(j)} \to \nabla\varphi$ $\mathcal{L}^d$*-a.e. on* $\mathbb{R}^d$*.*

Proof. We start the proof with the following trivial remark: if $u : \overline{B(0, R)} \to [-\infty, +\infty)$ is a concave function, then

$$\sup_{\overline{B(0,R)}} u = \sup_{B(0,R)} u. \tag{11.1.19}$$

In fact, if we take $y_0 \in \partial B(0, R)$, by concavity of u we have that

$$\lim_{t \uparrow 1} u(ty_0) \geqslant u(y_0)$$

and, since $ty_0 \in B(0, R)$ for every $t \in [0, 1)$, we get that

$$\sup_{y \in B(0,R)} u(y) \geqslant \sup_{y \in \partial B(0,R)} u(y)$$

thus giving (11.1.19).

By (a) and Arzelà-Ascoli Theorem, we can find a subsequence $j \mapsto n(j)$ and a convex and R-Lipschitz function $\varphi : \mathbb{R}^d \to \mathbb{R}$ such that $\varphi_{n(j)} \to \varphi$ locally uniformly on $\mathbb{R}^d$ and $\varphi(0) = 0$; this proves (i). In particular, $\varphi_{n(j)}$ Mosco converges (see e.g. [11, Definition 3.17, Proposition 3.19]) to φ. If we denote by $\tilde{\psi}_j$ the convex and lower semicontinuous envelope of the extension of $\psi_{n(j)}$ to $\mathbb{R}^d$ (set equal to $+\infty$ outside $B(0, R)$), we have that $\varphi_{n(j)}$ coincides with the Legendre transform $\mathcal{L}(\tilde{\psi}_j)$ of $\tilde{\psi}_j$; indeed, for every $x \in \mathbb{R}^d$, we have

$$\begin{aligned}
\mathcal{L}(\tilde{\psi}_j)(x) &:= \sup_{y \in \mathbb{R}^d} \{\langle x, y\rangle - \tilde{\psi}_j(y)\} \\
&= \sup_{y \in \overline{B(0,R)}} \{\langle x, y\rangle - \tilde{\psi}_j(y)\} \\
&= \sup_{y \in B(0,R)} \{\langle x, y\rangle - \tilde{\psi}_j(y)\} \\
&= \sup_{y \in B(0,R)} \{\langle x, y\rangle - \psi_{n(j)}(y)\} \\
&= \varphi_{n(j)}(x),
\end{aligned}$$

where the second equality comes from the fact that $\tilde{\psi}_j = +\infty$ outside $\overline{B(0, R)}$, the second follows by (11.1.19) and (b), the third is a consequence of the equality $\tilde{\psi}_j = \psi_{(n(j)}$ in $B(0, R)$, following from (b), and the last one is (c).
Thus (by e.g. [11, Theorem 3.18]) we get that $\tilde{\psi}_j$ Mosco converges to a proper, lower semicontinuous and convex function $\tilde{\psi} : \mathbb{R}^d \to (-\infty, +\infty]$ such that φ is the Legendre transform of $\tilde{\psi}$. By Mosco convergence we have that $\tilde{\psi} = +\infty$ outside $\overline{B(0, R)}$ so that, if we denote by ψ the restriction of $\tilde{\psi}$ to $B(0, R)$, we get that, for every $x \in \mathbb{R}^d$, it holds

$$\begin{aligned}
\varphi(x) &= \mathcal{L}(\tilde{\psi})(x) \\
&= \sup_{y \in \mathbb{R}^d} \{\langle x, y\rangle - \tilde{\psi}(y)\} \\
&= \sup_{y \in \overline{B(0,R)}} \{\langle x, y\rangle - \tilde{\psi}(y)\} \\
&= \sup_{y \in B(0,R)} \{\langle x, y\rangle - \psi(y)\},
\end{aligned}$$

where the last inequality follows again by (11.1.19); this proves (iii). By Mosco convergence and by $\varphi(0) = \inf_{B(0,R)} \psi$ we also get that $\liminf_j \psi_{n(j)}(y) \geqslant \psi(y) \geqslant 0$ for every $y \in B(0, R)$; this proves (ii). Finally, by [11, Theorem 3.66], the Mosco convergence of $\varphi_{n(j)}$ to φ, implies the graph convergence (see [11, Definition

3.58]) of the subdifferentials $\partial\varphi_{n(j)}$ to the subdifferential $\partial\varphi$. Being the functions $\varphi_{n(j)}$ and φ differentiable $\mathcal{L}^d$-a.e., we obtain that $\nabla\varphi_{n(j)} \to \nabla\varphi$ $\mathcal{L}^d$-a.e. in $\mathbb{R}^d$; this proves (iv) and concludes the proof of the Lemma. □

The following preliminary lemma provides a simple gradient estimate for the distance from the Dirac mass at 0, i.e. the quadratic moment of a measure.

Lemma 11.1.14. *Let $\vartheta \in \mathrm{Lip}(\mathbb{R}^d)$ be a L-Lipschitz function which is continuously differentiable in the open set $\Omega_\vartheta := \{x \in \mathbb{R}^d : \vartheta(x) \neq 0\}$. Then the map*

$$F : \mu \to \big(\mathsf{L}_{\vartheta^2}\big)^{1/2} = \Big(\int_{\mathbb{R}^d} \vartheta^2(x)\, d\mu(x)\Big)^{1/2} \tag{11.1.20}$$

is L-Lipschitz and belongs to $D^{1,2}(\mathbb{W}_2, \mathscr{A})$, in particular its $(2, \mathscr{A})$-relaxed gradient is bounded above by L and satisfies

$$|DF|^2_{\star,\mathscr{A}}(\mu) \leqslant \frac{1}{F^2(\mu)} \int_{\mathbb{R}^d} \vartheta^2 |\nabla\vartheta|^2\, d\mu \quad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d) \textit{ with } F(\mu) > 0. \tag{11.1.21}$$

Proof. Let $T \in C^\infty(\mathbb{R})$ be an odd, nondecreasing truncation function satisfying

$$T(x) = x \quad \text{if } |x| \leqslant 1/2, \quad |T(x)| = 1 \quad \text{if } |x| \geqslant 2, \quad |T'(x)| \leqslant 1, \tag{11.1.22}$$

and let us set $T_n(x) := nT(x/n)$, $\vartheta_n := T_n \circ \vartheta$, so that ϑ_n is L-Lipschitz and is continuously differentiable in Ω_ϑ, so that $\vartheta_n^2 \in C^1_b(\mathbb{R}^d)$.

We define $\psi_n(r) := (r + n^{-2})^{1/2}$ and $F_n := \psi_n \circ \mathsf{L}_{\vartheta_n^2}$. By construction $F_n \in \mathscr{A}$ with

$$\begin{aligned} DF_n(\mu, x) &= \frac{1}{F_n(\mu)} \vartheta_n(x) \nabla\vartheta_n(x), \\ \big(\mathrm{lip} F_n(\mu)\big)^2 = \|DF_n[\mu]\|^2 &= \frac{1}{F_n^2(\mu)} \int_{\mathbb{R}^d} \vartheta_n^2(x) |\nabla\vartheta_n(x)|^2\, d\mu(x) \leqslant L^2 \end{aligned} \tag{11.1.23}$$

Since $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ is a length space we deduce that F_n is L-Lipschitz. On the other hand $\lim_{n\to\infty} F_n(\mu) = F(\mu)$ pointwise everywhere, so that F is L-Lipschitz as well, it belongs to $D^{1,2}(\mathbb{W}_2, \mathscr{A})$ and $|DF|_{\star,\mathscr{A}} \leqslant L$. Passing eventually to the limit as $n \to \infty$ in (11.1.23) for μ in the open set $\{\mu \in \mathcal{P}_2(\mathbb{R}^d) : F(\mu) > 0\}$ we get (11.1.21). □

Given $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, we denote by

$$m_2^2(\mu) := \int_{\mathbb{R}^d} |x|^2\, d\mu(x) \tag{11.1.24}$$

its squared moment. Selecting $\vartheta(x) := |x|$ in the Lemma above, we immediately get the following result.

Corollary 11.1.15. *The function $m_2(\cdot)$ belongs to $D^{1,2}(\mathbb{W}_2, \mathscr{A})$ with*

$$|Dm_2|_{\star,\mathscr{A}}(\mu) \leqslant 1 \quad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.1.25}$$

We now use m_2 for localizing gradient estimates in $\mathcal{P}_2(\mathbb{R}^d)$.

Lemma 11.1.16. *Let F_n be a sequence of functions in $D^{1,2}(\mathbb{W}_2, \mathscr{A})$ such that F_n and $|DF_n|_{\star,\mathscr{A}}$ are uniformly bounded in every bounded set of $\mathcal{P}_2(\mathbb{R}^d)$ and let F, G be Borel function in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$, G nonnegative. If*

$$\lim_{n\to\infty} F_n(\mu) = F(\mu), \quad \limsup_{n\to\infty} |DF_n|_{\star,\mathscr{A}}(\mu) \leqslant G(\mu) \quad \mathfrak{m}\text{-a.e. in } \mathcal{P}_2(\mathbb{R}^d) \tag{11.1.26}$$

then $F \in H^{1,2}(\mathbb{W}_2, \mathscr{A})$ and $|DF|_{\star,\mathscr{A}} \leqslant G$.

Proof. Let us consider a smooth nonincreasing function $\theta \in C^\infty[0, +\infty)$ such that

$$\theta(r) = 1 \quad \text{if } 0 \leqslant r \leqslant 1, \quad \theta(r) = 0 \quad \text{if } r \geqslant 2, \quad |\theta'(r)| \leqslant 2 \tag{11.1.27}$$

and set

$$\chi_n(\mu) := \theta\big(m_2(\mu)/n\big) \tag{11.1.28}$$

By Corollary 11.1.15 we have

$$\begin{gathered} \chi_n \in H^{1,2}(\mathbb{W}_2, \mathscr{A}), \quad |D\chi_n|_{\star,\mathscr{A}} \leqslant 2/n, \\ |D\chi_n|_{\star,\mathscr{A}}(\mu) = 0 \text{ if } m_2(\mu) \leqslant n \text{ or } m_2(\mu) \geqslant 2n. \end{gathered} \tag{11.1.29}$$

Thanks to the Leibniz rule, setting $F_{n,m}(\mu) := F_n(\mu)\chi_m^2(\mu)$ and $G_n := |DF_n|_{\star,\mathscr{A}}$, we have

$$F_{n,m} \in D^{1,2}(\mathbb{W}_2, \mathscr{A}), \quad |DF_{n,m}|_{\star,\mathscr{A}}(\mu) \leqslant G_n(\mu)\chi_m^2(\mu) + 4/\mathfrak{m}F_n(\mu)\chi_m(\mu). \tag{11.1.30}$$

Since for every $m \in \mathbb{N}$ the sequence $n \mapsto G_n\chi_m^2$ is uniformly bounded, we can find an increasing subsequence $k \mapsto n(k)$ such that $k \mapsto G_{n(k)}\chi_m^2$ is weakly* convergent in $L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ and we denote by $\tilde{G}_m$ is weak* limit. By Faotu's lemma, for every Borel set $B \subset \mathcal{P}_2(\mathbb{R}^d)$ we get

$$\begin{aligned} \int_B \tilde{G}_m \, d\mathfrak{m} &= \lim_{k\to\infty} \int_B G_{n(k)}(\mu)\chi_m^2(\mu) \, d\mathfrak{m}(\mu) \\ &\leqslant \int_B \limsup_{k\to\infty} \Big(G_{n(k)}(\mu)\chi_m^2(\mu)\Big) \, d\mathfrak{m}(\mu) \\ &\leqslant \int_B G^2\chi_m^2 \, d\mathfrak{m} \end{aligned}$$

so that we deduce

$$\tilde{G}_m \leqslant G^2\chi_m^2 \quad \mathfrak{m}\text{-a.e. in } \mathcal{P}_2(\mathbb{R}^d), \text{ for every } m \in \mathbb{N}. \tag{11.1.31}$$

On the other hand, passing to the limit in (11.1.30) along the subsequence $n(k)$ and recalling that $\lim_{k\to\infty} F_{n(k),m} = F\chi_m^2$ $\mathfrak{m}$-a.e. we get

$$|D(F\chi_m^2)|_{\star,\mathscr{A}}(\mu) \leqslant \tilde{G}_m(\mu) + \frac{4}{m}F(\mu)\chi_m(\mu) \leqslant G(\mu)\chi_m^2(\mu) + \frac{4}{m}F(\mu)\chi_m(\mu) \tag{11.1.32}$$

for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. We eventually pass to the limit as $m \to \infty$ concluding the proof of the Lemma. □

We now derive a natural estimate, extending (11.1.7) to the case of quadratically coercive functions whose gradient has a linear growth.

Lemma 11.1.17. *Let $\phi \in C^1(\mathbb{R}^d)$ be satisfying the growth conditions*

$$\phi(x) \geqslant A|x|^2 - B, \quad |\nabla\phi(x)| \leqslant C(|x|+1) \quad \text{for every } x \in \mathbb{R}^d \tag{11.1.33}$$

for given positive constants $A, B, C > 0$ and let $\zeta : \mathbb{R} \to \mathbb{R}$ be a C^1 nondecreasing function whose derivative has compact support. Then the function $F(\mu) := \zeta \circ L_\phi$ is Lipschitz in $\mathcal{P}_2(\mathbb{R}^d)$, it belongs to $H^{1,2}(W_2, \mathscr{A})$, and

$$|DF|_{\star,\mathscr{A}}(\mu) \leqslant \zeta'(L_\phi(\mu))\Big(\int_{\mathbb{R}^d} |\nabla\phi(x)|^2 \,d\mu(x)\Big)^{1/2}. \tag{11.1.34}$$

Proof. We set $\zeta_a(z) := (z+a)^{1/2}$ and $\vartheta_a := \zeta_a \circ \phi$, with $a := A + B$, so that

$$\vartheta_a \in C^1(\mathbb{R}^d), \quad \vartheta_a \geqslant \big(A(|x|^2+1)\big)^{1/2}, \quad |\nabla\vartheta_a(x)| = \frac{|\nabla\phi(x)|}{2(\phi(x)+a)^{1/2}} \leqslant L,$$

where $L := A^{-1/2}C$, for every $x \in \mathbb{R}^d$.

We can then apply Lemma 11.1.14, observing that

$$\big(L_{\vartheta_a^2}(\mu)\big)^{1/2} = \zeta_a\big(L_\phi(\mu)\big);$$

we deduce that $F_a = \zeta_a \circ L_\phi$ is L-Lipschitz, it belongs to $D^{1,2}(W_2, \mathscr{A})$ and satisfies (recall (11.1.21))

$$|DF_a|_{\star,\mathscr{A}}(\mu) \leqslant \frac{1}{2F_a(\mu)}\Big(\int_{\mathbb{R}^d} |\nabla(\vartheta_a^2)|^2 \,d\mu\Big)^{1/2} = \zeta_a'(L_\phi(\mu))\Big(\int_{\mathbb{R}^d} |\nabla\phi|^2 \,d\mu\Big)^{1/2}. \tag{11.1.35}$$

We eventually observe that $F = \psi_a \circ F_a$ where $\psi_a(z) = \zeta(z^2 - a)$ which is still C^1 with derivative with compact support. Then (11.1.35) yields (11.1.34). □

Let $\kappa \in C_c^\infty(\mathbb{R}^d)$ be such that $\operatorname{supp}\kappa = \overline{B(0,1)}$, $\kappa(x) \geqslant 0$ for every $x \in \mathbb{R}^d$ and $\kappa(x) > 0$ for every $x \in B(0,1)$, $\int_{\mathbb{R}^d} \kappa \,d\mathcal{L}^d = 1$ and $\kappa(-x) = \kappa(x)$ for every $x \in \mathbb{R}^d$. Let us define, for every $0 < \varepsilon < 1$ the standard mollifiers

$$\kappa_\varepsilon(x) := \frac{1}{\varepsilon^d}\kappa(x/\varepsilon) \quad x \in \mathbb{R}^d.$$

Given $\sigma \in \mathcal{P}_2(\mathbb{R}^d)$ and $0 < \varepsilon < 1$, we define

$$\sigma_\varepsilon := \sigma * \kappa_\varepsilon, \tag{11.1.36}$$

$$\hat\sigma_\varepsilon := \frac{\sigma_\varepsilon \llcorner B(0,1/\varepsilon) + \varepsilon^{d+3}\mathcal{L}^d \llcorner B(0,1/\varepsilon)}{\sigma_\varepsilon(B(0,1/\varepsilon)) + \varepsilon^{d+3}\mathcal{L}^d(B(0,1/\varepsilon))}. \tag{11.1.37}$$

Notice that $\sigma_\varepsilon, \hat\sigma_\varepsilon \in \mathcal{P}_2^r(\mathbb{R}^d)$, $\operatorname{supp}\hat\sigma_\varepsilon = \overline{B(0,1/\varepsilon)}$ and $W_2(\sigma_\varepsilon,\sigma) \to 0$, $W_2(\hat\sigma_\varepsilon,\sigma) \to 0$ as $\varepsilon \downarrow 0$. Moreover, if $\sigma, \sigma' \in \mathcal{P}_2(\mathbb{R}^d)$, we have

$$W_2(\sigma_\varepsilon, \sigma'_\varepsilon) \leqslant W_2(\sigma,\sigma') \quad \text{for every } 0 < \varepsilon < 1 \tag{11.1.38}$$

and it is easy to check that, if we set

$$C_\varepsilon := m_2(\kappa_\varepsilon \mathcal{L}^d), \tag{11.1.39}$$

then we have

$$m_2(\mu_\varepsilon) \leqslant m_2(\mu) + C_\varepsilon \quad \text{for every } 0 < \varepsilon < 1. \tag{11.1.40}$$

Definition 11.1.18. Let $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, $0 < \varepsilon < 1$, $\zeta : \mathbb{R} \to \mathbb{R}$ be a C^1 nondecreasing function whose derivative has compact support. We define the continuous functions $W_\nu^\varepsilon, F_\nu^\varepsilon, W_\nu : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ as

$$F_\nu^\varepsilon(\mu) := \frac{1}{2} W_2^2(\mu_\varepsilon, \hat{\nu}_\varepsilon), \quad W_\nu^\varepsilon(\mu) := \zeta\Big(F_\nu^\varepsilon(\mu)\Big), \quad W_\nu(\mu) := W_2(\mu, \nu).$$

Proposition 11.1.19. *Let* $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, $\varepsilon \in (0,1)$ *and let* $\mathscr{A} = \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$. *We have*

$$|DW_\nu^\varepsilon|_{\star,\mathscr{A}}(\mu) \leqslant \zeta'(F_\nu^\varepsilon)\Big(\int_{\mathbb{R}^d} |x - \nabla(\varphi_\varepsilon^* * \kappa_\varepsilon)(x)|^2 \, d\mu(x)\Big)^{1/2} \quad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d), \tag{11.1.41}$$

where $\varphi_\varepsilon^* = \Phi^*(\hat{\nu}_\varepsilon, \mu_\varepsilon)$.

Proof. Let $\mathcal{G} := \{\mu^h\}_{h \in \mathbb{N}}$ be a dense and countable set in $\mathcal{P}_2(\mathbb{R}^d)$ and let us set, for every $h \in \mathbb{N}$, $\varphi_h := \Phi(\hat{\nu}_\varepsilon, \mu_\varepsilon^h)$, $\varphi_h^* := \Phi^*(\hat{\nu}_\varepsilon, \mu_\varepsilon^h)$ (see Theorem 11.1.1),

$$a_h := \int_{B(0,1/\varepsilon)} \left(\frac{1}{2}|y|^2 - \varphi_h(y)\right) d\hat{\nu}_\varepsilon(y), \quad u_h(x) := \frac{1}{2}|x|^2 - \varphi_h^*(x) + a_h, \quad x \in \mathbb{R}^d$$

and

$$G_k(\mu) := \max_{1 \leqslant h \leqslant k} \int_{\mathbb{R}^d} u_h d\mu_\varepsilon, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

We first observe that $\varphi_h^*(x)$ is $1/\varepsilon$-Lipschitz, so that $|\varphi_h^*(x)| \leqslant |x|/\varepsilon$ and

$$u_h(x) \geqslant \frac{1}{2}|x|^2 - \frac{1}{\varepsilon}|x| + a_h \geqslant \frac{1}{4}|x|^2 - \frac{1}{\varepsilon^2} + a_h \tag{11.1.42}$$

$$u_h(x) \leqslant |x|^2 + \frac{1}{\varepsilon^2} + a_h. \tag{11.1.43}$$

Claim 1. It holds

$$\lim_{k \to +\infty} G_k(\mu) = F_\nu^\varepsilon(\mu) \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Proof of claim 1. Since $G_{k+1}(\mu) \geqslant G_k(\mu)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, we have that

$$\lim_{k \to +\infty} G_k(\mu) = \sup_k G_k(\mu) = \sup_h \int_{\mathbb{R}^d} u_h d\mu_\varepsilon \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

By the definition of φ_h and φ_h^* we have, for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $h \in \mathbb{N}$, that

$$\begin{aligned}\int_{\mathbb{R}^d} u_h d\mu_\varepsilon &= \int_{\mathbb{R}^d} \left(\frac{1}{2}|x|^2 - \varphi_h^*(x)\right) d\mu_\varepsilon + \int_{B(0,1/\varepsilon)} \left(\frac{1}{2}|y|^2 - \varphi_h(y)\right) d\hat{\nu}_\varepsilon(y) \\ &\leqslant \frac{1}{2} W_2^2(\mu_\varepsilon, \hat{\nu}_\varepsilon) \\ &= F_\nu^\varepsilon(\mu).\end{aligned}$$

This proves that $\sup_k G_k(\mu) \leqslant F_\nu^\varepsilon(\mu)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Clearly, if $\mu \in \mathcal{G}$, this is an equality. Let now $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ and $h \in \mathbb{N}$ and observe that

$$\begin{aligned}\int_{\mathbb{R}^d} u_h d\mu_\varepsilon - \int_{\mathbb{R}^d} u_h d\mu'_\varepsilon &= \frac{1}{2} m_2^2(\mu_\varepsilon) - \frac{1}{2} m_2^2(\mu'_\varepsilon) - \int_{\mathbb{R}^d} \varphi_h^* d(\mu_\varepsilon - \mu'_\varepsilon) \\ &\leqslant \frac{1}{2}((m_2(\mu_\varepsilon) + m_2(\mu'_\varepsilon)) W_2(\mu_\varepsilon, \mu'_\varepsilon) + \frac{1}{\varepsilon} W_2(\mu_\varepsilon, \mu'_\varepsilon) \\ &\leqslant \frac{1}{2}(m_2(\mu) + m_2(\mu') + 2C_\varepsilon W_2(\mu, \mu') + \frac{1}{\varepsilon} W_2(\mu, \mu') \\ &\leqslant \left(m_2(\mu) + m_2(\mu') + 2C_\varepsilon + \frac{1}{\varepsilon}\right) W_2(\mu, \mu'),\end{aligned}$$

where we used (11.1.38), (11.1.40), the fact that φ_h^* is $1/\varepsilon$-Lipschitz continuous and (2.4.5). We hence deduce that for every $k \in \mathbb{N}$

$$\left|G_k(\mu) - G_k(\mu')\right| \leqslant \left(m_2(\mu) + m_2(\mu') + \frac{1}{\varepsilon} + 2C_\varepsilon\right) W_2(\mu, \mu') \tag{11.1.44}$$

for every $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$. Choosing $\mu' \in \mathcal{G}$ and passing to the limit as $k \to +\infty$ we get from (11.1.44)

$$\left|\lim_{k \to +\infty} G_k(\mu) - F_\nu^\varepsilon(\mu')\right| \leqslant \left(m_2(\mu) + m_2(\mu') + 2C_\varepsilon + \frac{1}{\varepsilon}\right) W_2(\mu, \mu')$$

for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\mu' \in \mathcal{G}$. Using the density of $\mathcal{G}$ and the continuity of $\mu' \mapsto F_\nu^\varepsilon(\mu')$ we deduce that

$$\lim_{k \to +\infty} G_k(\mu) = F_\nu^\varepsilon(\mu) \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d)$$

proving the first claim.

Claim 2. If $H_k := \zeta \circ G_k$ and $u_{h,\varepsilon} := u_h * \kappa_\varepsilon$ it holds

$$\begin{aligned}|DH_k|^2_{\star,\mathscr{A}}(\mu) &\leqslant \left(\zeta'(G_k(\mu))\right)^2 \int_{\mathbb{R}^d} |\nabla u_{h,\varepsilon}|^2 \, d\mu(x) \\ &= \left(\zeta'(G_k(\mu))\right)^2 \int_{\mathbb{R}^d} |x - \nabla(\varphi_h^* * \kappa_\varepsilon)(x)|^2 \, d\mu(x),\end{aligned}$$

for $\mathfrak{m}$-a.e. $\mu \in B_h^k$, where $B_h^k := \{\mu \in \mathcal{P}_2(\mathbb{R}^d) \mid G_k(\mu) = \int_{\mathbb{R}^d} u_h d\mu_\varepsilon\}$, $h \in \{1, \dots, k\}$.
Proof of claim 2. For every $h \in \mathbb{N}$, (11.1.42) yields

$$u_{h,\varepsilon}(x) \geqslant \frac{1}{4}|x|^2 + \frac{C_\varepsilon^2}{4} - \frac{1}{\varepsilon^2} + a_h, \quad |\nabla u_{h,\varepsilon}(x)| \leqslant |x| + \frac{1}{\varepsilon} + \varepsilon; \tag{11.1.45}$$

Since the map $\ell_h : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ defined as $\ell_h(\mu) := \int_{\mathbb{R}^d} u_h d\mu_\varepsilon$ satisfies

$$\ell_h(\mu) = \int_{\mathbb{R}^d} (u_h * \kappa_\varepsilon) d\mu = L_{u_{h,\varepsilon}}(\mu), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d),$$

Lemma 11.1.17 and the above estimates yield

$$|D(\zeta \circ \ell_h)|_\star(\mu) \leqslant \zeta'(\ell_h(\mu)) \Big(\int_{\mathbb{R}^d} |\nabla u_{h,\varepsilon}|^2 \, d\mu \Big)^{1/2} \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Since H_k can be written as

$$H_k(\mu) = \max_{1 \leqslant h \leqslant k} (\zeta \circ \ell_h)(\mu), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d),$$

we can apply Theorem 10.1.2 (8) and conclude the proof of the second claim.

Claim 3. For every $R > 0$ there exists a constant $C > 0$ independent of h such that

$$\Big(\int_{\mathbb{R}^d} |\nabla u_{h,\varepsilon}(x)|^2 \, d\mu(x) \Big)^{1/2} \leqslant C \quad \text{whenever } m_2(\mu) \leqslant R. \tag{11.1.46}$$

Proof of Claim 3. It is sufficient to use (11.1.45) obtaining $C := R + 1/\varepsilon + \varepsilon$.

Claim 4. Let $(h_n)_n \subset \mathbb{N}$ be an increasing sequence and let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. If $\lim_n \int_{\mathbb{R}^d} u_{h_n} d\mu_\varepsilon = F_\nu^\varepsilon(\mu)$, then

$$\lim_n \int_{\mathbb{R}^d} \left| x - \nabla(\varphi^*_{h_n} * \kappa_\varepsilon)(x) \right|^2 d\mu(x) = \int_{\mathbb{R}^d} |x - \nabla(\varphi^*_\varepsilon * \kappa_\varepsilon)(x)|^2 \, d\mu(x),$$

where $\varphi^*_\varepsilon = \Phi^*(\hat{\nu}_\varepsilon, \mu_\varepsilon)$.
Proof of claim 4. Let us set for every $n \in \mathbb{N}$

$$\phi_n := \varphi_{h_n}, \quad \phi^*_n := \varphi^*_{h_n}.$$

We will show that from any (non relabeled) increasing sequence it is possible to extract a further subsequence $j \mapsto n(j)$ such that

$$\lim_j \int_{\mathbb{R}^d} \left| x - \nabla(\phi^*_{n(j)} * \kappa_\varepsilon)(x) \right|^2 d\mu(x) = \int_{\mathbb{R}^d} |x - \nabla(\varphi^*_\varepsilon * \kappa_\varepsilon)(x)|^2 \, d\mu(x).$$

By Theorem 11.1.1, we have that, for every $n \in \mathbb{N}$, $\phi^*_n : \mathbb{R}^d \to \mathbb{R}$ is convex and $1/\varepsilon$-Lipschitz continuous with $\phi^*_n(0) = 0$, $\phi_n : B(0, 1/\varepsilon) \to \mathbb{R}$ is convex and lower semicontinuous and

$$\phi^*_n(x) = \sup_{y \in B(0,1/\varepsilon)} \{ \langle x, y \rangle - \phi_n(y) \} \quad \text{for every } x \in \mathbb{R}^d.$$

Thus by Lemma 11.1.13, we get the existence of a subsequence $j \mapsto n(j)$ and two convex and lower semicontinuous functions $\phi^* : \mathbb{R}^d \to \mathbb{R}$ and $\phi : B(0, 1/\varepsilon) \to (-\infty, +\infty]$ such that points (i), (ii), (iii) and (iv) of Lemma 11.1.13 hold. By points

(i) and (ii) we can use Fatou Lemma and the dominated convergence Theorem to conclude that

$$\liminf_j \int_{B(0,1/\varepsilon)} \phi_{n(j)} d\hat{\nu}_\varepsilon \geqslant \int_{B(0,1/\varepsilon)} \phi d\hat{\nu}_\varepsilon, \quad \lim_j \int_{\mathbb{R}^d} \phi^*_{n(j)} d\mu_\varepsilon = \int_{\mathbb{R}^d} \phi^* d\mu_\varepsilon.$$

We thus deduce that

$$\begin{aligned} \int_{\mathbb{R}^d} \left(\frac{1}{2}|x|^2 - \phi^*(x)\right) d\mu_\varepsilon(x) + \int_{B(0,1/\varepsilon)} \left(\frac{1}{2}|y|^2 - \phi(y)\right) d\hat{\nu}_\varepsilon(y) \\ \geqslant \limsup_j \int_{\mathbb{R}^d} u_{h_{n(j)}} d\mu_\varepsilon \\ = F^\varepsilon_\nu(\mu) \end{aligned}$$

proving that

$$\int_{B(0,1/\varepsilon)} \phi d\hat{\nu}_\varepsilon + \int_{\mathbb{R}^d} \phi^* d\mu_\varepsilon = \frac{1}{2} m_2^2(\hat{\nu}_\varepsilon) + \frac{1}{2} m_2^2(\mu_\varepsilon) - \frac{1}{2} W_2^2(\hat{\nu}_\varepsilon, \mu_\varepsilon).$$

By the uniqueness part of Theorem 11.1.1 and point (iii) in Lemma 11.1.13, we deduce that $\phi = \Phi(\hat{\nu}_\varepsilon, \mu_\varepsilon)$ and $\phi^* = \Phi^*(\hat{\nu}_\varepsilon, \mu_\varepsilon)$. Finally, the a.e. convergence of the gradient of ϕ^*_n to the gradient of ϕ^* given by point (iv) in Lemma 11.1.13 gives that $\nabla(\phi^*_{n(j)} * \kappa_\varepsilon) \to \nabla(\phi^* * \kappa_\varepsilon)$ pointwise everywhere. Moreover, since for every $x \in \mathbb{R}^d$ we have

$$\left| x - \nabla(\varphi^*_{n(j)} * \kappa_\varepsilon)(x) \right|^2 \leqslant |x|^2 + \varepsilon^2 + 1/\varepsilon^2 + 2 \in L^1(\mathbb{R}^d, \mu),$$

we can use the dominated convergence Theorem to conclude that

$$\lim_j \int_{\mathbb{R}^d} \left| x - \nabla(\phi^*_{n(j)} * \kappa_\varepsilon)(x) \right|^2 d\mu(x) = \int_{\mathbb{R}^d} |x - \nabla(\phi^* * \kappa_\varepsilon)(x)|^2 d\mu(x).$$

This concludes the proof of the fourth claim.

Claim 5. It holds

$$\limsup_k |DH_k|_{\star,\mathscr{A}}(\mu) \leqslant \zeta'(F^\varepsilon_\nu(\mu)) \left(\int_{\mathbb{R}^d} |x - \nabla(\varphi^*_\varepsilon * \kappa_\varepsilon)(x)|^2 d\mu(x) \right)^{1/2}$$

for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, where $\varphi^*_\varepsilon = \Phi^*(\hat{\nu}_\varepsilon, \mu_\varepsilon)$.
Proof of claim 5. Let $B \subset \mathcal{P}_2(\mathbb{R}^d)$ be defined as

$$B := \bigcap_k \bigcup_{h=1}^k A^k_h,$$

where A^k_h is the full $\mathfrak{m}$-measure subset of B^k_h where claim 2 holds. Notice that B has full $\mathfrak{m}$-measure. Let $\mu \in B$ be fixed and let us pick an increasing sequence $k \mapsto h_k$ such that

$$G_k(\mu) = \int_{\mathbb{R}^d} u_{h_k} d\mu_\varepsilon.$$

By claim 1 we know that $G_n(\mu) \to F_\nu^\varepsilon(\mu)$ so that we can apply claim 4 and conclude that

$$\zeta'(F_\nu^\varepsilon(\mu)) \int_{\mathbb{R}^d} |x - \nabla(\varphi_\varepsilon^* * \kappa_\varepsilon)(x)|^2 d\mu(x) = \lim_k \zeta'(G_k(\mu)) \int_{\mathbb{R}^d} \left|x - \nabla(\varphi_{h_k}^* * \kappa_\varepsilon)(x)\right|^2 d\mu(x).$$

By claim 2, the right hand side is greater than $\limsup_k |DH_k|^2_{\star,\mathscr{A}}(\mu)$; this concludes the proof of the fifth claim.

Eventually, we observe that by Claim 1

$$W_\nu^\varepsilon(\mu) = \lim_{k\to\infty} \zeta(G_k(\mu)) = \lim_{k\to\infty} H_k(\mu). \tag{11.1.47}$$

Moreover, passing to the limit the estimate in Claim 3, we see that

$$\zeta'(F_\nu^\varepsilon(\mu)) \int_{\mathbb{R}^d} |x - \nabla(\varphi_\varepsilon^* * \kappa_\varepsilon)(x)|^2 \, d\mu(x)$$

is uniformly bounded. We can then combine the expression of Claim 2, the uniform estimate of Claim 3, the limit of Claim 5 with Lemma 11.1.16 to get (11.1.41). □

Corollary 11.1.20. *Let* $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. *Then*

$$|DW_\nu|_{\star,\mathscr{A}}(\mu) \leqslant 1 \quad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.1.48}$$

Proof. First of all we prove that

$$\int_{\mathbb{R}^d} |x - \nabla(\varphi_\varepsilon^* * \kappa_\varepsilon)(x)|^2 \, d\mu(x) \leqslant W_2^2(\mu_\varepsilon, \hat{\nu}_\varepsilon) \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.1.49}$$

Since

$$|x - \nabla(\varphi_\varepsilon^* * \kappa_\varepsilon)(x)|^2 \leqslant |x - \nabla\varphi_\varepsilon^*(x)|^2 * \kappa_\varepsilon(x) \quad \text{for every } x \in \mathbb{R}^d,$$

we get, also using Proposition 11.1.19, that

$$\begin{aligned} \int_{\mathbb{R}^d} |x - \nabla(\varphi_\varepsilon^* * \kappa_\varepsilon)(x)|^2 \, d\mu(x) &\leqslant \int_{\mathbb{R}^d} \left(|x - \nabla\varphi_\varepsilon^*(x)|^2 * \kappa_\varepsilon(x)\right) d\mu(x) \\ &= \int_{\mathbb{R}^d} |x - \nabla\varphi_\varepsilon^*(x)|^2 \, d\mu_\varepsilon(x) \\ &= W_2^2(\mu_\varepsilon, \hat{\nu}_\varepsilon), \end{aligned}$$

for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, where the last equality comes from Theorem 11.1.1. This proves (11.1.49). It then follows that

$$|D(\zeta \circ F_\nu^\varepsilon)|_{\star,\mathscr{A}}(\mu) \leqslant \zeta'(F_\nu^\varepsilon(\mu))\sqrt{2F_\nu^\varepsilon(\mu)}$$

Setting $\vartheta(r) = \zeta(\frac{1}{2}r^2)$ so that $\vartheta'(r) = r\zeta'(\frac{1}{2}r^2)$ and $\zeta \circ F_\nu^\varepsilon = \vartheta(W_\nu^\varepsilon)$ we get

$$|D(\vartheta \circ W_\nu^\varepsilon)|_{\star,\mathscr{A}}(\mu) \leqslant \zeta'(\frac{1}{2}(W_\nu^\varepsilon)^2(\mu))W_\nu^\varepsilon(\mu) = \vartheta'(W_\nu^\varepsilon)(\mu). \tag{11.1.50}$$

Let now $\zeta_n : \mathbb{R} \to \mathbb{R}$ be a sequence of C^1 functions having derivatives with compact support such that

$$\zeta_n(t) \to u(t), \quad \zeta_n'(t) \to u'(t), \quad 0 \leqslant \zeta_n'(t) \leqslant 2 \text{ for every } t \in \mathbb{R},$$

where $u : \mathbb{R} \to \mathbb{R}$ is equal to $\arctan(t)$ if $t \geqslant 0$ and equal to 0 if $t \leqslant 0$. Then, setting $\theta_n(t) := \zeta_n(\frac{1}{2}t^2)$, we get by (11.1.50) that

$$|D(\vartheta_n \circ W_\nu^\varepsilon)|_{\star,\mathscr{A}}(\mu) \leqslant \vartheta_n'(W_\nu^\varepsilon)(\mu). \tag{11.1.51}$$

Passing to the limit as $n \to +\infty$ (11.1.51) using Theorem 10.1.2(1)-(3), we obtain

$$|D(\nu \circ W_\nu^\varepsilon)|_{\star,\mathscr{A}}(\mu) \leqslant \nu'(W_\nu^\varepsilon)(\mu),$$

where $\nu : \mathbb{R} \to \mathbb{R}$ is equal to $\arctan(\frac{1}{2}t^2)$ if $t \geqslant 0$ and equal to 0 if $t \leqslant 0$. We thus conclude that

$$|DW_\nu^\varepsilon|_{\star,\mathscr{A}} \leqslant 1 \tag{11.1.52}$$

by Lemma 10.1.9. Choosing $\varepsilon = 1/k$, we have $\lim_{k\to+\infty} W_\nu^{1/k}(\mu) = W_\nu(\mu)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$; using Theorem 10.1.2 (1)-(3), we obtain (11.1.48). □

The **proof of Theorem 11.1.11** then easily follows by Corollary 11.1.20 and Theorem 10.2.1.

We conclude this section with a simple but useful density property, which shows the possibility to use smaller algebras of cylindrical functions to operate in $H^{1,2}(\mathbb{W}_2)$.

Proposition 11.1.21. *Let $\mathscr{F}$ be a subset of $C_b^1(\mathbb{R}^d)$ satisfying the following property: for every $f \in C_b^1(\mathbb{R}^d)$ there exists a sequence $f_n \in \mathscr{F}$, $n \in \mathbb{N}$, such that*

$$\sup_{\mathbb{R}^d} |f_n| + |Df_n| < \infty, \quad \lim_{n\to\infty} \int_{\mathbb{R}^d} |f_n - f| + |\nabla(f_n - f)|\, d\mu = 0 \quad \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.1.53}$$

Then the algebra $\mathscr{A} \subset \mathrm{Cyl}\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ generated by the set of cylindrical functions $\big\{L_f : f \in \mathscr{F}\big\}$ is dense in $H^{1,2}(\mathbb{W}_2)$ and satisfies the strong approximation property of Theorem 11.1.11.

In particular the algebra $\mathrm{Cyl}^\infty(\mathcal{P}_2(\mathbb{R}^d))$ generated by $\big\{L_f : f \in C_c^\infty(\mathbb{R}^d)\big\}$ is strongly dense in $H^{1,2}(\mathbb{W}_2)$ and satisfies the approximation property of Theorem 11.1.11.

Proof. Thanks to Theorem 11.1.11 and a simple diagonal argument, it is sufficient to prove that for every cylindrical function $F \in \mathrm{Cyl}\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ there exists a sequence $F_n \in \mathscr{A}$ such that

$$F_n \to F \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \quad \text{and} \quad \mathrm{pCE}_2(F_n - F) \to 0 \quad \text{as } n \to \infty. \tag{11.1.54}$$

In the case $F = L_f$ with $f \in C_b^1(\mathbb{R}^d)$, (11.1.53) and Lebesgue Dominated Convergence Theorem show that we can find a sequence $f_n \in \mathscr{F}$ such that, setting $F_n := L_{f_n}$, we have

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} |F_n - F|^2 \, d\mathfrak{m} = \int_{\mathcal{P}_2(\mathbb{R}^d)} \Big| \int_{\mathbb{R}^d} (f_n(x) - f(x))\, d\mu(x) \Big|^2 d\mathfrak{m}(\mu) \to 0,$$

$$\mathrm{pCE}_2(F_n - F) = \int_{\mathcal{P}_2(\mathbb{R}^d)} \int |\nabla f_n(x) - \nabla f(x)|^2 \, d\mu(x)\, d\mathfrak{m}(x) \to 0$$

as $n \to +\infty$. Let us now consider a general $F = \psi \circ L_f$ as in (11.1.5), where $f = (f_1, \cdots, f_N)$ is a vector of functions in $C^1_b(\mathbb{R}^d)$ and $\psi \in C^1_b(\mathbb{R}^N)$. If we consider $\tilde{f} := (1, f_1, \ldots, f_N)$ and $\tilde{\psi} \in C^1_b(\mathbb{R}^{N+1})$ defined as

$$\tilde{\psi}(x_0, x_1, \ldots, x_n) := \psi(0)x_0 - \psi(0) + \psi(x_1, x_2, \ldots, x_N), \quad (x_0, x_1, \ldots, x_N) \in \mathbb{R}^{N+1},$$

we have that $\tilde{\psi}(0) = 0$ and $\tilde{\psi} \circ L_{\tilde{f}} = F$. For this reason we can always suppose that $f_1 \equiv 1$ and $\psi(0) = 0$. It is also not restrictive to assume that ψ is a polynomial with $\psi(0) = 0$: in fact, setting $R := \sup_{\mathbb{R}^d, 1 \leqslant k \leqslant N} \Big(|f_k| + |\nabla f_k|\Big)$, we can find a sequence of polyonomials $(P_h)_h$ in $\mathbb{R}^N$ such that

$$P_h(0) = 0, \quad \sup_{|z| \leqslant R} |P_h(z) - \psi(z)| + |\nabla P_h(z) - \nabla \psi(z)| \to 0 \quad \text{as } h \to \infty. \tag{11.1.55}$$

It follows that $F_h := P_h \circ L_f$ satisfies

$$\lim_{h \to \infty} \sup_{\mathcal{P}_2(\mathbb{R}^d)} \Big(|F_h(\mu) - F(\mu)| + \|DF_h[\mu] - DF[\mu]\|_\mu\Big) = 0. \tag{11.1.56}$$

Applying (11.1.53) we can find sequences $(f_{k,n})_{n \in \mathbb{N}}$, $k = 1, \cdots, N$, approximating f_k as in (11.1.53). In particular, there exists $R > 0$ such that $\sup_{\mathbb{R}^d} \Big(|f_{k,n}| + |Df_{k,n}| + |f_k| + |Df_k|\Big) \leqslant R$ for every $n \in \mathbb{N}$, $k \in \{1, \cdots, N\}$. If ψ is a polynomial in $\mathbb{R}^N$ with $\psi(0) = 0$ then the function $F_n := \psi \circ L_{f_n}$ belongs to $\mathscr{A}$, where $f_n = (f_{1,n}, f_{2,n}, \ldots, f_{N,n})$. Denoting by L the maximum of the Lipschitz constants of ψ and $\partial_k \psi$ in the cube $[-R, R]^N$ with respect to the ∞-norm, it is easy to see that

$$\begin{aligned}
|F_n(\mu) - F(\mu)| &= \Big|\psi(L_{f_n}(\mu))\psi(L_{f_n}(\mu)))\Big| \leqslant L \sup_k |L_{f_{k,n}}(\mu) - L_{f_k}(\mu)| \to 0, \\
\|DF_n[\mu] - DF[\mu]\|_\mu &= \Big\| \sum_k \Big(\partial_k \psi(L_{f_n}(\mu)) \nabla f_{k,n} - \partial_k \psi(L_f(\mu)) \nabla f_k\Big) \Big\|_\mu \\
&\leqslant \sum_k \Big\| \partial_k \psi(L_{f_n}(\mu)) \nabla f_{k,n} - \partial_k \psi(L_{f_n}(\mu)) \nabla f_k \Big\|_\mu \\
&\quad + \sum_k \Big\| \Big(\partial_k \psi(L_{f_n}(\mu)) - \partial_k \psi(L_f(\mu))\Big) \nabla f_k \Big\|_\mu \\
&\leqslant L \sum_k \Big(\Big\| \nabla f_{k,n} - \nabla f_k \Big\|_\mu + R \Big| \langle f_{k,n} - f_k, \mu \rangle \Big| \Big).
\end{aligned}$$

Both terms are uniformly bounded w.r.t. μ and n, and converge to 0 as $n \to \infty$. We deduce that (11.1.54) holds. □

11.2 CALCULUS RULES AND EXAMPLES

Let us now show how we can give a more precise description of CE_2 and to establish useful calculus rules.

Theorem 11.2.1 ($\mathfrak{m}$-Wasserstein differential)**.** *For every $F \in D^{1,2}(\mathbb{W}_2)$ there exists a unique vector field $D_\mathfrak{m} F \in L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$ (the $\mathfrak{m}$-Wassesrstein differential* *of* F*) such that for every sequence* $F_n \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$, $n \in \mathbb{N}$, *satisfying* (11.1.18) *we have*

$$\mathrm{D}F_n \to \mathrm{D_m}F \quad \textit{strongly in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d). \tag{11.2.1}$$

Moreover:

(a) *The map* $F \mapsto \mathrm{D_m}F$ *from* $D^{1,2}(\mathbb{W}_2)$ *to* $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d)$ *is linear and for every* $F, G \in D^{1,2}(\mathbb{W}_2)$ *we have*

$$\mathsf{CE}_2(F,G) = \int \mathrm{D_m}F(\mu,x) \cdot \mathrm{D_m}G(\mu,x)\, \mathrm{d}\mathbf{m}(\mu,x), \quad \mathsf{CE}_2(F) = \int |\mathrm{D_m}F(\mu,x)|^2\, \mathrm{d}\mathbf{m}(\mu,x), \tag{11.2.2}$$

where $\mathsf{CE}_2(\cdot,\cdot)$ *denotes the quadratic form associated to* $\mathsf{CE}_2(\cdot)$ *as in Remark* 10.2.5.

(b) *The map* $F \mapsto (F, \mathrm{D_m}F)$ *is a linear isometric (thus continuous) immersion of* $H^{1,2}(\mathbb{W}_2)$ *into* $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \times L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d)$.

(c) *The graph of* $\mathrm{D_m}$ *in* $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \times L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d)$ *is (weakly) closed: for every sequence* $F_n \in H^{1,2}(\mathbb{W}_2)$

$$\left.\begin{aligned} F_n &\rightharpoonup F \textit{ in } L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \\ \mathrm{D_m}F_n &\rightharpoonup \mathbf{G} \textit{ in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d) \end{aligned}\right\} \quad \Rightarrow \quad F \in H^{1,2}(\mathbb{W}_2),\ \mathbf{G} = \mathrm{D_m}F. \tag{11.2.3}$$

Proof. The proof uses well known arguments of the theory of quadratic forms. If F_n, $n \in \mathbb{N}$, is a sequence in $\mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ for every $m, n \in \mathbb{N}$ we have

$$\frac{1}{4}\mathsf{pCE}_2(F_m - F_n) = \frac{1}{2}\Big(\mathsf{pCE}_2(F_m) + \mathsf{pCE}(F_n)\Big) - \mathsf{pCE}_2\Big(\frac{1}{2}(F_m + F_n)\Big). \tag{11.2.4}$$

If (11.1.18) holds, observing that $\lim_{m,n\to\infty} \frac{1}{2}(F_m + F_n) = F$, we can pass to the limit as $m, n \to \infty$ and therefore by (10.1.17) $\liminf_{m,n\to\infty} \mathsf{pCE}_2\left(\frac{1}{2}(F_m + F_n)\right) \geqslant \mathsf{CE}_2(F)$; we thus obtain

$$\limsup_{m,n\to\infty} \frac{1}{4}\mathsf{pCE}_2(F_m - F_n) = \limsup_{m,n\to\infty} \frac{1}{4}\int |\mathrm{D}F_m(\mu,x) - \mathrm{D}F_n(\mu,x)|^2\, \mathrm{d}\mathbf{m}(\mu,x) \leqslant 0 \tag{11.2.5}$$

which shows that $n \mapsto \mathrm{D}F_m$ is a Cauchy sequence in $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d)$ and therefore converges to some element $\mathbf{V}$.

If $\tilde{F}_n$ is another sequence satisfying (11.1.18), we can use the identity

$$\frac{1}{4}\mathsf{pCE}_2(F_n - \tilde{F}_n) = \frac{1}{2}\Big(\mathsf{pCE}_2(F_n) + \mathsf{pCE}(\tilde{F}_n)\Big) - \mathsf{pCE}_2\Big(\frac{1}{2}(F_n + \tilde{F}_n)\Big) \tag{11.2.6}$$

and the same argument to conclude that $\lim_{n\to\infty} \mathsf{pCE}_2(F_n - \tilde{F}_n) = 0$, so that the limit $\mathbf{V}$ is independent of the approximating sequence and we are authorized to call it $\mathrm{D_m}F$.

Concering claim (a), the linearity of D_m follows immediately from the linearity of D as a map from $\mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ to $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d)$.

If $F, G \in D^{1,2}(\mathbb{W}_2)$ and $(F_n)_n, (G_n)_n \subset \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ are sequences satisfying (11.1.18) for F and G respectively, we can see that $\mathrm{pCE}_2(F_n, G_n) \to \mathrm{CE}_2(F, G)$; indeed

$$\begin{aligned}\mathrm{pCE}_2(F_n, G_n) &= \frac{1}{2}\mathrm{pCE}_2(F_n + G_n) - \frac{1}{2}\mathrm{pCE}_2(F_n) - \frac{1}{2}\mathrm{pCE}_2(G_n),\\ &= -\frac{1}{2}\mathrm{pCE}_2(F_n - G_n) + \frac{1}{2}\mathrm{pCE}_2(F_n) + \frac{1}{2}\mathrm{pCE}_2(G_n).\end{aligned}$$

Passing the first equality to the $\liminf_n$, the second one to the $\limsup_n$ and using (10.1.17), we get that $\mathrm{pCE}_2(F_n, G_n) \to \mathrm{CE}_2(F, G)$. Passing then to the limit in (11.2.2) we immediately see that

$$\mathrm{CE}_2(F, G) = \int D_m F(\mu, x) \cdot D_m G(\mu, x) \, d\mathbf{m}(\mu, x) \tag{11.2.7}$$

which, together with (10.2.25), shows that $F \mapsto (F, D_m F)$ is an isometry from $H^{1,2}(\mathbb{W}_2)$ into $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \times L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d)$ (claim (b)).

Claim (c) then follows by claim (b) and the fact that $H^{1,2}(\mathbb{W}_2)$ is an Hilbert space. □

Let us now collect a few properties of $D_m F$, which follow by the corresponding metric versions of Theorem 10.1.2 and the approximation property of Theorem 11.2.1.

Proposition 11.2.2 (Calculus properties of $D_m F$)**.** *The* $\mathfrak{m}$*-Wasserstein differential satisfies the following properties:*

(1) (Relaxed gradient and asymptotic Lipschitz constant) For every $F \in D^{1,2}(\mathbb{W}_2)$ *we have*

$$\|D_m F[\mu]\|_\mu^2 = \int |D_m F(\mu, x)|^2 \, d\mu(x) = |DF|_\star^2(\mu) \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.2.8}$$

In particular, for every $F \in \mathrm{Lip}_b(\mathcal{P}_2(\mathbb{R}^d))$

$$\|D_m F[\mu]\|_\mu^2 = \int |D_m F(\mu, x)|^2 \, d\mu(x) \leqslant \left|\mathrm{lip} F(\mu)\right|^2 \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d), \tag{11.2.9}$$

and if $F \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$

$$\int |D_m F(\mu, x)|^2 \, d\mu(x) \leqslant \int |DF(\mu, x)|^2 \, d\mu(x) \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.2.10}$$

(2) (Leibniz rule) If $F, G \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \cap D^{1,2}(\mathbb{W}_2)$*, then* $H := FG \in D^{1,2}(\mathbb{W}_2)$ *and*

$$D_m H(\mu, x) = F(\mu) D_m G(\mu, x) + G(\mu) D_m F(\mu, x) \quad \text{for } \mathbf{m}\text{-a.e. } (\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d. \tag{11.2.11}$$

(3) *(Locality) If* $F \in D^{1,2}(W_2)$ *then for any* $\mathscr{L}^1$*-negligible Borel subset* $N \subset \mathbb{R}$ *we have*

$$D_m F[\mu] = 0 \quad \textit{in } L^2(\mathbb{R}^d, \mu; \mathbb{R}^d) \quad m\textit{-a.e. on } F^{-1}(N). \tag{11.2.12}$$

(4) *(Truncations) If* $F_j \in D^{1,2}(W_2)$*,* $j = 1, \cdots, J$*, then also the functions*

$$F_+ := \max(F_1, \cdots, F_J) \textit{ and } F_- := \min(F_1, \cdots, F_J)$$

belong to $D^{1,2}(W_2)$ *and*

$$D_m F_+ = D_m F_j \quad \mathbf{m}\textit{-a.e. on } \{(\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d : F_+(\mu) = F_j(\mu)\}, \tag{11.2.13}$$

$$D_m F_- = D_m F_j \quad \mathbf{m}\textit{-a.e. on } \{(\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d : F_-(\mu) = F_j(\mu)\}. \tag{11.2.14}$$

(5) *(Chain rule) If* $F \in D^{1,2}(W_2)$ *and* $\phi \in \mathrm{Lip}(\mathbb{R})$ *then* $\phi \circ F \in D^{1,2}(W_2)$ *and*

$$D_m(\phi \circ F) = \phi'(F)\, D_m F \quad \mathbf{m}\textit{-a.e. on } \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d. \tag{11.2.15}$$

Remark 11.2.3. Notice that the product in (11.2.15) is well defined since there exists a $\mathscr{L}^1$-negligible Borel set $N \subset \mathbb{R}$ such that ϕ is differentiable in $\mathbb{R} \setminus N$ and $D_m F$ vanishes m-a.e. in $F^{-1}(N)$ thanks to the locality property (11.2.12).

Proof. Claim (a) is an immediate consequence of the fact that (11.1.18) yields $\mathrm{lip} F_n \to |DF|_\star$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), m)$; up to extracting a suitable (not relabeled) subsequence we get $\int |DF_n|^2 \, d\mu \to |DF|_\star^2(\mu)$ for m-a.e. μ. On the other hand (11.2.1) yields

$$\int \Big| |DF_n(\mu, x)|^2 - |D_m F(\mu, x)|^2 \Big| \, d\mathbf{m}(\mu, x) \to 0 \quad \text{as } n \to \infty \tag{11.2.16}$$

so that Fubini's Theorem yields, up to extracting a suitable subsequence,

$$\int |DF_n(\mu, x)|^2 \, d\mu \to \int |D_m F(\mu, x)|^2 \, d\mu \quad \text{for } m\text{-a.e.} \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.2.17}$$

(11.2.9) and (11.2.10) then follows by the general properties of the minimal relaxed gradients.

Claim (c) follows by (10.1.9) and (11.2.8).

Claim (d) is just a consequence of the locality property (11.2.12).

Claim (e) is true if $\phi \in C^1_b(\mathbb{R})$ just by passing to the limit in the corresponding formula for a cylindrical function. In fact if $F_n \in \mathrm{Cyl}\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ is a sequence as in (11.1.18) and (11.2.1) we have

$$D(\phi \circ F_n) = (\phi' \circ F_n) DF_n \quad \text{in } \mathcal{D}. \tag{11.2.18}$$

Since ϕ' is bounded and continuous we get

$$D(\phi \circ F_n) \to \mathbf{G} = (\phi' \circ F) D_m F \quad \text{strongly in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d) \quad \text{as } n \to \infty.$$

(11.2.19)

Integrating w.r.t. $\mathbf{m}$ and recalling (11.2.8) and Theorem 10.1.2(7) we get

$$\begin{aligned}\int |\mathbf{G}|^2 \, d\mathbf{m} &= \int |\phi'(F(\mu))|^2 |D_m F(\mu,x)|^2 \, d\mathbf{m}(\mu,x) \\ &= \int |\phi'(F(\mu))| \, |DF|_\star^2 \, dm \\ &= CE_2(\phi \circ F)\end{aligned}$$

so that

$$\lim_{n\to\infty} pCE_2(\phi \circ F_n) = CE_2(\phi \circ F).$$

We conclude by Theorem 11.2.1 that $\mathbf{G} = (\phi' \circ F) D_m F$ coincides with $D_m(\phi \circ F)$.

Let us now consider the case of a general Lipschitz function ϕ; by truncation and Claim (d) it is not restrictive to assume that ϕ is also bounded. We can find a sequence $\phi_n \in C^1_b(\mathbb{R})$ such that $\sup_{\mathbb{R}} |\phi_n| + |\phi_n'| \leqslant L < \infty$, $\phi_n \to \phi$ uniformly, and $\phi_n'(x) \to \phi'(x)$ for every $x \in \mathbb{R} \setminus N$ for a Borel set N with $\mathscr{L}^1(N) = 0$. We have

$$D_m(\phi_n \circ F) = \phi_n'(F) D_m F \quad m\text{-a.e. in } \mathcal{P}_2(\mathbb{R}^d). \tag{11.2.20}$$

Setting $\tilde N := \{(\mu,x) \in \overline{\mathcal{D}} : F(\mu) \in N\}$, Fubini's Theorem and the locality property (11.2.12) yields $D_m F(\mu,x) = 0$ for $\mathbf{m}$-a.e. $(\mu,x) \in \tilde N$. On the other hand $\phi_n'(F(\mu)) \to \phi'(F(\mu))$ for every $(\mu,x) \in \mathcal{D} \setminus \tilde N$; since ϕ_n' is uniformly bounded, we deduce that

$$\phi_n'(F) D_m F \to \phi'(f) D_m F \quad \text{strongly in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d). \tag{11.2.21}$$

We conclude by Theorem 11.2.1(b) that $D_m(\phi \circ F) = \phi'(F) D_m F$.

Claim (b) follows by claim (e); indeed, since $F, G \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), m)$, we can find a constant $M > 0$ such that

$$|F|(\mu) \leqslant M, \quad |G|(\mu) \leqslant M, \quad |F+G|(\mu) \leqslant M \quad \text{for } m\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Let $\phi \in \mathrm{Lip}(\mathbb{R})$ be such that $\phi(x) = x^2$ for every $x \in [-M-1, M+1]$; then we have

$$\begin{aligned} D_m FG &= \frac12 D_m((F+G)^2) - \frac12 D_m(F^2) - \frac12 D(G^2) \\ &= \frac12 D_m(\phi\circ(F+G)) - \frac12 D_m(\phi\circ F) - \frac12 D(\phi \circ G) \\ &= \frac12 \phi'(F+G) D_m(F+G) - \frac12 \phi'(F) D_m F - \frac12 \phi'(G) DG \\ &= (F+G) D_m(F+G) - F D_m F - G D G \\ &= F D_m G + G D_m F \end{aligned}$$

for $\mathbf{m}$-a.e. $(\mu,x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$.

□

Corollary 11.2.4. CE_2 *is a local Dirichlet form in* $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ *[22, p. 3.1.1] enjoying* Γ*-calculus with Carré du champs* Γ *given by*

$$\Gamma(F, G)[\mu] := \int D_{\mathfrak{m}} F(\mu, x) \cdot D_{\mathfrak{m}} G(\mu, x) \, d\mu(x) \quad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.2.22}$$

In particular, for every $F, G \in H^{1,2}(\mathbb{W}_2)$ *we have*

$$\begin{aligned} \mathsf{CE}_2(F, G) &= \int_{\mathcal{P}_2(\mathbb{R}^d)} \Gamma(F, G)[\mu] \, d\mathfrak{m}(\mu) = \int D_{\mathfrak{m}} F(\mu, x) \cdot D_{\mathfrak{m}} G(\mu, x) \, d\boldsymbol{\mathfrak{m}}(\mu, x), \\ \mathsf{CE}_2(F) &= \int_{\mathcal{P}_2(\mathbb{R}^d)} \Gamma(F, F) \, d\mathfrak{m}(\mu) = \int |D_{\mathfrak{m}} F(\mu, x)|^2 \, d\boldsymbol{\mathfrak{m}}(\mu, x). \end{aligned} \tag{11.2.23}$$

Proof. The fact that CE_2 is a Dirichlet form follows by the truncation property (11.2.15) with $\phi(r) := r \wedge 1$. Since $\mathsf{CE}_2(1) = 0$, the same property with $\phi(r) = |r|$ also shows that CE_2 is local (see [22, Corollary 5.1.4]).

Using the Leibniz rule (11.2.11) one can also easily show that the Γ-operator (11.2.22) is the Carré du champ associated to $\frac{1}{2}\mathsf{CE}_2$ [22, Definition 4.1.2]. □

11.2.1 *Tangent bundle, residual differentials and relaxation*

In general $\mathsf{CE}_2(F)$ doesn't coincide with $\mathsf{pCE}_2(F)$ if $F \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$, or, equivalently, $D_{\mathfrak{m}} F$ is not equal to DF: this equality corresponds to the closability of pCE_2. We can however investigate the relations between DF and $D_{\mathfrak{m}} F$: two useful tools are represented by the closure of the graph of D and by the collection of all the weak limits of Wasserstein differentials along vanishing sequences.

Definition 11.2.5 (Multivaled gradient)**.** We denote by $\mathbf{G} \subset L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \times L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ the closure of the space $\left\{(F, DF) : F \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)\right\}$. The multivalued gradient $\mathbf{D}_{\mathfrak{m}} : H^{1,2}(\mathbb{W}_2) \rightrightarrows L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ is the operator whose graph is $\mathbf{G}$.

It is clear that $\mathbf{G}$ is a closed vector space of $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \times L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$, which can also be obtained as the weak closure of $\left\{(F, DF) : F \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)\right\}$. Thus $\mathbf{V} \in \mathbf{D}_{\mathfrak{m}} F$ if and only if there exists a sequence $F_n \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ such that

$$F_n \to F \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}), \quad DF_n \rightharpoonup \mathbf{V} \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d). \tag{11.2.24}$$

The set $\mathbf{D}_{\mathfrak{m}} 0$ plays a crucial role.

Definition 11.2.6 (Residual gradients)**.** The set of residual gradients

$$G_0 \subset L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$$

is defined as

$$\begin{aligned} G_0 := \Big\{ \mathbf{V} \in L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d) : \text{ there exists } (F_n)_{n \in \mathbb{N}} \subset \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right) : \\ F_n \to 0 \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}), \ DF_n \rightharpoonup \mathbf{V} \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d) \Big\}. \end{aligned} \tag{11.2.25}$$

Notice that pCE_2 is closable if and only if G_0 is trivial. A third important space is the L^2 tangent bundle of $\mathcal{P}_2(\mathbb{R}^d)$. In the following, given a Borel map $\mathbf{G} \in \mathcal{L}^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$, we denote, for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, by $\mathbf{G}[\mu]$ the map $x \mapsto \mathbf{G}(\mu, x)$.

Definition 11.2.7. We denote by $\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ the subspace of $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$ of vector fields $\mathbf{V}$ satisfying

$$\mathbf{V}[\mu] \in \mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.2.26}$$

Lemma 11.2.8. $\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ *is a closed subspace of* $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$ *which is a* $L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ *module:*

$$\text{for every } \mathbf{V} \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}),\ H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}): \quad H\mathbf{V} \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}). \tag{11.2.27}$$

For every function $H \in H^{1,2}(\mathbb{W}_2)$ *(resp.* $F \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$*) we have that* $D_\mathfrak{m} F \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ *(resp.* $DF \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$*). Finally, if* $\mathscr{C} \subset C^\infty_c(\mathbb{R}^d)$ *is a countable set dense in* $C^\infty_c(\mathbb{R}^d)$ *with respect to the Lipschitz norm* $\|\zeta\|_{\mathrm{Lip}} := \sup_{\mathbb{R}^d} |\zeta| + |\nabla\zeta|$ *and* $\mathscr{L}$ *is a countable set dense in* $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ *then the set*

$$\mathscr{T} = \mathrm{span}\left\{ H\nabla\zeta : H \in \mathscr{L},\ \zeta \in \mathscr{C} \right\} \quad \textit{is dense in } \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}). \tag{11.2.28}$$

Proof. Let $(\mathbf{V}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ strongly converging to $\mathbf{V}$ in L^2; it is not restrictive to assume that $\mathbf{V}_n$ are Borel maps satisfying $\mathbf{V}_n[\mu] \in \mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}$ for a $\mathfrak{m}$-negligible set of $\mathcal{P}_2(\mathbb{R}^d)$. Up to extracting a suitable subsequence, we can also assume that $\sum_{n=1}^\infty \|\mathbf{V}_n - \mathbf{V}\|^2_{L^2} < \infty$. Applying Fubini's Theorem it follows that

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\sum_{n=1}^\infty \int_{\mathbb{R}^d} |\mathbf{V}_n[\mu](x) - \mathbf{V}[\mu](x)|^2 \,\mathrm{d}\mu(x) \Big) \,\mathrm{d}\mathfrak{m} < +\infty$$

so that there exists a $\mathfrak{m}$-negligible set $\mathcal{N}' \supset \mathcal{N}$ such that

$$\sum_{n=1}^\infty \int_{\mathbb{R}^d} |\mathbf{V}_n[\mu](x) - \mathbf{V}[\mu](x)|^2 \,\mathrm{d}\mu(x) < \infty \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}';$$

and this implies that $\mathbf{V}_n[\mu] \to \mathbf{V}[\mu]$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mu)$, so that $\mathbf{V}[\mu] \in \mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}'$.

(11.2.27) is obvious. Since for every $F = L_\phi$, $\phi \in C^1_b$ $DF[\mu] = \nabla\phi \in \mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, it is immediate to check that $DF \in \mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ for every cylindrical function. The closure property of $\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ then yields the analogous conclusion for the Wasserstein differential of $D_\mathfrak{m} F$ of a Sobolef function $F \in H^{1,2}(\mathbb{W}_2)$.

Let us eventually consider (11.2.28): it is sufficient to prove that any $\mathbf{V} \in \mathscr{T}^\perp$ belongs to $\left(\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})\right)^\perp$. If $\mathbf{V} \in \mathscr{T}^\perp$ is a Borel vector field, then

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int \langle \nabla\zeta, \mathbf{V}(\mu, x) \rangle \,\mathrm{d}\mu(x) \Big) H(\mu) \,\mathrm{d}\mathfrak{m}(\mu) = 0$$

for every $\zeta \in \mathscr{C}$, $H \in \mathscr{L}$. Since $\mathscr{L}$ is dense in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ we have for every $\zeta \in \mathscr{C}$

$$\int \langle \nabla\zeta, \mathbf{V}(\mu, x)\rangle \, d\mu(x) = 0 \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d)$$

Since $\mathscr{C}$ is countable, we can find a $\mathfrak{m}$-negligible set $\mathcal{N} \subset \mathcal{P}_2(\mathbb{R}^d)$ such that

$$\int \langle \nabla\zeta, \mathbf{V}(\mu, x)\rangle \, d\mu(x) = 0 \quad \text{for every } \zeta \in \mathscr{C} \text{ and every } \mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}$$

which shows that $\mathbf{V}[\mu] \in \Big(\mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)\Big)^\perp$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}$, so that for every $\boldsymbol{W} \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}})$

$$\begin{aligned}\int \langle \mathbf{V}(\mu, x), \boldsymbol{W}(\mu, x)\rangle \, d\boldsymbol{\mathfrak{m}} &= \int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} \langle \mathbf{V}[\mu](x), \boldsymbol{W}[\mu](x)\rangle \, d\mu(x) \Big) \, d\mathfrak{m}(\mu) \\ &= 0.\end{aligned}$$

□

Let us collect a few simple properties of G_0.

Lemma 11.2.9. *Let* G_0 *be as in* (11.2.25).

(1) G_0 *is a closed subspace of* $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ *and coincides with the set*

$$\mathbf{D}_\mathfrak{m} 0 = \big\{ \mathbf{V} \in L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d) : (0, \mathbf{V}) \in G \big\}. \tag{11.2.29}$$

(2) *For every* $\mathbf{V} \in G_0$ *there exists a sequence* $F_n \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$, $n \in \mathbb{N}$, *such that*

$$F_n \to 0 \textit{ in } L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}), \ DF_n \to \mathbf{V} \textit{ strongly in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d). \tag{11.2.30}$$

Every element $\mathbf{V} \in G_0$ *is therefore characterized by the property*

$$\begin{gathered}\textit{for every } \varepsilon > 0 \textit{ there exists } F \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right) \textit{ s.t.} \\ \|F\|_{L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})} \leqslant \varepsilon, \quad \|DF - \mathbf{V}\|_{L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)} \leqslant \varepsilon.\end{gathered} \tag{11.2.31}$$

(3) G_0 *satisfies the locality property*

$$\textit{for every } \mathbf{V} \in G_0, \ H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) : \quad H\mathbf{V} \in G_0. \tag{11.2.32}$$

Proof. We have already observed that G is a closed vector space, coinciding with the weak closure of $\big\{(F, DF) : F \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)\big\}$; in view of (11.2.24), (11.2.25) precisely characterizes the elements $\mathbf{V}$ for which $(0, \mathbf{V}) \in G$. Therefore the first two claims are obvious.

Let us eventually prove the last claim. We first consider the case when $H \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$. If $\mathbf{V} \in G_0$ we can find a sequence $F_n \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ such that

(11.2.30) holds. Setting $G_n := HF_n$, since H is bounded we clearly have $G_n \to 0$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$; moreover, by the Leibnitz rule we get

$$DG_n = HDF_n + F_n DH \to H\boldsymbol{V} \tag{11.2.33}$$

since $DH \in L^\infty(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ and $F_n \to 0$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$. We deduce that $H\boldsymbol{V} \in G_0$ as well.

If now H is a function in $L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ we can find by (10.1.4) a uniformly bounded sequence $H_n \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ converging to H $\mathfrak{m}$-a.e. in $\mathcal{P}_2(\mathbb{R}^d)$, so that $H_n\boldsymbol{V} \to H\boldsymbol{V}$ in $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$. Being G_0 a closed subspace and $H_n\boldsymbol{V} \in G_0$ by the previous step, we deduce that $H\boldsymbol{V} \in G_0$. □

We now define

$$\begin{aligned} T &:= \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}}) \cap G_0^\perp \\ &= \left\{ \boldsymbol{V} \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}}) : \langle \boldsymbol{V}, \boldsymbol{W} \rangle_{L^2} = 0 \text{ for every } \boldsymbol{W} \in G_0 \right\} \end{aligned} \tag{11.2.34}$$

We can now obtain our main structure result.

Theorem 11.2.10. *For every $F \in H^{1,2}(\mathbb{W}_2)$ we have $D_\mathfrak{m} F \in T$ and for every $\boldsymbol{V} \in G_0$ we have the pointwise orthogonality property*

$$\int_{\mathbb{R}^d} D_\mathfrak{m} F(\mu, x) \cdot \boldsymbol{V}(\mu, x) \, d\mu(x) = 0 \quad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{11.2.35}$$

If $\boldsymbol{V} \in \boldsymbol{D}_\mathfrak{m} F$ then $\boldsymbol{V} - D_\mathfrak{m} F \in G_0$. In particular for every $F \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ $DF - D_\mathfrak{m} F \in G_0$ and for every $G \in H^{1,2}(\mathbb{W}_2)$

$$\int_{\mathbb{R}^d} D_\mathfrak{m} F(\mu, x) \cdot D_\mathfrak{m} G(\mu, x) \, d\mu(x) = \int_{\mathbb{R}^d} DF(\mu, x) \cdot D_\mathfrak{m} G(\mu, x) \, d\mu(x) \tag{11.2.36}$$

for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Finally, $D_\mathfrak{m} F$ is the element of minimal L^2-norm in $\boldsymbol{D}_\mathfrak{m} F$.

Proof. Let us first observe that if $F_n \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ satisfies (11.2.30) and $\tilde{F}_n \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ satisfies (11.2.1), we have $F_n + \tilde{F}_n \to F$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$, with $D(F_n + \tilde{F}_n) \to D_\mathfrak{m} F + \boldsymbol{V}$, so that by (10.1.17) we get

$$\mathsf{CE}_2(F) = \int |D_\mathfrak{m} F|^2 \, d\boldsymbol{\mathfrak{m}} \leqslant \int |D_\mathfrak{m} F + \boldsymbol{V}|^2 \, d\boldsymbol{\mathfrak{m}}. \tag{11.2.37}$$

Since $\boldsymbol{V}$ is arbitrary in G_0 we deduce that

$$\int D_\mathfrak{m} F \cdot \boldsymbol{V} \, d\boldsymbol{\mathfrak{m}} = 0 \quad \text{for every } \boldsymbol{V} \in G_0.$$

Replacing $\boldsymbol{V}$ with $H\boldsymbol{V}$, $H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ we get

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} D_\mathfrak{m} F \cdot \boldsymbol{V} \, d\mu(x) \Big) H(\mu) \, d\mathfrak{m}(\mu) = 0 \quad \text{for } \boldsymbol{V} \in G_0, \ H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}), \tag{11.2.38}$$

which yields (11.2.35).

If now $F_n \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ converges strongly to F with $DF \rightharpoonup \boldsymbol{V}$, selecting $\tilde{F}_n$ as above, we have $F_n - \tilde{F}_n \to 0$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ and $D(F_n - \tilde{F}_n) \rightharpoonup \boldsymbol{G} - D_\mathfrak{m} F$ weakly in $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$, so that $\boldsymbol{G} - D_\mathfrak{m} F \in G_0$. By (11.2.37) we conclude that $D_\mathfrak{m} F$ is the element of minimal norm in $\boldsymbol{D}_\mathfrak{m} F = D_\mathfrak{m} F + G_0$. □

We can give a “pointwise” interpretation of the orthogonality properties of the previous Theorem. Let us select an orthonormal set $O_0 := \{\mathbf{V}_n : n \in \mathbb{N}\} \subset \mathcal{L}^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$ dense in G_0 (we are thus assuming that $\mathbf{V}_n$ are Borel vector fields everywhere defined). Since

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} |\mathbf{V}_n(\mu, x)|^2 \, d\mu(x) \Big) \, d\mathfrak{m}(\mu) = 1$$

we deduce that there exists a $\mathfrak{m}$-negligible set $\mathcal{N} \subset \mathcal{P}_2(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} |\mathbf{V}_n(\mu, x)|^2 \, d\mu(x) < \infty \quad \text{for every } n \in \mathbb{N},\ \mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}. \tag{11.2.39}$$

We thus define for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}$

$$G_0[\mu] := \overline{\operatorname{span}\{\mathbf{V}_n[\mu] : n \in \mathbb{N}\}} \subset L^2(\mathbb{R}^d, \mu; \mathbb{R}^d), \tag{11.2.40}$$

$$T[\mu] := \big(G_0[\mu]\big)^{\perp} \cap \operatorname{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) \tag{11.2.41}$$

Theorem 11.2.11. *Let* $F \in H^{1,2}(\mathbb{W}_2)$ *and* $\mathbf{V} \in L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$.

(1) $\mathbf{V}$ *belongs to* G_0 *if and only if for* $\mathfrak{m}$*-a.e.* μ $\mathbf{V}[\mu] \in G_0[\mu]$.

(2) $\mathbf{V}$ *belongs to* T *if and only if for* $\mathfrak{m}$*-a.e.* μ $\mathbf{V}[\mu] \in T[\mu]$.

(3) $D_\mathfrak{m} F[\mu] \in T[\mu]$ *for* $\mathfrak{m}$*-a.e.* μ.

(4) If $F \in \operatorname{Cyl}\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ *then* $D_\mathfrak{m} F[\mu]$ *is the* $L^2(\mathbb{R}^d, \mu)$*-orthogonal projection of* $DF[\mu]$ *on* $T[\mu]$ *for* $\mathfrak{m}$*-a.e.* $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

Proof. If $\mathbf{V} \in G_0$ we can write $\mathbf{V} = \lim_{N\to\infty} \mathbf{V}^N$ in $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$ where $\mathbf{V}^N = \sum_{n=1}^N u_n \mathbf{V}_n$ is the orthogonal projection of $\mathbf{V}$ on the space generated by $\{\mathbf{V}_1, \cdots, \mathbf{V}_N\}$, with $u_n := \langle V, V_n \rangle$. Clearly $\mathbf{V}^N[\mu] \in G_0[\mu]$ for every $N \in \mathbb{N}$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}$. Moreover we can find a subsequence, not relabeled, and a $\mathfrak{m}$-negligible set $\mathcal{N}' \supset \mathcal{N}$ such that $\mathbf{V}^N[\mu] \to \mathbf{V}[\mu]$ in $L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}'$, so that $\mathbf{V}[\mu] \in G_0[\mu]$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}'$.

Let now $\mathbf{V} \in L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$ be a vector field such that $\mathbf{V}[\mu] \in G_0[\mu]$ for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Since G_0 is a closed subspace, in order to show that $\mathbf{V} \in G_0$, it is sufficient to prove that the scalar product with every element $\boldsymbol{W} \in G_0^\perp$ vanishes.

If $\boldsymbol{W} \in G_0^\perp$ then for every $H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ and every $n \in \mathbb{N}$ we get

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} \boldsymbol{W} \cdot \mathbf{V}_n \, d\mu(x) \Big) H(\mu) \, d\mathfrak{m}(\mu) = 0,$$

since $H\mathbf{V}_n \in G_0$ by (11.2.32). Being H arbitrary, we find that there exists a $\mathfrak{m}$-negligible set $\mathcal{N}'' \subset \mathcal{P}_2(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} \boldsymbol{W}[\mu] \cdot \mathbf{V}_n[\mu] \, d\mu = 0 \quad \text{for every } n \in \mathbb{N},\ \mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}'',$$

so that $\boldsymbol{W}[\mu] \in \big(G_0[\mu]\big)^{\perp}$ for $\mathfrak{m}$-a.e.$\mu \in \mathcal{P}_2(\mathbb{R}^d)$. We then deduce that

$$\int_{\mathbb{R}^d} \boldsymbol{W}[\mu] \cdot \mathbf{V}[\mu] \, d\mu = 0 \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d),$$

and therefore

$$\langle \mathbf{W}, \mathbf{V} \rangle_{L^2} = \int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} \mathbf{W} \cdot \mathbf{V} \, d\mu(x) \Big) \, dm(\mu) = 0.$$

The previous argument also shows that a vector field $\mathbf{V}$ belongs to $G_0^\perp$ if and only if $\mathbf{V}[\mu] \in \big(G_0[\mu]\big)^\perp$ for m-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. This fact, together with the very definition of $\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathbf{m})$ (11.2.26), yields claim (2).

Claim (3) just follows by Theorem 11.2.10, since (11.2.35) shows that, for every $F \in H^{1,2}(\mathbb{W}_2)$, $D_m F[\mu] \in T[\mu]$ for m-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

If $F \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ we also have $DF[\mu] - D_m F[\mu] \in G_0[\mu] = \big(\mathcal{G}[\mu]\big)^\perp$ m-a.e., so that $D_m F[\mu]$ is the $L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$-orthogonal projection of $DF[\mu]$ on $G[\mu]$, as stated in Claim (4). □

We can now interpret the above results in terms of the nonsmooth tangent and cotangent structures introduced and developed by Gigli in [55]. Since we are in the Hilbertian case, we can identify the cotangent module $L^2(T^*\mathcal{P}_2(\mathbb{R}^d))$ and dual tangent module $L^2(T\mathcal{P}_2(\mathbb{R}^d))$ with the Hilbert space T defined by (11.2.34). Let us report a useful characterization of the cotangent module $L^2(T^*X)$ [57, Theorem 4.1.1] for a general metric measure space (X, d, m).

Theorem 11.2.12. *Let* (X, d, m) *be a metric measure space. Then there exists a unique pair* $((\mathcal{M}, \|\cdot\|_{\mathcal{M}}, \cdot_{\mathcal{M}}, |\cdot|_{\mathcal{M}}), \mathsf{diff})$ *such that* $(\mathcal{M}, \|\cdot\|_{\mathcal{M}}, \cdot_{\mathcal{M}}, |\cdot|_{\mathcal{M}})$ *is a* $L^2(X, m)$*-normed* $L^\infty(X, m)$ *module (cf. [57, Definition 3.1.1]) and* $\mathsf{diff} : D^{1,2}(X, d, m) \to \mathcal{M}$ *is a linear operator such that*

- *(i)* $|\mathsf{diff}(f)|_{\mathcal{M}} = |Df|_\star$ *m-a.e. in* X *for every* $f \in D^{1,2}(X, d, m)$.
- *(ii)* $\mathcal{M}$ *is generated by* $\big\{\mathsf{diff}(f) : f \in D^{1,2}(X, d, m)\big\}$.

Uniqueness is intended in the following sense: if $((\tilde{\mathcal{M}}, \|\cdot\|_{\tilde{\mathcal{M}}}, \cdot_{\tilde{\mathcal{M}}}, |\cdot|_{\tilde{\mathcal{M}}}), \tilde{\mathsf{diff}})$ *is another pair with the above properties, then there exists a module isomorphism* $\mathcal{J} : \mathcal{M} \to \tilde{\mathcal{M}}$ *such that* $\tilde{\mathsf{diff}} = \mathcal{J} \circ \mathsf{diff}$.

We thus have the following result.

Theorem 11.2.13. *There exists a module isomorphism* $\mathcal{I} : T \to L^2(T^*\mathcal{P}_2(\mathbb{R}^d)) \cong L^2(T\mathcal{P}_2(\mathbb{R}^d))$ *such that* $\mathcal{I} \circ D_m$ *coincides with the abstract differential operator taking values in* $L^2(T^*\mathcal{P}_2(\mathbb{R}^d))$ *as in [55, Defintion 2.2.2].*

Proof. It is enough to show that T (with an appropriate module structure) and the map D_m satisfy the properties listed in Theorem 11.2.12.
If as $\|\cdot\|_T$ we take the $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d)$ norm, it is clear that $(T, \|\cdot\|_T)$ is a Banach space, being closed by Lemma 11.2.8. The pointwise product $\cdot_T : L^\infty(\mathcal{P}_2(\mathbb{R}^d), m) \times T \to T$ is well defined by (11.2.27) and (11.2.32), bilinear and associative in $L^\infty(\mathcal{P}_2(\mathbb{R}^d), m)$ by definition. Defining the pointwise norm $|\cdot|_T$ as the map sending $\mathbf{V} \in T$ to $\|\mathbf{V}[\mu]\|_\mu$, we immediately have that $\|\mathbf{V}\|_T = \||\mathbf{V}|_T\|_{L^2(\mathcal{P}_2(\mathbb{R}^d), m)}$ and $|H \cdot_T \mathbf{V}|_T = |H||\mathbf{V}|_T$ m-a.e. in $\mathcal{P}_2(\mathbb{R}^d)$ for every $\mathbf{V} \in T$ and every $H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), m)$. This shows that $(T, \|\cdot\|_T, \cdot_T, |\cdot|_T)$ is a $L^2(\mathcal{P}_2(\mathbb{R}^d), m)$-normed $L^\infty(\mathcal{P}_2(\mathbb{R}^d), m)$ module.

Taking as diff the map $D_{\mathfrak{m}} : D^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m}) \to T$, we see that it is well defined and linear by Theorem 11.2.1 and Theorem 11.2.10. Property (i) of Theorem 11.2.12 follows by (11.2.8). Finally property (ii) of Theorem 11.2.12, meaning that ([57, Definition 3.1.13]) T coincides with the $\|\cdot\|_T$-closure of

$$\operatorname{span}\left\{HD_{\mathfrak{m}}F : H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}),\ F \in D^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})\right\},$$

follows by (11.2.28) and the definition of T. This shows the existence of the module isomorphism $\mathcal{I} : T \to L^2(T^*\mathcal{P}_2(\mathbb{R}^d))$.
Finally, notice that $L^2(T^*\mathcal{P}_2(\mathbb{R}^d)) \cong L^2(T\mathcal{P}_2(\mathbb{R}^d))$ since $(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ is infinitesimally Hilbertian by Corollary 11.1.12 (see also [57, Theorem 4.3]). □

11.2.2 *Examples*

Isometric embedding of Euclidean Sobolev spaces

Let Ω be a Lipschitz bounded open set in $\mathbb{R}^d$. For every $\omega \in \Omega$ let us consider the Dirac mass δ_ω concentrated at ω. The map $\iota : \omega \mapsto \delta_\omega$ is an isometry between $\mathbb{R}^d$ and $\iota(\mathbb{R}^d) \subset \mathcal{P}_2(\mathbb{R}^d)$. Setting $\mathfrak{m} := \iota_\sharp \mathscr{L}^d \llcorner \Omega$ we easily see that $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ is isomorphic to $H^{1,2}(\Omega)$.

In this case only Dirac masses are involved and cylindrical functions are of the form $F(\delta_\omega) = \psi(\boldsymbol{\phi}(\omega))$, so that the Wasserstein gradient reduces to the usual gradient of $\psi \circ \boldsymbol{\phi}$.

Another isometric embedding is also possible: we fix a reference measure $\lambda \in \mathcal{P}_2(\mathbb{R}^d)$ symmetric w.r.t. the origin and we consider the map $\iota : \Omega \to \mathcal{P}_2(\mathbb{R}^d)$ given by

$$\iota(\omega) := \lambda(\cdot - \omega) = (t_\omega)_\sharp \lambda, \quad t_\omega(x) := x + \omega, \quad \omega \in \Omega. \tag{11.2.42}$$

Every function $F : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ corresponds to a map

$$\hat{F}(\omega) = F((t_\omega)_\sharp \lambda). \tag{11.2.43}$$

In the case of a cylindrical function as in (11.1.5) we get

$$\begin{aligned} \hat{F}(\omega) &= \psi\Big(\int \phi_1(x+\omega)\,d\lambda(x), \cdots, \int \phi_N(x+\omega)\,d\lambda(x)\Big) \\ &= \psi\Big(\phi_1 * \lambda(\omega), \cdots, \phi_N * \lambda(\omega)\Big). \end{aligned} \tag{11.2.44}$$

In this case (identifying $\iota(\omega)$ with ω) we have

$$DF(\omega, x) = \sum_{j=1}^N \partial\psi_j(\phi_1 * \lambda(\omega), \cdots, \phi_N * \lambda(\omega))\nabla\phi_j(x) \tag{11.2.45}$$

and

$$\|DF[\omega]\|_\omega^2 = \int_{\mathbb{R}^d} \Big|\sum_{j=1}^N \partial\psi_j(\phi_1 * \lambda(\omega), \cdots, \phi_N * \lambda(\omega))\nabla\phi_j(x+\omega)\Big|^2 d\lambda(x). \tag{11.2.46}$$

Moreover, ι is an isometry of $\mathbb{R}^d$ into $\mathcal{P}_2(\mathbb{R}^d)$, so that $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ is still isomorphic to $H^{1,2}(\Omega)$. It follows that the $\mathfrak{m}$-Wasserstein gradient of F is

$$D_{\mathfrak{m}} F(\omega, x) = \sum_{j=1}^{N} \partial\psi_j(\phi_1 * \lambda(\omega), \cdots, \phi_N * \lambda(\omega)) \nabla\phi_j * \lambda(\omega) \tag{11.2.47}$$

independent of x and the minimal relaxed gradient is

$$|D_{\mathfrak{m}} F|_{\star}^2(\omega) = \left| \sum_{j=1}^{N} \partial\psi_j(\phi_1 * \lambda(\omega), \cdots, \phi_N * \lambda(\omega)) \nabla\phi_j * \lambda(\omega) \right|^2. \tag{11.2.48}$$

Gaussian distributions

Let now $\kappa = N(\omega, \Sigma) := (\det(2\pi\Sigma))^{-1/2} e^{-\frac{1}{2}\langle x, \Sigma^{-1} x\rangle} \mathscr{L}^d$ be a Gaussian measure with mean ω and covariance matrix $\Sigma \in \mathrm{Sym}^+(d)$, the space of symmetric and positive definite $d \times d$-matrices; we consider the set

$$\mathcal{N}^d := \left\{ \mathcal{N}(\omega, \Sigma) : \omega \in \mathbb{R}^d,\ \Sigma \in \mathrm{Sym}^+(d) \right\}, \tag{11.2.49}$$

endowed with the Wasserstein distance and a Borel probability measure $\mathfrak{m}$ concentrated on $\mathcal{N}^d$. Since

$$W_2^2(N(\omega_1, \Sigma_1), N(\omega_2, \Sigma_2)) = |\omega_1 - \omega_2|^2 + \mathrm{tr}\Sigma_1 + \mathrm{tr}\Sigma_2 - 2\mathrm{tr}\left(\Sigma_1^{1/2} \Sigma_2 \Sigma_1^{1/2}\right)^{1/2}, \tag{11.2.50}$$

$H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ is isometric to $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), d, \hat{\mathfrak{m}})$ where $\mathcal{P}_2(\mathbb{R}^d) = \mathbb{R}^d \times \mathrm{Sym}^+(d) \subset \mathbb{R}^d \times \mathbb{R}^{d \times d}$ endowed with the distance d induced by the formula (11.2.50) and $\hat{\mathfrak{m}}$ is the measure induced by $\mathfrak{m}$.

Closable case

Following [44], we assume the following integration by parts formula: for every $G \in \mathrm{Cyl}^\infty(\mathcal{P}_2(\mathbb{R}^d))$ and $w \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$ there exists $D_w^* G \in L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ such that for every $F \in \mathrm{Cyl}^\infty(\mathcal{P}_2(\mathbb{R}^d))$ it holds

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} DF(\mu, x) \cdot w(x) \, d\mu(x) \right) G(\mu) \, d\mathfrak{m}(\mu) = \int_{\mathcal{P}_2(\mathbb{R}^d)} D_w^* G(\mu) F(\mu) \, d\mathfrak{m}(\mu).$$

This equality implies that $G_0 = \{0\}$ i.e. that pCE_2 is closable. We notice that the measure $\mathfrak{m}$ induced by the immersion in the space of delta measure considered at the beginning of this section satisfies the integration by parts formula above (see also Example 5.4 in [44]). In [44], in case the base space is a compact Riemannian manifold, are reported a few example of measures $\mathfrak{m}$ satisfying the (Riemannian analogue of the) integration by parts formula: the normalized mixed Poisson measure (Example 5.11 in [44] and [2, 103]), the entropic measure over $\mathbb{S}^1$ (Example 5.15 in [44] and [101]) and the Malliavin–Shavgulidze image measure (Example 5.18 in [44] and [81]).

11.3 EXTENSION TO RIEMANNIAN MANIFOLDS AND HILBERT SPACES

The aim of this Section is to extend the density result stated in Theorem 11.1.11 from the finite dimensional and flat space $\mathbb{R}^d$ to Riemannian manifolds and (possibly infinite dimensional) separable Hilbert spaces. Our first step deals with manifolds embedded in some Euclidean space $\mathbb{R}^d$ and in fact we will consider more general closed subsets of $\mathbb{R}^d$.

11.3.1 *Intrinsic Wasserstein spaces on closed subsets of* $\mathbb{R}^d$

In this subsection we denote by ρ the Euclidean distance on $\mathbb{R}^d$. $\mathcal{P}_2(\mathbb{R}^d)$ still denotes the subset of Borel probability measure on $\mathbb{R}^d$ with finite second ρ-moment and W_2 is the Wasserstein distance on $\mathcal{P}_2(\mathbb{R}^d)$ induced by ρ.

We assume that $C \subset \mathbb{R}^d$ is a closed set and that σ is a distance on C such that (C, σ) is a complete and separable metric space and

$$\rho(x_1, x_2) \leqslant \sigma(x_1, x_2) \leqslant \rho_{C,\ell}(x_1, x_2) \quad \text{for every } x_1, x_2 \in C, \tag{11.3.1}$$

where $\rho_{C,\ell}$ is defined as in (10.3.3) with respect to the distance $d := \rho$. Since the topology induced by σ is stronger than the Euclidean topology and they are both Polish topologies, the Borel sets of (C, σ) coincide with the Borel sets of C as a subset of the Euclidean space $\mathbb{R}^d$. This means that every Borel probability measure on $\mathbb{R}^d$ with support contained in C can be identified with a Borel probability measure in (C, σ). Conversely any probability measure on (C, σ) extends to a probability measure on $\mathbb{R}^d$. We can thus denote unambiguously by $\mathcal{P}(C)$ the set of Borel probability measures on C and by $\mathcal{P}_{2,\sigma}(C)$ the elements of $\mathcal{P}(C)$ with finite second σ-moment.

$\mathcal{P}_{2,\sigma}(C)$ can be identified with the subset of $\mathcal{P}_2(\mathbb{R}^d)$

$$\left\{\mu \in \mathcal{P}_2(\mathbb{R}^d) : \mathrm{supp}(\mu) \subset C, \quad \int_C \sigma^2(x_0, x)\, d\mu(x) < +\infty \text{ for some } x_0 \in C\right\}.$$

We will denote by $\iota : C \to \mathbb{R}^d$ the inclusion map; $\boldsymbol{\iota} : \mathcal{P}_{2,\sigma}(C) \to \mathcal{P}_2(\mathbb{R}^d)$ is the corresponding continuous injection given by $\boldsymbol{\iota}(\mu) := \iota_\sharp \mu$, which may be identified with the inclusion map of $\mathcal{P}_{2,\sigma}(C)$ into $\mathcal{P}_{2,\sigma}(\mathbb{R}^d)$.

Since $(\mathcal{P}_2(C), W_{2,\sigma})$ is a complete and separable metric space and the topology induced by $W_{2,\sigma}$ is stronger than the topology induced by W_2, we deduce that $\mathcal{P}_{2,\sigma}(C)$ is a Lusin (and therefore Borel) subset of $\mathcal{P}_2(\mathbb{R}^d)$.

If $\mathfrak{m}$ is a Borel probability measure on $\mathcal{P}_{2,\sigma}(C)$, $\boldsymbol{\iota}_\sharp \mathfrak{m}$ is the Borel measure in $\mathcal{P}_2(\mathbb{R}^d)$ which is concentrated on $\mathcal{P}_{2,\sigma}(C)$ and satisfies $\boldsymbol{\iota}_\sharp \mathfrak{m}(Z) = \mathfrak{m}(Z \cap \mathcal{P}_{2,\sigma}(C))$ for every Borel set $Z \subset \mathcal{P}_2(\mathbb{R}^d)$.

In a similar way, if $F : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is a Borel (or $\boldsymbol{\iota}_\sharp \mathfrak{m}$- measurable) map, we will set $\boldsymbol{\iota}^* F := F \circ \boldsymbol{\iota} : \mathcal{P}_{2,\sigma}(C) \to \mathbb{R}$.

Theorem 11.3.1. *We have* $H^{1,2}(\mathcal{P}_{2,\sigma}(C), W_{2,\sigma}, \mathfrak{m}) \cong H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \boldsymbol{\iota}_\sharp \mathfrak{m})$ *with equal minimal relaxed gradient, meaning that*

$$|D(\boldsymbol{\iota}^* F)|_\star = \boldsymbol{\iota}^* \left(|DF|_\star\right) \textit{ for every } F \in H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \boldsymbol{\iota}_\sharp \mathfrak{m}). \tag{11.3.2}$$

In particular $H^{1,2}(\mathcal{P}_{2,\sigma}(C), W_{2,\sigma}, \mathfrak{m})$ *is an Hilbert space and the algebra of cylindrical functions* $\iota^*\big(\mathrm{Cyl}\,(\mathcal{P}_2(\mathbb{R}^d))\big)$ *is dense in* $H^{1,2}(\mathcal{P}_{2,\sigma}(C), W_{2,\sigma}, \mathfrak{m})$ *in the sense of* (11.1.18).

Proof. We want to apply Theorem 10.3.3 where $X := \mathcal{P}_2(\mathbb{R}^d)$, $\mathsf{d} := W_2$, $Y := \mathcal{P}_{2,\sigma}(C)$, and $\delta := W_{2,\sigma}$. The first assumption of Condition (A), $\iota_\sharp\mathfrak{m}(\mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{P}_{2,\sigma}(C)) = 0$, is clearly satisfied by construction.

In order to prove (10.3.9) we consider a W_2-Lipschitz curve $\mu : [0, \ell] \to \mathcal{P}_2(\mathbb{R}^d)$ parametrized by the W_2-arclength such that $\mu(s) \in \mathcal{P}_{2,\sigma}(C)$ for $\mathscr{L}^1$-a.e. $s \in [0, \ell]$. Since the map μ is continuous in $\mathcal{P}_2(\mathbb{R}^d)$, C is a closed set, and $\mu_s(\mathbb{R}^d \setminus C) = 0$ for $\mathscr{L}^1$-a.e. $s \in [0, \ell]$, we conclude that $\mu_s(\mathbb{R}^d \setminus C) = 0$ for every $s \in [0, \ell]$.

By [5, Theorem 8.2.1, Theorem 8.3.1]) there exists a measure $\boldsymbol{\eta} \in \mathcal{P}(C([0, \ell]; \mathbb{R}^d))$ concentrated on absolutely continuous curves such that $(\mathsf{e}_t)_\sharp(\boldsymbol{\eta}) = \mu(t)$ for every $t \in [0, \ell]$ and

$$\int |\gamma'(t)|^2 \,\mathrm{d}\boldsymbol{\eta}(\gamma) = \int |\dot\gamma|_\rho^2(t)\,\mathrm{d}\boldsymbol{\eta}(\gamma) = 1 \quad \text{for a.e. } t \in [0, \ell]. \tag{11.3.3}$$

Let us also consider the function $\zeta(x) := \mathrm{dist}(x, C) \wedge 1$, $x \in \mathbb{R}^d$, where $\mathrm{dist}(x, C) := \min_{z \in C} \rho(x, z)$. ζ is a bounded Lipschitz function which vanishes precisely on C. Fubini's Theorem yields

$$\int \Big(\int_0^\ell \zeta(\gamma(t))\,\mathrm{d}t\Big)\,\mathrm{d}\boldsymbol{\eta}(\gamma) = \int_0^\ell \int \zeta(\mathsf{e}_t(\gamma))\,\mathrm{d}\boldsymbol{\eta}(\gamma)\,\mathrm{d}t = \int_0^\ell \int_{\mathbb{R}^d} \zeta\,\mathrm{d}\mu_t\,\mathrm{d}t = 0$$

since $\int \zeta(x)\,\mathrm{d}\mu_t = 0$ for $\mathscr{L}^1$-a.e. $t \in (0, \ell)$.

It follows that $\int_0^\ell \zeta(\gamma(t))\,\mathrm{d}t = 0$ for $\boldsymbol{\eta}$-a.e. γ, so that the set of $t \in [0, \ell]$ for which $\gamma(t) \in C$ is dense in $[0, \ell]$. Being C closed, we conclude that γ takes values in C for $\boldsymbol{\eta}$-a.e. γ.

We can now estimate the $W_{2,\sigma}$ distance between the two measures μ_{t_0} and μ_{t_1}, where $0 \leqslant t_0 < t_1 \leqslant \ell$:

$$\begin{aligned} W_{2,\sigma}^2(\mu_{t_0}, \mu_{t_1}) &\leqslant \int \sigma^2(\gamma(t_0), \gamma(t_1))\,\mathrm{d}\boldsymbol{\eta}(\gamma) \leqslant \int \Big(\int_{t_0}^{t_1} |\dot\gamma|_\sigma(s)\,\mathrm{d}s\Big)^2 \mathrm{d}\boldsymbol{\eta}(\gamma) \\ &= \int \Big(\int_{t_0}^{t_1} |\dot\gamma|_\rho(s)\,\mathrm{d}s\Big)^2 \mathrm{d}\boldsymbol{\eta}(\gamma) \leqslant (t_1 - t_0) \int\int_{t_0}^{t_1} |\dot\gamma|_\rho^2\,\mathrm{d}s\,\mathrm{d}\boldsymbol{\eta}(\gamma) \\ &= (t_1 - t_0)\int_{t_0}^{t_1}\int |\dot\gamma|_\rho^2\,\mathrm{d}\boldsymbol{\eta}(\gamma)\,\mathrm{d}s = (t_1 - t_0)^2, \end{aligned}$$

where we have used (11.3.1) and Remark 10.3.2 to say that $|\dot\gamma|_\rho(s) = |\dot\gamma|_\sigma(s)$.

Choosing $t_0 \in [0, \ell]$ such that $\mu_{t_0} \in \mathcal{P}_{2,\sigma}(C)$ we deduce that $\mu_{t_1} \in \mathcal{P}_{2,\sigma}(C)$ as well for every $t_1 \in [0, \ell]$. This concludes the proof of property (A).

Condition (B) corresponds to

$$W_2(\mu_0, \mu_1) \leqslant W_{2,\sigma}(\mu_0, \mu_1) \leqslant (W_2)_{Y,\ell}(\mu_0, \mu_1) \quad \text{for every } \mu_0, \mu_1 \in Y = \mathcal{P}_{2,\sigma}(C), \tag{11.3.4}$$

where $(W_2)_{Y,\ell}(\mu_0, \mu_1)$ is defined as in (10.3.3) with W_2 in place of d. The first inequality immediately follows by (11.3.1); to prove the second one, we use (10.3.4) and the above estimate with $t_0 = 0$ and $t_1 = \ell$ for a a W_2-Lipshitz curve $\mu : [0,\ell] \to Y$ such that $|\dot\mu|_{W_2} = 1$ a.e. in $[0,\ell]$ with $\mu(0) = \mu_0$ and $\mu(\ell) = \mu_1$. Taking the infimum w.r.t. ℓ we obtain (11.3.4). □

11.3.2 *Wasserstein-Sobolev space on complete Riemannian manifolds*

In this subsection, we discuss the case of the Sobolev space $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,d_\mathbb{M}}, \mathfrak{m})$ where $(\mathbb{M}, d_\mathbb{M})$ is a smooth and complete Riemannian manifold endowed with the canonical Riemannian distance $d_\mathbb{M}$ (inducing the Wasserstein distance $W_{2,d_\mathbb{M}}$) and $\mathfrak{m}$ is a Borel probability measure on $\mathcal{P}_2(\mathbb{M})$. We will denote by $\mathscr{A}$ the unital algebra generated by $\{L_f : f \in C^1_c(\mathbb{M})\}$.

Theorem 11.3.2. $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,d_\mathbb{M}}, \mathfrak{m})$ *is an Hilbert space and the algebra* $\mathscr{A}$ *is (strongly) dense: for every* $F \in H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,d_\mathbb{M}}, \mathfrak{m})$ *there exists a sequence* $F_n \in \mathscr{A}$, $n \in \mathbb{N}$ *such that*

$$F_n \to F, \quad \mathrm{lip}(F_n) \to |DF|_\star \quad \textit{strongly in } L^2(\mathcal{P}_2(\mathbb{M}), \mathfrak{m}). \tag{11.3.5}$$

Proof. By Nash isometric embedding Theorem [87] we can find a dimension d, and an isometric embedding $\jmath : \mathbb{M} \to \jmath(\mathbb{M}) \subset \mathbb{R}^d$. On $M := \jmath(\mathbb{M})$ we can define the (Riemannian) metric d_M inherited by $d_\mathbb{M}$: $d_M(\jmath(x), \jmath(y)) = d_\mathbb{M}(x,y)$ so that $\jmath$ is an isometry and (M, d_M) is a complete and separable metric space. We denote by $\boldsymbol{\jmath} := \jmath_\sharp$ the corresponding isometry between $(\mathcal{P}_2(\mathbb{M}), W_{2,d_\mathbb{M}})$ and $(\mathcal{P}_2(M), W_{2,d_M})$ and we also set $\tilde{\mathfrak{m}} := \boldsymbol{\jmath}_\sharp \mathfrak{m} \in \mathcal{P}(\mathcal{P}_2(M))$.

It is clear that the map $\boldsymbol{\jmath}^* : F \mapsto F \circ \boldsymbol{\jmath}$ induces a linear isometric isomorphism between $H^{1,2}(\mathcal{P}_2(M), W_{2,d_M}, \tilde{\mathfrak{m}})$ and $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,d_\mathbb{M}}, \mathfrak{m})$.

Since $\mathbb{M}$ is complete and $\jmath$ is an embedding, M is a closed subset of $\mathbb{R}^d$ and d_M induces on M the relative topology of $\mathbb{R}^d$. Since $\jmath$ is isometric, we also have

$$\rho(y_1, y_2) \leqslant d_M(y_1, y_2) = \rho_{M,\ell}(y_1, y_2) \quad \text{for every } y_1, y_2 \in M, \tag{11.3.6}$$

where $\rho_{M,\ell}$ is as in (10.3.3) and ρ denotes the Euclidean distance on $\mathbb{R}^d$.

As in Section 11.3.1, we can introduce the inclusion map $\iota : M \to \mathbb{R}^d$ and the corresponding $\boldsymbol{\iota} = \iota_\sharp : \mathcal{P}_{2,d_M}(M) \to \mathcal{P}_2(\mathbb{R}^d)$. By Theorem 11.3.1 we have that the map $\boldsymbol{\iota}^* : F \mapsto F \circ \boldsymbol{\iota}$ provides a linear isometric isomorphism between $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \iota_\sharp \tilde{\mathfrak{m}})$ and $H^{1,2}(\mathcal{P}_{2,d_M}(M), W_{2,d_M}, \tilde{\mathfrak{m}})$ satisfying (11.3.2); we conclude that the map $\boldsymbol{\kappa}^* := \boldsymbol{\jmath}^* \circ \boldsymbol{\iota}^* = (\boldsymbol{\iota} \circ \boldsymbol{\jmath})^*$ is a isometric isomorphism between $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \boldsymbol{\kappa}_\sharp \mathfrak{m})$ (notice that $\boldsymbol{\kappa}_\sharp = \iota_\sharp \circ \jmath_\sharp$) and $H^{1,2}(\mathcal{P}_{2,d_\mathbb{M}}(M), W_{2,d_M}, \mathfrak{m})$ satisfying

$$|D(\boldsymbol{\kappa}^* F)|_\star = \boldsymbol{\kappa}^* (|DF|_\star) \text{ for every } F \in H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \boldsymbol{\kappa}_\sharp \mathfrak{m}). \tag{11.3.7}$$

This property in particular yields the Hilbertianity of $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,d_\mathbb{M}}, \mathfrak{m})$.

In order to prove that $\mathscr{A}$ is dense in $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,d_\mathbb{M}}, \mathfrak{m})$ we consider the algebra $\tilde{\mathscr{A}}$ generated by $\{L_{\tilde f} : \tilde f \in C^\infty_c(\mathbb{R}^d)\}$; Proposition 11.1.21 shows that $\tilde{\mathscr{A}}$ is strongly dense in $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \tilde{\mathfrak{m}})$, so that $\mathscr{A}' := \boldsymbol{\kappa}^*(\tilde{\mathscr{A}})$ is strongly dense in $H^{1,2}(\mathcal{P}_{2,d_\mathbb{M}}(\mathbb{M}), W_{2,d_M}, \mathfrak{m})$.

$\mathscr{A}'$ is generated by functions of the form $\boldsymbol{\kappa}^* L_{\tilde f}$, $\tilde f \in C^\infty_c(\mathbb{R}^d)$. Since

$$\boldsymbol{\kappa}^* L_{\tilde f}(\mu) = L_{\tilde f}(\boldsymbol{\kappa}(\mu)) = \int_{\mathbb{R}^d} \tilde f(\kappa(x)) \, d\mu(x) \quad \text{for every } \mu \in \mathcal{P}_{2,d_\mathbb{M}}(\mathbb{M}),$$

where $\kappa = \iota \circ \jmath$, we see that $\mathscr{A}'$ is generated by functions of the form $L_{\tilde{f}\circ\kappa}$, so that $\mathscr{A}' \subset \mathscr{A}$ and a fortiori $\mathscr{A}$ is strongly dense in $H^{1,2}(\mathcal{P}_{2,d_{\mathbb{M}}}(\mathbb{M}), W_{2,d_{\mathbb{M}}}, \mathfrak{m})$ as well.

To prove (11.3.5) (involving the asymptotic Lipschitz constants of functions in $\mathscr{A}$ with respect to the Riemannian metric) we observe that for every $\tilde{F} \in \tilde{\mathscr{A}}$ [108, Lemma 3.1.14]

$$\kappa^*\tilde{F} \in \mathscr{A}' \subset \mathscr{A}, \quad \kappa^*(\mathrm{lip}_{W_2}\tilde{F}) \geqslant \mathrm{lip}_{W_{2,d_{\mathbb{M}}}}\kappa^*\tilde{F}. \tag{11.3.8}$$

Let now $F = \kappa^*\tilde{F} \in H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,d_{\mathbb{M}}}, \mathfrak{m})$ with $\tilde{F} \in H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \tilde{\mathfrak{m}})$; there exists a sequence $\tilde{F}_n \in \tilde{\mathscr{A}}$ such that

$$\tilde{F}_n \to \tilde{F}, \quad \mathrm{lip}_{W_2}\tilde{F}_n \to |D\tilde{F}|_\star \quad \text{in } L^2(\mathcal{P}_2(\mathbb{R}^d), \tilde{\mathfrak{m}}).$$

Applying the linear isometric isomorphism κ^*, we deduce that the sequence $\kappa^*F_n \in \mathscr{A}'$ satisfies

$$\kappa^*\tilde{F}_n \to F, \quad \kappa^*\big(\mathrm{lip}_{W_2}\tilde{F}_n\big) \to \kappa^*\big(|D\tilde{F}|_\star\big) = |DF|_\star \quad \text{in } L^2(\mathcal{P}_{2,d_{\mathbb{M}}}(\mathbb{M}), \mathfrak{m}). \tag{11.3.9}$$

Up to extracting a suitable (not relabelled) subsequence and using (11.3.8), we can suppose that $\mathrm{lip}_{W_{2,d_{\mathbb{M}}}}\kappa^*\tilde{F}_n$ converges weakly in $L^2(\mathcal{P}_2(\mathbb{M}), W_{2,d_{\mathbb{M}}})$ to some $G \in L^2(\mathcal{P}_2(\mathbb{M}), W_{2,d_{\mathbb{M}}})$ relaxed gradient of F. (11.3.8) and (11.3.9) also yield

$$\begin{aligned}\int G^2\,d\mathfrak{m} &\leqslant \limsup_{n\to\infty}\int(\mathrm{lip}_{W_{2,d_{\mathbb{M}}}}\kappa^*F_n)^2\,d\mathfrak{m}\\ &\leqslant \limsup_{n\to\infty}\int\Big(\kappa^*(\mathrm{lip}_{W_2}\tilde{F}_n)\Big)^2\,d\mathfrak{m}\\ &= \int |DF|_\star^2\,d\mathfrak{m},\end{aligned}$$

showing that $G = |DF|_\star$ and $\mathrm{lip}_{W_{2,d_{\mathbb{M}}}}\kappa^*F_n \to |DF|_\star$ strongly in $L^2(\mathcal{P}_{2,d_{\mathbb{M}}}(\mathbb{M}), \mathfrak{m})$. □

11.3.3 *Wasserstein-Sobolev space on Hilbert spaces*

In this last section we will consider the case of a separable Hilbert space $(H, |\cdot|)$; as usual, the space $\mathcal{P}_2(H)$ will be endowed with the Wasserstein distance W_2 induced by the Hilbertian norm of H and we will assume that $\mathfrak{m}$ is a Borel probability measure on $\mathcal{P}_2(H)$.

We select a complete orthonormal system $E := (e_n)_{n\in\mathbb{N}}$ and the collection of maps $\pi_d : H \to \mathbb{R}^d$, $d \in \mathbb{N}$, given by

$$\pi^d(x) := \big(\langle x, e_1\rangle, \cdots, \langle x, e_d\rangle\big). \tag{11.3.10}$$

The adjoint map $\pi^{d*} : \mathbb{R}^d \to H$ is given by

$$\pi^{d*}(y_1, \cdots, y_d) := \sum_{j=1}^{d} y_j\, e_j. \tag{11.3.11}$$

The map $\hat{\pi}^d := \pi^{d*} \circ \pi^d$ is the orthogonal projection of H onto $\mathrm{span}\{e_1, \cdots, e_d\}$. We say that a function $\phi : H \to \mathbb{R}$ belongs to $C^1_b(H, E)$ if it can be written as

$$\phi := \varphi \circ \pi^d \quad \text{for some } d \in \mathbb{N},\ \varphi \in C^1_b(\mathbb{R}^d). \tag{11.3.12}$$

Clearly $\phi \in C^1_b(H)$ and its gradient $\nabla\phi$ can be written as

$$\nabla\phi = \pi^{d*} \circ \nabla\varphi \circ \pi^d, \quad \nabla\phi(x) := \sum_{j=1}^{d} \partial_j\varphi(\pi^d(x))e_j. \tag{11.3.13}$$

We then consider the algebra $\mathrm{Cyl}\,(\mathcal{P}_2(H))$ generated by $\{L_\phi : \phi \in C^1_b(H,E)\}$. For every $F \in \mathrm{Cyl}\,(\mathcal{P}_2(H))$ we can find $N \in \mathbb{N}$, a polynomial $\psi : \mathbb{R}^N \to \mathbb{R}$ and functions $\phi_n \in C^1_b(H,E)$, $n = 1, \cdots, N$, such that

$$F(\mu) = (\psi \circ L_{\boldsymbol{\phi}})(\mu), \tag{11.3.14}$$

where $\boldsymbol{\phi} = (\phi_1, \dots, \phi_N)$. As in (11.1.7) we can set

$$DF(\mu, x) := \sum_{n=1}^{N} \partial_n\psi(L_{\boldsymbol{\phi}}(\mu))\nabla\phi_n(x). \tag{11.3.15}$$

It is also easy to check that a function F belongs to $\mathrm{Cyl}\,(\mathcal{P}_2(H))$ if and only if there exists $d \in \mathbb{N}$ and $\tilde{F} \in \mathrm{Cyl}\,\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ such that

$$F(\mu) = \tilde{F}(\pi^d_\sharp(\mu)) \quad \text{for every } \mu \in \mathcal{P}_2(H), \tag{11.3.16}$$

so that

$$DF(\mu, x) = \pi^{d*}\Big(D\tilde{F}(\pi^d_\sharp\mu, \pi^d(x))\Big), \quad \|DF[\mu]\|_\mu = \|D\tilde{F}(\pi^d_\sharp\mu)\|_{\pi^d_\sharp\mu}. \tag{11.3.17}$$

By Proposition 11.1.10 and using (11.3.17) it is not difficult to check that

$$\|DF[\mu]\|_\mu = \mathrm{lip} F(\mu) \quad \text{for every } \mu \in \mathcal{P}_2(H). \tag{11.3.18}$$

Adapting in an obvious way the definitions in (11.1.16) and (11.1.17) to the Hilbertian framework, we have the following result.

Theorem 11.3.3. $H^{1,2}(\mathcal{P}_2(H), W_2, \mathfrak{m})$ *is an Hilbert space and the algebra* $\mathrm{Cyl}\,(\mathcal{P}_2(H))$ *is (strongly) dense: for every* $F \in H^{1,2}(\mathcal{P}_2(H), W_2, \mathfrak{m})$ *there exists a sequence* $F_n \in \mathrm{Cyl}\,(\mathcal{P}_2(H))$, $n \in \mathbb{N}$ *such that*

$$F_n \to F, \quad \mathrm{lip}(F_n) \to |DF|_\star \quad \textit{strongly in } L^2(\mathcal{P}_2(H), \mathfrak{m}). \tag{11.3.19}$$

Proof. Let us set $\mathscr{A} := \mathrm{Cyl}\,(\mathcal{P}_2(H))$; we use Theorem 10.2.1 and we want to prove that for every $\nu \in \mathcal{P}_2(H)$ the function

$$F(\mu) := W_2(\nu, \mu) \quad \text{satisfies} \quad |DF|_{\star,\mathscr{A}} \leqslant 1 \quad \mathfrak{m}\text{-a.e.}. \tag{11.3.20}$$

We split the proof in two steps.

Step 1: it is sufficient to prove that, for every $h \in \mathbb{N}$, the function $F_h : \mathcal{P}_2(H) \to \mathbb{R}$

$$F_h(\mu) := W_2(\hat{\pi}^h_\sharp\nu, \hat{\pi}^h_\sharp\mu) \quad \text{satisfies} \quad |DF_h|_{\star,\mathscr{A}} \leqslant 1 \quad \mathfrak{m}\text{-a.e.} \tag{11.3.21}$$

In fact, using the continuity property of the Wasserstein distance, it is clear that for every $\mu \in \mathcal{P}_2(H)$

$$\lim_{n\to\infty} F_h(\mu) = F(\mu), \tag{11.3.22}$$

so that it is enough to apply Theorem 10.1.2(1)-(3) to obtain (11.3.20).

Step 2: Let $h \in \mathbb{N}$ be fixed and let us denote by $W_{2,h}$ the Wasserstein distance on $\mathcal{P}_2(\mathbb{R}^h)$; it is easy to check that

$$W_{2,h}(\pi^h_\sharp\mu_0, \pi^h_\sharp\mu_1) = W_2(\hat{\pi}^h_\sharp\mu_0, \hat{\pi}^h_\sharp\mu_1) \quad \text{for every } \mu_0, \mu_1 \in \mathcal{P}_2(H).$$

Thus, if we define the function $\tilde{F}_h : \mathcal{P}_2(\mathbb{R}^h) \to \mathbb{R}$ as

$$\tilde{F}_h(\mu) := W_{2,h}(\pi^h_\sharp\nu, \mu)$$

we get that

$$F_h(\mu) = \tilde{F}_h(\pi^h_\sharp\mu).$$

We also introduce the measure $\mathfrak{m}_h \in \mathcal{P}(\mathcal{P}_2(\mathbb{R}^h))$ which is the push-forward of $\mathfrak{m}$ through the (1-Lipschitz) map $P^h : \mathcal{P}_2(H) \to \mathcal{P}_2(\mathbb{R}^h)$ defined as $P^h(\mu) := \pi^h_\sharp\mu$. By Theorem 11.1.11 applied to $H^{1,2}(\mathcal{P}_2(\mathbb{R}^h), W_{2,h}, \mathfrak{m}_h)$, we can find a sequence of cylindrical functions $\tilde{F}_{h,n} \in \mathrm{Cyl}\left(\mathcal{P}_2(\mathbb{R}^h)\right)$, $n \in \mathbb{N}$, such that

$$\tilde{F}_{h,n} \to \tilde{F}_h \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^h), \mathfrak{m}_h), \tag{11.3.23}$$

$$\mathrm{lip}_{\mathcal{P}_2(\mathbb{R}^h)}\tilde{F}_{h,n} \to g_h \quad \text{in } L^2(\mathcal{P}_2(\mathbb{R}^h), \mathfrak{m}_h) \quad \text{with } g_h \leqslant 1 \ \mathfrak{m}_h\text{-a.e.} \tag{11.3.24}$$

We thus consider the functions $F_{h,n} \in \mathrm{Cyl}\,(\mathcal{P}_2(H))$ defined as in (11.3.16) by

$$F_{h,n}(\mu) := \tilde{F}_{h,n}(\pi^h_\sharp\mu) = \tilde{F}_{h,n}(P^h(\mu)) \quad \text{for every } \mu \in \mathcal{P}_2(H). \tag{11.3.25}$$

We immediately have that $F_{h,n} \to F_h$ in $L^2(\mathcal{P}_2(H), \mathfrak{m})$; on the other hand, (11.3.17) and (11.3.18) yield

$$\mathrm{lip}F_{h,n}(\mu) = \mathrm{lip}_{\mathcal{P}_2(\mathbb{R}^h)}\tilde{F}_{h,n}(P^h(\mu)$$

so that

$$\mathrm{lip}F_{h,n} \to g_h \circ P^h \quad \text{in } L^2(\mathcal{P}_2(H), \mathfrak{m})$$

and $g_h \circ P^h \leqslant 1$ $\mathfrak{m}$-a.e. in $\mathcal{P}_2(H)$. By Theorem 10.1.2(1)-(3), we obtain (11.3.21), concluding the proof. □

Remark 11.3.4. We remark that the results in Sections 11.1.2 and 11.2.1 can be extended to $\mathcal{P}_2(\mathbb{M})$ and $\mathcal{P}_2(H)$ in an analogous way.

BIBLIOGRAPHY

[1] Martial Agueh. “Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory.” In: *Adv. Differential Equations* 10.3 (2005), pp. 309–360.

[2] Sergio Albeverio, Yuri Kondratiev, and Michael Röckner. “Analysis and geometry on configuration spaces: the Gibbsian case.” In: *J. Funct. Anal.* 157.1 (1998), pp. 242–291.

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

[4] Luigi Ambrosio, Elia Brué, and Daniele Semola. *Lectures on optimal transport*. Vol. 130. Unitext. La Matematica per il 3+2. Springer, Cham, 2021, pp. ix+250.

[5] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Second edition. Lectures in Mathematics ETH Zürich. Basel: Birkhäuser Verlag, 2008, pp. x+334.

[6] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. “Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces.” In: *Rev. Mat. Iberoam.* 29.3 (2013), pp. 969–996.

[7] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. “Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below.” In: *Invent. Math.* 195.2 (2014), pp. 289–391.

[8] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. “Metric measure spaces with Riemannian Ricci curvature bounded from below.” In: *Duke Math. J.* 163.7 (2014), pp. 1405–1490.

[9] Luigi Ambrosio and Aldo Pratelli. “Existence and stability results in the L^1 theory of optimal transportation.” In: *Optimal transportation and applications (Martina Franca, 2001)*. Vol. 1813. Lecture Notes in Math. Springer, Berlin, 2003, pp. 123–160.

[10] Hedy Attouch. “Familles d’opérateurs maximaux monotones et mesurabilite.” In: *Ann. Mat. Pura Appl. (4)* 120 (1979), pp. 35–111.

[11] Hedy Attouch. *Variational convergence for functions and operators*. Boston, MA: Pitman (Advanced Publishing Program), 1984, pp. xiv+423.

[12] Jean-Pierre Aubin and Arrigo Cellina. *Differential Inclusions. Set-Valued Maps and Viability Theory*. Springer-Verlag Berlin Heidelberg, 1984, pp. xx+342.

[13] Jean-Pierre Aubin and Hélène Frankowska. *Set-valued analysis*. Modern Birkhäuser Classics. Reprint of the 1990 edition [MR1048347]. Birkhäuser Boston Inc., 2009, pp. xx+461.

[14] Jean-David Benamou. "Numerical resolution of an "unbalanced" mass transport problem." In: *M2AN Math. Model. Numer. Anal.* 37.5 (2003), pp. 851–868.

[15] Jean-David Benamou and Yann Brenier. "A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem." In: *Numer. Math.* 84.3 (2000), pp. 375–393.

[16] Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré. "Iterative Bregman projections for regularized transportation problems." In: *SIAM J. Sci. Comput.* 37.2 (2015).

[17] Jean-David Benamou, Guillaume Carlier, Simone Di Marino, and Luca Nenna. "An entropy minimization approach to second-order variational mean-field games." In: *Math. Models Methods Appl. Sci.* 29.8 (2019), pp. 1553–1583.

[18] Jean-David Benamou, Guillaume Carlier, and Luca Nenna. "A numerical method to solve multi-marginal optimal transport problems with Coulomb cost." In: *Splitting methods in communication, imaging, science, and engineering*. Sci. Comput. Springer, Cham, 2016, pp. 577–601.

[19] Anders Björn and Jana Björn. *Nonlinear potential theory on metric spaces*. Vol. 17. EMS Tracts in Mathematics. European Mathematical Society (EMS), Zürich, 2011, pp. xii+403.

[20] Vladimir I. Bogachev. *Measure Theory*. Vol. II. Springer Berlin Heidelberg, 2006.

[21] Benoît Bonnet and Hélène Frankowska. "Differential inclusions in Wasserstein spaces: the Cauchy-Lipschitz framework." In: *J. Differential Equations* 271 (2021), pp. 594–637.

[22] Nicolas Bouleau and Francis Hirsch. *Dirichlet forms and analysis on Wiener spaces*. Vol. 14. De Gruyter studies in Mathematics. De Gruyter, 1991.

[23] Yann Brenier. "Polar factorization and monotone rearrangement of vector-valued functions." In: *Comm. Pure Appl. Math.* 44.4 (1991), pp. 375–417.

[24] Yann Brenier and Wilfrid Gangbo. "L^p approximation of maps by diffeomorphisms." In: *Calc. Var. Partial Differential Equations* 16.2 (2003), pp. 147–164.

[25] Haim Brezis. "Propriétés régularisantes de certains semi-groupes non linéaires." In: *Israel J. Math.* 9 (1971), pp. 513–534.

[26] Haim Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50). North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973, pp. vi+183.

[27] Martin Burger, Marzena Franek, and Carola-Bibiane Schönlieb. "Regularized Regression and Density Estimation based on Optimal Transport." In: *Applied Mathematics Research Express* 2012 (2012), pp. 209–253.

[28] Luis A. Caffarelli. "Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation." In: *Ann. of Math. (2)* 131.1 (1990), pp. 135–150.

[29] Luis A. Caffarelli, Mikhail Feldman, and Robert J. McCann. "Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs." In: *J. Amer. Math. Soc.* 15.1 (2002), pp. 1–26.

[30] Luis A. Caffarelli and Robert J. McCann. "Free boundaries in optimal transport and Monge-Ampère obstacle problems." In: *Ann. of Math. (2)* 171.2 (2010), pp. 673–730.

[31] Fabio Camilli, Giulia Cavagnari, Raul De Maio, and Benedetto Piccoli. "Superposition principle and schemes for measure differential equations." In: *Kinet. Relat. Models* 14.1 (2021), pp. 89–113.

[32] Guillaume Carlier, Vincent Duval, Gabriel Peyré, and Bernhard Schmitzer. "Convergence of entropic schemes for optimal transport and gradient flows." In: *SIAM J. Math. Anal.* 49.2 (2017), pp. 1385–1418.

[33] Giulia Cavagnari, Stefano Lisini, Carlo Orrieri, and Giuseppe Savaré. "Lagrangian, Eulerian and Kantorovich formulations of multi-agent optimal control problems: equivalence and gamma-convergence." In: *J. Differential Equations* 322 (2022), pp. 268–364.

[34] Giulia Cavagnari, Giuseppe Savaré, and Giacomo E. Sodini. "Dissipative probability vector fields and generation of evolution semigroups in Wasserstein spaces." Accepted for publication in *Probability Theory and Related Fields*, May 2022.

[35] Jeff Cheeger. "Differentiability of Lipschitz functions on metric measure spaces." In: *Geom. Funct. Anal.* 9.3 (1999), pp. 428–517.

[36] Lénaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. "An interpolating distance between optimal transport and Fisher-Rao metrics." In: *Found. Comput. Math.* 18.1 (2018), pp. 1–44.

[37] Lénaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. "Unbalanced optimal transport: dynamic and Kantorovich formulations." In: *J. Funct. Anal.* 274.11 (2018), pp. 3090–3123.

[38] Michael G. Crandall and Lawrence C. Evans. "On the relation of the operator $\partial/\partial s + \partial/\partial\tau$ to evolution governed by accretive operators." In: *Israel J. Math.* 21.4 (1975), pp. 261–278.

[39] Michael G. Crandall and Thomas M. Liggett. "Generation of semi-groups of nonlinear transformations on general Banach spaces." In: *Amer. J. Math.* 93 (1971), pp. 265–298.

[40] Gianni Dal Maso. *An Introduction to Γ-Convergence*. Vol. 8. Progress in Nonlinear Differential Equations and Their Applications. Boston: Birkhäuser, 1993.

[41] Ennio De Giorgi. "New problems on minimizing movements." In: *Boundary value problems for partial differential equations and applications*. Vol. 29. RMA Res. Notes Appl. Math. Masson, Paris, 1993, pp. 81–98.

[42] Nicolò De Ponti. "Metric properties of homogeneous and spatially inhomogeneous F-divergences." In: *IEEE Trans. Inform. Theory* 66.5 (2020), pp. 2872–2890.

[43] Claude Dellacherie and Paul-Andre Meyer. *Probabilités et potentiel. Chap. I à IV. Ed. entièrement refondue*. French. 1975.

[44] Lorenzo Dello Schiavo. "A Rademacher-type theorem on L^2-Wasserstein spaces over closed Riemannian manifolds." In: *J. Funct. Anal.* 278.6 (2020), pp. 108397, 57.

[45] Simone Di Marino and Augusto Gerolin. "An optimal transport approach for the Schrödinger bridge problem and convergence of Sinkhorn algorithm." In: *J. Sci. Comput.* 85.2 (2020), Paper No. 27, 28.

[46] Ivar Ekeland and Roger Témam. *Convex analysis and variational problems*. 1st ed. Classics in Applied Mathematics. Society for Industrial Mathematics, 1987.

[47] Lawrence C. Evans and Wilfrid Gangbo. "Differential equations methods for the Monge-Kantorovich mass transfer problem." In: *Mem. Amer. Math. Soc.* 137.653 (1999), pp. viii+66.

[48] Alessio Figalli. "The optimal partial transport problem." In: *Arch. Ration. Mech. Anal.* 195.2 (2010), pp. 533–560.

[49] Alessio Figalli and Federico Glaudo. *An invitation to optimal transport, Wasserstein distances, and gradient flows*. EMS Textbooks in Mathematics. EMS Press, Berlin, 2021, pp. vi+136.

[50] Gerald B. Folland. *Real analysis*. Second edition. Pure and Applied Mathematics (New York). Modern techniques and their applications, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999, pp. xvi+386.

[51] István S. Gál. "On the fundamental theorems of the calculus." In: *Trans. Amer. Math. Soc.* 86 (1957), pp. 309–320.

[52] Wilfrid Gangbo and Robert J. McCann. "The geometry of optimal transportation." In: *Acta Math.* 177.2 (1996), pp. 113–161.

[53] Nicola Gigli. "On the inverse implication of Brenier-McCann theorems and the structure of $(\mathscr{P}_2(M), W_2)$." In: *Methods Appl. Anal.* 18.2 (2011), pp. 127–158.

[54] Nicola Gigli. "On the differential structure of metric measure spaces and applications." In: *Mem. Amer. Math. Soc.* 236.1113 (2015), pp. vi+91.

[55] Nicola Gigli. "Nonsmooth differential geometry—an approach tailored for spaces with Ricci curvature bounded from below." In: *Mem. Amer. Math. Soc.* 251.1196 (2018), pp. v+161.

[56] Nicola Gigli, Kazumasa Kuwada, and Shin-Ichi Ohta. "Heat flow on Alexandrov spaces." In: *Comm. Pure Appl. Math.* 66.3 (2013), pp. 307–331.

[57] Nicola Gigli and Enrico Pasqualetto. *Lectures on nonsmooth differential geometry*. Vol. 2. SISSA Springer Series. Springer, Cham, 2020, pp. xi+204.

[58] Kevin Guittet. "Extended Kantorovich norms: a tool for optimization." In: *Tech. Rep. 4402, INRIA* (2002).

[59] Leonid G. Hanin. "Kantorovich-Rubinstein norm and its application in the theory of Lipschitz spaces." In: *Proc. Amer. Math. Soc.* 115.2 (1992), pp. 345–352.

[60] Leonid G. Hanin. "An extension of the Kantorovich norm." In: *Monge Ampère equation: applications to geometry and optimization (Deerfield Beach, FL, 1997)*. Vol. 226. Contemp. Math. Amer. Math. Soc., Providence, RI, 1999, pp. 113–130.

[61] Juha Heinonen and Pekka Koskela. "Quasiconformal maps in metric spaces with controlled geometry." In: *Acta Math.* 181.1 (1998), pp. 1–61.

[62] Juha Heinonen, Pekka Koskela, Nageswari Shanmugalingam, and Jeremy T. Tyson. *Sobolev spaces on metric measure spaces*. Vol. 27. New Mathematical Monographs. An approach based on upper gradients. Cambridge University Press, Cambridge, 2015, pp. xii+434.

[63] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex analysis and minimization algorithms. I*. Vol. 305. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Fundamentals. Springer-Verlag, Berlin, 1993, pp. xviii+417.

[64] Richard Jordan, David Kinderlehrer, and Felix Otto. "The variational formulation of the Fokker-Planck equation." In: *SIAM J. Math. Anal.* 29.1 (1998), pp. 1–17.

[65] Leonid Kantorovich and Gennady S. Rubinstein. "On a space of totally additive functions." In: *Vestnik Leningrad. Univ* 13 (1958), pp. 52–59.

[66] Leonid Kantorovitch. "On the translocation of masses." In: *C. R. (Doklady) Acad. Sci. URSS (N.S.)* 37 (1942), pp. 227–229.

[67] Leonid Kantorovitch. "On a problem of Monge." In: *Uspekhi Mat. Nauk* 3 (1948), pp. 225–226.

[68] Hans G. Kellerer. "Duality theorems for marginal problems." In: *Z. Wahrsch. Verw. Gebiete* 67.4 (1984), pp. 399–432.

[69] Stanislav Kondratyev, Léonard Monsaingeon, and Dmitry Vorotnikov. "A new optimal transport distance on the space of finite Radon measures." In: *Adv. Differential Equations* 21.11-12 (2016), pp. 1117–1164.

[70] Pekka Koskela and Paul MacManus. "Quasiconformal mappings and Sobolev spaces." In: *Studia Math.* 131.1 (1998), pp. 1–17.

[71] Stanislav N. Kružkov. "First order quasilinear equations with several independent variables." In: *Mat. Sb. (N.S.)* 81 (123) (1970), pp. 228–255.

[72] Christian Léonard. "A saddle-point approach to the Monge-Kantorovich optimal transport problem." In: *ESAIM Control Optim. Calc. Var.* 17.3 (2011), pp. 682–704.

[73] Christian Léonard. "From the Schrödinger problem to the Monge-Kantorovich problem." In: *J. Funct. Anal.* 262.4 (2012), pp. 1879–1920.

[74] Christian Léonard. "A survey of the Schrödinger problem and some of its connections with optimal transport." In: *Discrete Contin. Dyn. Syst.* 34.4 (2014), pp. 1533–1574.

[75] Matthias Liero, Alexander Mielke, and Giuseppe Savaré. "Optimal transport in competition with reaction: the Hellinger-Kantorovich distance and geodesic curves." In: *SIAM J. Math. Anal.* 48.4 (2016), pp. 2869–2911.

[76] Matthias Liero, Alexander Mielke, and Giuseppe Savaré. "Optimal entropy-transport problems and a new Hellinger-Kantorovich distance between positive measures." In: *Invent. Math.* 211.3 (2018), pp. 969–1117.

[77] Damiano Lombardi and Emmanuel Maitre. "Eulerian models and algorithms for unbalanced optimal transport." In: *ESAIM Math. Model. Numer. Anal.* 49.6 (2015), pp. 1717–1744.

[78] John Lott. "Some geometric calculations on Wasserstein space." In: *Comm. Math. Phys.* 277.2 (2008), pp. 423–437.

[79] John Lott and Cédric Villani. "Ricci curvature for metric-measure spaces via optimal transport." In: *Ann. of Math. (2)* 169.3 (2009), pp. 903–991.

[80] Jan Maas, Martin Rumpf, Carola Schönlieb, and Stefan Simon. "A generalized model for optimal transport of images including dissipation and density modulation." In: *ESAIM Math. Model. Numer. Anal.* 49.6 (2015), pp. 1745–1769.

[81] Marie-Paule Malliavin and Paul Malliavin. "An infinitesimally quasi-invariant measure on the group of diffeomorphisms of the circle." In: *Special functions (Okayama, 1990)*. ICM-90 Satell. Conf. Proc. Springer, Tokyo, 1991, pp. 234–244.

[82] Robert J. McCann. "A convexity principle for interacting gases." In: *Adv. Math.* 128.1 (1997), pp. 153–179.

[83] Gaspard Monge. *Mémoire sur la théorie des déblais et des remblais*. De l'Imprimerie Royale, 1781.

[84] James Munkres. *Topology*. 2nd ed. Prentice Hall, Inc, 2000.

[85] Matteo Muratori and Giuseppe Savaré. "Gradient flows and evolution variational inequalities in metric spaces. I: Structural properties." In: *J. Funct. Anal.* 278.4 (2020), pp. 108347, 67.

[86] Emanuele Naldi and Giuseppe Savaré. "Weak topology and Opial property in Wasserstein spaces, with applications to gradient flows and proximal point algorithms of geodesically convex functionals." In: *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* 32.4 (2021), pp. 725–750.

[87] John Nash. "C^1 isometric imbeddings." In: *Ann. of Math. (2)* 60 (1954), pp. 383–396.

[88] Luca Natile and Giuseppe Savaré. "A Wasserstein approach to the one-dimensional sticky particle system." In: *SIAM J. Math. Anal.* 41.4 (2009), pp. 1340–1365.

[89] Ricardo H. Nochetto and Giuseppe Savaré. "Nonlinear evolution governed by accretive operators in Banach spaces: error control and applications." In: *Math. Models Methods Appl. Sci.* 16.3 (2006), pp. 439–477.

[90] Felix Otto. "The geometry of dissipative evolution equations: the porous medium equation." In: *Comm. Partial Differential Equations* 26.1-2 (2001), pp. 101–174.

[91] Ofir Pele and Michael Werman. "A Linear Time Histogram Metric for Improved SIFT Matching." In: Oct. 2008, pp. 495–508.

[92] Gabriel Peyré and Marco Cuturi. "Computational Optimal Transport." In: *Foundations and Trends in Machine Learning* 11 (5-6) (2019), pp. 355–602.

[93] Benedetto Piccoli. "Measure differential inclusions." In: *2018 IEEE Conference on Decision and Control (CDC)* (2018), pp. 1323–1328.

[94] Benedetto Piccoli. "Measure differential equations." In: *Arch. Ration. Mech. Anal.* 233.3 (2019), pp. 1289–1317.

[95] Benedetto Piccoli and Francesco Rossi. "Generalized Wasserstein distance and its application to transport equations with source." In: *Arch. Ration. Mech. Anal.* 211.1 (2014), pp. 335–358.

[96] Benedetto Piccoli and Francesco Rossi. "On properties of the generalized Wasserstein distance." In: *Arch. Ration. Mech. Anal.* 222.3 (2016), pp. 1339–1365.

[97] Aldo Pratelli. "On the equality between Monge's infimum and Kantorovich's minimum in optimal mass transportation." In: *Annales de l'Institut Henri Poincare (B) Probability and Statistics* 43 (Feb. 2007), pp. 1–13.

[98] Liqun Qi. "Uniqueness of the maximal extension of a monotone operator." English. In: *Nonlinear Anal., Theory Methods Appl.* 7 (1983), pp. 325–332.

[99] Svetlozar T. Rachev and Ludger Rüschendorf. *Mass transportation problems. Vol. I.* Probability and its Applications (New York). Theory. Springer-Verlag, New York, 1998, pp. xxvi+508.

[100] Svetlozar T. Rachev and Ludger Rüschendorf. *Mass transportation problems. Vol. II.* Probability and its Applications (New York). Applications. Springer-Verlag, New York, 1998, pp. xxvi+430.

[101] Max-K. von Renesse and Karl-Theodor Sturm. "Entropic measure and Wasserstein diffusion." In: *Ann. Probab.* 37.3 (2009), pp. 1114–1191.

[102] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Landmarks in Mathematics. Reprint of the 1970 original, Princeton Paperbacks. Princeton University Press, Princeton, NJ, 1997, pp. xviii+451.

[103] Michael Röckner and Alexander Schied. "Rademacher's theorem on configuration spaces and applications." In: *J. Funct. Anal.* 169.2 (1999), pp. 325–356.

[104] Halsey L. Royden. *Real analysis*. Third edition. Macmillan Publishing Company, New York, 1988, pp. xx+444.

[105] Yossi Rubner, Leonidas Guibas, and Carlo Tomasi. "The earth mover's distance, multi-dimensional scaling, and color-based image retrieval." In: *Proceedings of the ARPA Image Understanding Workshop* (1997).

[106] Jim Rulla. "Error analysis for implicit approximations to solutions to Cauchy problems." In: *SIAM J. Numer. Anal.* 33.1 (1996), pp. 68–87.

[107] Filippo Santambrogio. *Optimal transport for applied mathematicians*. Vol. 87. Progress in Nonlinear Differential Equations and their Applications. Calculus of variations, PDEs, and modeling. Birkhäuser/Springer, Cham, 2015, pp. xxvii+353.

[108] Giuseppe Savaré. "Sobolev Spaces in Extended Metric-Measure Spaces." In: *New Trends on Analysis and Geometry in Metric Spaces*. Springer, 2022, pp. 117–276.

[109] Giuseppe Savaré and Giacomo E. Sodini. "A simple relaxation approach to duality for Optimal Transport problems in completely regular spaces." In: *Journal of Convex Analysis* 29.1 (2022), pp. 1–12.

[110] Laurent Schwartz. *Radon measures on arbitrary topological spaces and cylindrical measures*. OUP, Tata Institute Monographs on Mathematics, 1973.

[111] Nageswari Shanmugalingam. "Newtonian spaces: an extension of Sobolev spaces to metric measure spaces." In: *Rev. Mat. Iberoamericana* 16.2 (2000), pp. 243–279.

[112] Daniel W. Stroock. *Probability theory*. Second edition. An analytic view. Cambridge University Press, Cambridge, 2011, pp. xxii+527.

[113] Karl-Theodor Sturm. "On the geometry of metric measure spaces. I." In: *Acta Math.* 196.1 (2006), pp. 65–131.

[114] Karl-Theodor Sturm. "On the geometry of metric measure spaces. II." In: *Acta Math.* 196.1 (2006), pp. 133–177.

[115] Karl-Theodor Sturm. "Entropic measure on multidimensional spaces." In: *Seminar on Stochastic Analysis, Random Fields and Applications VI*. Vol. 63. Progr. Probab. Birkhäuser/Springer Basel AG, Basel, 2011, pp. 261–277.

[116] Flemming Topsoe. "Compactness in spaces of measures." In: *Studia Mathematica* 36 (1970), pp. 195–212.

[117] Neil S. Trudinger and Xu-Jia Wang. "On the Monge mass transfer problem." In: *Calc. Var. Partial Differential Equations* 13.1 (2001), pp. 19–31.

[118] Cédric Villani. *Topics in optimal transportation*. Vol. 58. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2003, pp. xvi+370.

[119] Cédric Villani. *Optimal transport. Old and new*. Vol. 338. Grundlehren der Mathematischen Wissenschaften. Berlin: Springer-Verlag, 2009, pp. xxii+973.