Scuola Normale Superiore, Pisa
Classe di Scienze Matematiche, Fisiche e Naturali

SOME APPLICATIONS OF OPTIMAL TRANSPORT THEORY TO EVOLUTION AND SHAPE OPTIMIZATION PROBLEMS

PhD Thesis

EDOARDO MAININI

ADVISOR:
Prof. LUIGI AMBROSIO

2007 - 2009

Contents

Introduction

Monge's formulation of the optimal transport problem, going back to the eighteenth century, can be described as follows (the original reference is [55]). We are given two distribution functions f_1 and f_2, corresponding to the height profile of a certain amount of material in a region Ω_1 and to an excavation to be filled in a region Ω_2. We are given a cost c on $\Omega_1 \times \Omega_2$, accounting for the work to be spent for transporting a unit of mass from $x \in \Omega_1$ to $y \in \Omega_2$. Let $\mathbf{t} : \Omega_1 \to \Omega_2$ be the (Borel) rule of transportation, that is, the map which identifies which arrival point $y \in \Omega_2$ corresponds to a starting point $x \in \Omega_1$. The quantity

$$\int_{\Omega_1} c(x, \mathbf{t}(x)) f_1(x) dx$$

represents the total cost for the operation. The optimization problem consists in finding the best way (i.e. the best $\mathbf{t}$) to move the material, so that the total cost is minimized. Of course, the conservation of mass constraint has to be taken into account, that is

$$\int_{\Omega_1} f_1(x)\, dx = \int_{\Omega_2} f_2(y)\, dy.$$

Moreover the mass corresponding to each portion of the material moved is conserved too, that is

$$\int_A f_2(y)\, dy = \int_{\mathbf{t}^{-1}(A)} f_1(x)\, dx,$$

for any (Borel) subset A of Ω_2. The latter conditions will be summarized with the notation $\mathbf{t}_\# f_1 = f_2$. Hence, we can give the following formulation:

$$\min \left\{ \int_{\Omega_1} c(x, \mathbf{t}(x)) f_1(x)\, dx \;:\; \mathbf{t}_\# f_1 = f_2 \right\}.$$

From the mathematical point of view, it is natural to state the optimal transport problem in a measure setting. Instead of density functions f_1, f_2 supported in some regions, we can consider two measures μ_1, μ_2 over a suitable ambient space X. The conservation of total mass (then normalized to 1 without loss of generality) is taken into account by asking μ_1, μ_2 to be probability measures over X, that is $\mu_1, \mu_2 \in \mathscr{P}(X)$.

This problem constitutes the basic tool of our analysis. In its more general form, due to Kantorovich (we refer to Chapter 2 for details), it gives rise to the notion of optimal transportation distance in the space of probability measures. This way, two measures can be seen as close if the relative transport cost is small. The nice structure induced is one of the most relevant features, since $\mathscr{P}(X)$ turns out to be a suitable setting for developing not only a metric, but even a differential calculus.

The main topic in this work will be the analysis of evolutionary equations (we refer to Chapters 3,4,5) with optimal transportation tools. In this case, De Giorgi's minimizing movements theory and the optimal transport can be combined to provide general existence and stability results. Also, it turns out that solutions to equations of diffusion type are seen to correspond to curves of maximal slope of suitable energy functionals in $\mathscr{P}(X)$, and an important parallel variational formulation is established. This is the approach introduced in the papers [43, 57, 58, 59, 60], and then developed by many other authors, for instance in [54, 26, 27, 1, 8, 9].

Of course, there are many other applications of optimal transport. In particular, we will also discuss some different formulations, related to "optimal transport networks". See [21, 23] and the other references in Chapter 6, where a problem of this kind is analyzed.

We go into details with the following summary.

The first three chapters are meant to give a short, but as much as possible self-contained, presentation of the theoretic background that is needed for the applications starting from Chapter 4. The main reference is the book [4].

Chapter 1 introduces the measure theoretic framework and the other mathematical tools. We consider the probability space $\mathscr{P}(X)$. Regarding X, we will assume it to be a separable metric space satisfying the Radon property (see (1.0.1) below). For the moment, we let X be a Hilbert space with norm $|\cdot|$ and orthonormal basis $\{\mathbf{e}_j\}_{j\in\mathbb{N}}$. We introduce the basic concepts like the moments $\int_X |x|^p\,d\mu$ of μ, and the corresponding subspace of probability measures with finite p-moments $\mathscr{P}_p(X)$. The standard narrow convergence of measures

$$\int_X \varphi\,d\mu_n \to \int_X \varphi\,d\mu \quad \forall\varphi\in C_b^0(X)$$

is considered over $\mathscr{P}(X)$, and denoted by $\mu_n \rightharpoonup \mu$. One of the basic tools is Prokhorov theorem, characterizing compactness by means of tightness, that is, uniform approximation by measures of compact sets of X. We refer to the monographs [3, 13, 14, 33] and to the standard real analysis texts like [63].

We then introduce the first important concepts related to optimal transportation. A Borel (or even μ_1-measurable) map $\mathbf{t}: X \to X$ such that $\mathbf{t}_\#\mu_1 = \mu_2$, i.e. $\mu_2(A) = \mu_1(\mathbf{t}^{-1}(A))$, is called a *transport map*. A measure $\gamma\in\mathscr{P}(X\times X)$ such that the first and second marginals are respectively μ_1 and μ_2 (i.e. $\gamma(A\times X)=\mu_1(A)$, $\gamma(X\times A)=\mu_2(A)$) is said to be a *transport plan*. We let

$$\Gamma(\mu_1,\mu_2) := \left\{\gamma\in\mathscr{P}(X\times X) : \pi^1_\#\gamma=\mu_1, \pi^2_\#\gamma=\mu_2\right\}.$$

Plans constitute the basic ingredient for generalizing the Monge problem and constructing a metric and differential structure on the whole of $\mathscr{P}_p(X)$, with no restriction on the measures involved.

Next we introduce the standard disintegration theorem and the cylindrical projections for measures and functions on X. Hence, for $\mu \in \mathscr{P}(X)$ and a μ-measurable function f, we have the slicing formula

$$\int_X f(x)\, d\mu = \int_{X^d} \left(\int_{(\Pi^d)^{-1}(x_d)} f(x)\, d\mu_{x_d}(x) \right) d\Pi^d_{\#}\mu(x_d).$$

Here $x_d \in X^d$ and X^d denotes a d-dimensional subspace, Π^d being the corresponding orthogonal projection. The internal integral is taken on the fiber $(\Pi^d)^{-1}(x_d)$, with respect to the disintegrated measure μ_{x_d}.

We also let $\mathrm{Cyl}(X)$ denote the set of smooth cylindrical functions over X, i.e. $\zeta(x) \in \mathrm{Cyl}(X)$ is smooth and depends on a finite number of components of x.

In order to gain compactness in X, we need to introduce a weaker topology, but still metrizable. For, we consider (as in [4]) the norm

$$\|x\|^2_{\varpi} = \sum_{j=1}^{\infty} \frac{1}{j^2} \langle x, \mathbf{e}_j \rangle^2.$$

X_ϖ denotes the space X endowed with the new topology, which induces a weaker topology also on $\mathscr{P}(X)$, to which we refer as $\mathscr{P}(X_\varpi)$.

We then prove some convergence lemmas involving transport maps and plans. We analyze the behavior of sequences of the form $(\rho_{n\#}\mu_n)$. Particular importance has to be paid to sequences (ρ_n) of functions with $\rho_n \in L^p(\mu_n)$, so that each element belongs to a different space. We define the convergence by duality with cylindrical functions, letting $\rho_n \to \rho$ if

$$\int_X \zeta(x)\rho_n(x)\, d\mu_n(x) \to \int_X \zeta(x)\rho(x)\, d\mu(x) \quad \forall \zeta \in \mathrm{Cyl}(X).$$

This notion of convergence will be important when dealing with finite dimensional approximations.

In the last section we introduce some tools from geometric measure theory. We refer for instance to [3, 35, 36]. First of all we recall the definition of Hausdorff measures $\mathcal{H}^k$ on $\mathbb{R}^n$ and the definition of $\mathcal{H}^k$-rectifiable set. Then we introduce the approximate tangent space, the tangent gradient and the distributional curvature, along with a result about the differentiation of the $\mathcal{H}^k$ measure of sets along smooth vector fields.

Chapter 2 deals with the rigorous formulation of the optimal transport problem. For the theory in the first two sections, we refer to the results in [15, 16] and to the monographs [72, 73]. In particular, we introduce the generalized Kantorovich version ([44, 45]), that is

$$\inf \left\{ \int_{X\times Y} c(x,y)\, d\gamma(x,y) : \gamma \in \Gamma(\mu_1, \mu_2) \right\}.$$

Unlike the standard Monge problem, this linearized version is well-posed, as easily seen by direct methods. As a consequence, we can define the set on which the infimum is attained $\Gamma_0(\mu_1, \mu_2)$, and refer to its elements as optimal transport plans. It is clear that this formulation contains the former, since any map $\mathbf{t}$ corresponds to the plan $(\mathbf{I}, \mathbf{t})_\# \mu_1$, where $(\mathbf{I}, \mathbf{t}) : X \to X \times X$ is the product map. Of course, if we are to transport a Dirac mass to a diffuse mass, the map from the support of the Dirac mass to the support of the target one has to be multivalued, making the Monge problem ill-posed. On the other hand, it is clear that no such restriction arises when dealing with plans, which include also multivalued cases. In particular, plans correspond to maps when they are concentrated on graphs.

We briefly recall some general results, like Kantorovich duality and c-monotonicity of the support of optimal plans. We do not enter in the details, and we address the reader to [72, 73], or also to [2, 4] for an exhaustive overview on optimal transportation. Here we focus the attention on the case of the p-cost, that is $c(x, y) = |x - y|^p$, $p > 1$, and we state the result about existence of optimal maps. After the heuristics above, we see that the only obstacle could be the concentration of the starting measure. Hence we let μ_1 be a *regular* measure, that is, null on Gauss null sets, or simply absolutely continuous with respect to Lebesgue measure in the Euclidean case. The classical result says that the optimal transport between μ_1 and μ_2 is induced by a unique map. In the case $p = 2$ on Euclidean spaces, the map is also given by the gradient of a convex l.s.c. function (Brenier's theorem, see [15, 16]). We also show an injectivity result.

In the second part of the chapter we introduce the optimal transportation distance. It is defined on $\mathscr{P}_p(X)$ by

$$W_p^p(\mu, \nu) = \inf \left\{ \int_{X \times X} d^p(x_1, x_2)\, d\gamma(x_1, x_2) : \gamma \in \Gamma(\mu, \nu) \right\}.$$

We need a lemma about composition of plans in order to show the triangle inequality. The argument is based on the disintegration theorem recalled in Chapter 1. After that, we prove some other facts about the distance. It metrizes the natural weak topology of $\mathscr{P}_p(X)$, that is

$$W_p(\mu_n, \mu) \to 0 \quad \Leftrightarrow \quad \mu_n \rightharpoonup \mu \text{ and } \int_X |x|^p \, d\mu_n \to \int_X |x|^p \, d\mu,$$

so that from now on $(\mathscr{P}_p(X), W_p)$ becomes the ambient space for the analysis. Thanks to this result, we are able to state some refined convergence lemmas, in particular improving the ones of Chapter 1. Most importantly, we characterize the Wasserstein (constant speed) geodesics, that is, curves $t \in [0, 1] \mapsto \mu_t \in \mathscr{P}_2(X)$ such that

$$W_2(\mu_s, \mu_t) = (t - s) W_2(\mu_0, \mu_1) \quad \forall\, 0 \le s \le t \le 1.$$

The characterization says that a curve is a geodesic if and only if there exists a plan $\gamma \in \Gamma_0(\mu_0, \mu_1)$ such that

$$\mu_t = \big((1 - t)\pi^1 + t\pi^2\big)_\# \gamma.$$

The geodesical interpolation between μ_0 and μ_1 possesses good properties. For instance, if μ_0 is regular and μ_t is a geodesic, then μ_t is regular for any $t < 1$, even if μ_1 is not (with the consequent existence of optimal maps).

After introducing the Wasserstein structure, in the third part of the chapter we go on showing a first relation with PDE's. We cite [4, 11, 59] as basic references. In the $(\mathscr{P}_2(X), W_2)$ framework, we will see how any absolutely continuous curve μ_t satisfies the continuity equation

$$\partial_t \mu_t + \operatorname{div}(\mathbf{v}_t \mu_t) = 0 \tag{0.0.1}$$

in the sense of distributions, where $\mathbf{v}_t$ is a suitable $L^2(X, \mu_t; X)$ vector field. The proof will in fact show that $\mathbf{v}_t$ can be chosen as an element of

$$\mathbf{v}_t \in \overline{\{\nabla\varphi : \varphi \in \mathrm{Cyl}(X)\}}^{L^2(X,\mu_t;X)}, \tag{0.0.2}$$

for almost any t. This is a characterization of optimality in norm for the velocity field, and it turns out that this is actually a tangent vector to μ_t in a suitable sense, as argued in [59]. For a more detailed discussion of these geometric properties we refer to [40].

We conclude the chapter devoted to optimal transport by introducing some related formulations, concerning urban planning (see for instance [21, 23, 24]). Thinking to the basic form of Monge problem in $\mathbb{R}^2$, we are given again two measures μ_1, μ_2, with same mass. In this case they represents quantity of specific products, amounts of some material, or densities of population in specific regions. For $\gamma \in \Gamma(\mu_1, \mu_2)$, we introduce the cost

$$I_\Sigma = \int_{\Omega\times\Omega} d_\Sigma(x, y)\, d\gamma(x, y), \tag{0.0.3}$$

where d_Σ is a distance depending on the presence of a transport line Σ, which can be for instance a curve connecting the two reference regions. An example could be

$$d(x, y) \wedge (\operatorname{dist}(x, \Sigma) + \operatorname{dist}(y, \Sigma)),$$

where d is the Euclidean distance and $\operatorname{dist}(x, \Sigma) = \inf\{d(x, \sigma) : \sigma \in \Sigma\}$. Here the goal is to minimize the optimal transport cost with respect to Σ, subject to a constraint accounting for the cost for constructing the network (for instance one may fix the length of Σ).

Chapter 3 is concerned with the differential structure of $(\mathscr{P}_2(X), W_2)$ and with the theory of gradient flows. Here we show the main results to be applied for existence and uniqueness of PDEs having the form of (0.0.1). We refer to [32] for the minimizing movements scheme. We also refer to [53] for the definition of convexity in this framework, and to [43, 58, 59, 60] for the application of the Wasserstein variational approach to evolution equations. The theory covered in this chapter is systematically developed in [4].

We begin with the natural (sub)differential definition in the Wasserstein space. Given a functional $\phi : \mathscr{P}_2(X) \to \mathbb{R}$, we say that $\boldsymbol{\xi} \in L^2(X, \mu; X)$ is in the subdifferential $\partial\phi(\mu)$

of the point μ if

$$\phi(\nu) - \phi(\mu) \geq \int_X \langle \boldsymbol{\xi}, \mathbf{t}_\mu^\nu(x) - x \rangle \, d\mu(x) + o(W_2(\mu, \nu)), \tag{0.0.4}$$

where $\mathbf{t}_\mu^\nu$ denotes the optimal transport map between μ and ν. For this definition to make sense, we need existence of optimal maps. Hence, we ask that any μ in the domain of $\partial\phi$ (such that $\partial\phi(\mu)$ is nonempty) is regular. This assumption is not strictly needed, since we could also define a generalized differential in terms of transport plans, but simplifies much the notation.

Our goal is to extend the natural steepest descent curves equation of Euclidean spaces, that is

$$u' = -\nabla(\phi(u)), \tag{0.0.5}$$

to more general spaces. We consider two strategies. The first one takes advantage of the definition of tangent vector field and of subdifferential in the new context, as given by (0.0.2) and (0.0.4). So that we can write the *gradient flow equation*

$$\mathbf{v}_t \in -\partial\phi(\mu), \tag{0.0.6}$$

where $\mathbf{v}_t$ satisfies (0.0.2). The second one is a purely metric approach, which consists in regarding the Euler implicit discretization (with time step τ) for (0.0.5) as the Euler-Lagrange equation for the minimization problem

$$\min_u \phi(u) + \frac{1}{2\tau}|u - u^0|^2, \quad u^0 \text{ given.}$$

We see that this relation is suitable to be generalized in $(\mathscr{P}_2(X), W_2)$, by

$$\min_{\mu \in \mathscr{P}_2(X)} \phi(\mu) + \frac{1}{2\tau} W_2^2(\mu, \mu^0). \tag{0.0.7}$$

We minimize recursively, producing a sequence (μ_τ^k), $k \in \mathbb{N}$, of discrete minimizers. We construct a piecewise constant interpolation $\mu_\tau(t) := \mu_\tau^{\lceil t/\tau \rceil}$ and we try passing to the limit as $\tau \to 0$, in the sense of measures, for any t, thus obtaining a curve μ_t in $\mathscr{P}_2(X)$. In a Hilbertian setting, this procedure is classical and known as implicit Euler scheme. De Giorgi in [32] studied and generalized this in general metric spaces, calling this minimizing movements scheme. We will show that, under suitable assumptions on functional ϕ (besides lower semicontinuity, coercivity and regularity of measures in the domain of $\partial\phi$), these two approaches produce in fact the same solution.

We refer to two main results in Chapter 3. The first one (Theorem 3.4.4) assumes that the sublevels of ϕ are compact. Then, there exists a solution of the minimizing movements scheme, which also produces a limiting velocity which satisfies a relaxed gradient flow relation. In some particular cases such relation can be shown to coincide with (0.0.6), hence giving a gradient flow.

On the other hand, the most powerful results are obtained in the case of a convex functional. Here the convexity relation has to be understood in the Wasserstein sense, that is, ϕ is λ-geodesically convex if, for any $\mu^1, \mu^2 \in \mathscr{P}_p(X)$, there exist $\gamma \in \Gamma_0(\mu^1, \mu^2)$ such that the inequality

$$\phi\left(((1-t)\pi^1 + t\pi^2)_{\#}\gamma\right) \leq (1-t)\phi(\mu^1) + t\phi(\mu^2) - \frac{1}{2}\lambda t(1-t)W_2^2(\mu^1, \mu^2) \tag{0.0.8}$$

holds for any $t \in [0,1]$. Such notion was introduced in [53]. Actually we will sometimes need a stronger convexity property, we refer to Definition 3.1.4. In the convex case (and even if ϕ does not have compact sublevels) both the minimizing movements scheme and the gradient flow inclusion (0.0.6) inherit many additional properties. For the latter, we mention the fact that it can be translated in a system of *evolution variational inequalities*, that is

$$\frac{1}{2}\frac{d}{dt}W_2^2(\mu_t, \nu) + \frac{1}{2}\lambda W_2^2(\mu_t, \nu) \leq \phi(\nu) - \phi(\mu_t) \quad \text{for } \mathcal{L}^1\text{-a.e. } t \in (0,T),$$

for any $\nu \in D(\phi)$. This characterization holds only in the distinguished case $p = 2$, and we point out that it is the basic tool for uniqueness and contractivity of gradient flows (see Theorem 3.3.4). Moreover, we are able to give a characterization for the minimal norm in $\partial\phi(\mu)$, as

$$\limsup_{\nu \to \mu \text{ in} \mathscr{P}_2(X)} \frac{(\phi(\mu) - \phi(\nu))^+}{W_2(\mu, \nu)} = \min\left\{\|\boldsymbol{\xi}\|_{L^2(X,\mu;X)} : \boldsymbol{\xi} \in \partial\phi(\mu)\right\}$$

The left hand side is called the *metric slope* of the functional. Concerning the minimizing movements scheme in the convex case, if μ_τ is a discrete minimizer, the element of minimal norm of $\partial\phi(\mu)$ can be suitably approximated with subdifferentials at the points μ_τ, as $\tau \to 0$ (a closure property of the subdifferential). Moreover, strong convexity yields uniqueness of each discrete minimizer. Finally, the main result asserts that there is a unique gradient flow/minimizing movement for ϕ, starting from each point $\mu \in \mathscr{P}_2(X)$, and that such solution is given by the action of a contractive semigroup $S(t)$ on $\mathscr{P}_2(X)$. More properties are given in the statement, see Theorem 3.5.8.

Chapter 4 gives the first new application of the theory illustrated so far. The results here are contained in the paper [5]. We consider the PDE

$$\partial_t\mu_t - \operatorname{div}\left(\frac{\nabla(L \circ \rho_t)}{\rho_t}\mu_t\right) \quad \text{in } X \times (0,T). \tag{0.0.9}$$

Here L represents the nonlinearity, and ρ_t denotes the density of the unknown μ_t with respect to a reference measure on $\gamma \in \mathscr{P}(X)$. In the Euclidean case, we can choose $\gamma = e^{-V}\mathcal{L}^n$, and if L is the identity we obtain the standard Fokker-Planck equation with potential V. In the case $L(z) = z^m$, $m > 1$, we obtain the porous media equation. In the pioneering paper [43], the authors studied the Fokker-Planck equation in the framework of minimizing movements with respect to the Wasserstein distance. The reference functional for the scheme is the *entropy*

$$\mathcal{H}(\rho) = \int_X \rho \log \rho \, d\gamma.$$

Hence this kind of problem motivates the development of the theory illustrated in Chapter 3. In [9] an infinite dimensional framework is considered, and the authors prove existence and uniqueness of a gradient flow solution when γ is a *log-concave measure*, that is, for any couple of open sets A, B in X, there holds

$$\log \gamma((1-t)A + tB) \geq (1-t) \log \gamma(A) + t \log \gamma(B).$$

In this chapter, we are concerned with the generalization to the nonlinear case, corresponding to different choices of the function L. We are in fact able to prove the same result.

We consider the *relative internal energy functional*

$$\mathscr{F}(\mu|\gamma) = \int_X F\left(\frac{d\mu}{d\gamma}\right) d\gamma.$$

Here $F : \mathbb{R} \to \mathbb{R}$ is a strictly convex function, related to L by $L(z) = zF'(z) - F(z)$. Our goal is to show that the element of minimal norm in the Wasserstein subdifferential of $\mathscr{F}$ is exactly the velocity vector field appearing in (0.0.9). This way the theory developed in Chapter 3, and in particular Theorem 3.5.8, is seen to apply. For, we take advantage of the validity of the same result in Euclidean spaces (see [4]), we take projections on finite dimensional subspaces of X and we pass to the limit. We need a definition of derivative in Hilbert spaces, that is, if $\mathbf{e}_j$ is a basis for X, a function $u \in L^1(X, \gamma)$ is said to have partial derivative $\eta_j = \partial_{\mathbf{e}_j} u \in L^1(X, \gamma)$ if

$$\int_X \partial_{\mathbf{e}_j} \zeta(x) u(x) \, d\gamma(x) = -\int_X \eta_j(x) \zeta(x) \, d\gamma(x) - \int_X u(x) \zeta(x) \, d(\partial_{\mathbf{e}_j} \gamma)(x) \quad \forall \zeta \in \mathrm{Cyl}(X),$$

and then

$$\nabla u := \sum_{j=1}^{\infty} (\partial_{\mathbf{e}_j} u) \mathbf{e}_j.$$

In order to give sense to the latter definition, we need the distributional derivatives $\partial_{\mathbf{e}_j}$ of γ to be absolutely continuous measures themselves. Also we should ask that u is bounded, not to have ambiguities with the last integral. In fact, the latter assumption can be avoided considering a generalized notion of differentiability, asking that all the truncates $(-n) \vee u \wedge n$ are weakly differentiable. On the other hand, the former one is strictly needed, and by slightly modifying the argument of Chapter 2, we are also able to show that optimal transport maps exist from such measures even if they are not regular. Moreover, log-concavity of γ and the McCann hypothesis (see [53]) on F, that is, $e^z F(e^{-z})$ is convex and nonincreasing, ensure the geodesical convexity of the energy, so that we are in the framework of Theorem 3.5.8.

Finally, for our argument we need to generalize one of the results of Chapter 3 holding in the convex case. We have mentioned that, by closure of the subdifferential, the element of minimal norm in $\partial\mathscr{F}(\mu)$ can be approximated in L^2 (in the suitable sense) by a sequence of elements in $\partial\mathscr{F}(\mu_{\tau_n})$, for τ_n going to zero. Here μ_{τ_n} minimize the perturbed functionals of the discrete problem (0.0.7). We show that the same holds true if μ_{τ_n} are minimizers of suitable functionals Γ-converging to $\mathscr{F}$. Hence, a section of the chapter will be focused on Γ-convergence.

The main result can be summarized as follows. For all $\mu^0 \in \mathscr{P}_2(X)$ there exists a unique distributional solution $\mu_t = \rho_t\gamma$ to (0.0.9), satisfying $L_F \circ \rho_t \in W^{1,1}(X,\gamma)$ for a.e. $t > 0$ and:

$$\left\| \frac{\nabla(L_F \circ \rho_t)}{\rho_t} \right\|_{L^2(X,\mu_t;X)} \in L^2_{loc}(0,+\infty).$$

Furthermore, reasoning as done in [1] we are able to show that if $\mu^0 \leq C\gamma$, then $\rho_t \leq C$ γ-a.e. for all $t > 0$.

Chapter 5 presents a second application to PDEs. The results here are contained in [8, 50, 6]. We are concerned with a model for the evolution of the so-called Ginzburg-Landau vortices in a superconductor. The model, proposed in [28], is the following.

$$\frac{d}{dt}\mu_t - \operatorname{div}(\nabla h_{\mu(t)}\mu_t) = 0 \quad in\ \mathcal{D}'((0,+\infty)\times\Omega), \tag{0.0.10}$$

where Ω is an open, bounded subset of $\mathbb{R}^2$ and the velocity vector field is coupled with μ_t, for any t, by

$$\begin{cases} -\Delta h_\mu + h_\mu = \mu & \text{in } \Omega \\ h_\mu = 1 & \text{on } \partial\Omega. \end{cases} \tag{0.0.11}$$

We look for a solution μ_t which is a measure in $P(\overline{\Omega}) \cap H^{-1}(\Omega)$. We are working with measures on $\overline{\Omega}$ in order to treat masses in Ω which vary during the evolution and may concentrate on the boundary. For measures μ on $\overline{\Omega}$ we make use of the notation $\mu = \widehat{\mu} + \widetilde{\mu}$, where $\widehat{\mu} = \chi_\Omega\mu$ and $\widetilde{\mu} = \chi_{\partial\Omega}\mu$. The reference functional is

$$\Phi_\lambda(\mu) = \Phi_\lambda = \frac{\lambda}{2}\mu(\Omega) + \frac{1}{2}\int_\Omega |\nabla h_\mu|^2 + |h_\mu - 1|^2, \quad \lambda \geq 0. \tag{0.0.12}$$

It is related to the standard Ginzburg-Landau energy. In fact, it is shown in [64] that it is the limit, in a suitable sense, of the Ginzburg-Landau functional in some asymptotic range of parameters. In this case we do not have the geodesical convexity property, so we will apply the theory of Section 3.4. Since we are in $\mathbb{R}^2$, we do not have any problem for compactness of sublevels.

After a brief physical introduction of the problem, we recall some results contained in the seminal paper [8]. The main result therein reads as follows: given an initial datum μ^0 with $\|\widehat{\mu}\|_{L^p} = C$, there exists a solution for (0.0.10) such that, for any t, $\|\widehat{\mu}_t\|_{L^p} \leq C$. This result is based on the following facts. An *entropy* function φ, behaving like the p-power, can be exhibited such that

$$\int_\Omega \varphi(\widehat{\mu}_t) \le \int_\Omega \varphi(\widehat{\mu}^0).$$

This is the essential element for the regularity part of the statement. Moreover, an Euler-Lagrange equation for the minimizing movements problem, starting from a generic $\mu \in \mathscr{P}_2(\overline{\Omega})$, is derived. Indeed, if μ_τ is the corresponding discrete minimizer, there holds

$$\frac{\mathbf{I} - \mathbf{t}_{\mu_\tau}^{\mu}}{\tau} = -\nabla h_{\mu_\tau}.$$

Since the left hand side is the discrete velocity of the scheme, passing to the limit as $\tau \to 0$, after constructing the piecewise constant interpolation $\mu_\tau^{\lceil t/\tau \rceil}$, one obtains, for any t, a characterization of the limiting velocity as $\mathbf{v}_t = -\nabla h_{\mu_t}$. Then, by Theorem 3.4.4, the couple $(\mu_t, \mathbf{v}_t)$ satisfies the continuity equation.

The main result of Chapter 5 is the global uniqueness of solutions. The main difficulty is the potential presence of mass on the boundary. The question of giving natural boundary conditions for the model is also strictly related. Hence, the main part of the analysis consist in taking a new variation for the discrete minimization problem, taking into account transportation of mass from $\partial\Omega$. Given the initial point μ and being μ_τ the minimizer, we compare it with

$$\mu_\tau^\varepsilon := \widehat{\nu} + T_{\varepsilon\#}(\alpha^2\sigma) + (1-\alpha^2)\widetilde{\nu},$$

where $\alpha \sim 0$ and $\varepsilon \sim 1$ are suitable parameters and σ is an arbitrary diffuse measure, T is the optimal transport map between σ and $\widetilde{\nu}$, and

$$T_\varepsilon = (1-\varepsilon)I + \varepsilon T, \quad \varepsilon \in [0,1].$$

We see that μ_τ^ε is built taking a portion α of the mass of μ_τ on the boundary and moving it inside, in correspondence of σ. With this kind of variation we are able to show that there exists a discrete minimizer μ_τ such that

$$\langle \nabla h_{\mu_\tau}(x), y - x\rangle \ge 0 \quad \forall (x,y) \in \operatorname{supp}(\widetilde{\mu}_\tau) \times \overline{\Omega}. \tag{0.0.13}$$

This condition says that, at least if Ω is convex, the velocity $-\nabla h_{\mu_\tau}$ near the boundary tends to be directed towards the exterior of the domain (and normally to the boundary). This fact suggests that no mass is going to enter in Ω from $\partial\Omega$. Also, it suggests that the actual velocity for the equation should be $-\chi_\Omega \nabla h_{\mu_\tau}$, hence zero on the boundary. Moreover, passing to the limit in (0.0.13) we obtain the condition which ensures global uniqueness of L^∞ solutions. The argument of the proof extends the one used for Theorem 3.3.4.

The actual model of [28] involves signed measures. In fact, Ginzburg-Landau vortices possess a topological charge. In the second part of Chapter 5 we analyze the evolution of (0.0.12) in the signed case. Of course we can no more apply the theory of the first chapters, suitable only for positive measures with the same mass. Given two signed measures $\mu, \nu \in \mathcal{M}(\overline{\Omega})$ with same integral, we have to define suitable cost functionals. A first choice is

$$\mathbb{W}_2(\mu,\nu) := W_2(\mu^+ + \nu^-, \nu^+ + \mu^-). \tag{0.0.14}$$

This is not a distance, but it can be bounded from below with a distance (if we have a uniform bound on total masses), and we will show some properties of the transportation given by $\mathbb{W}_2$. In order not to have too many degrees of freedom, we also introduce, for fixed $\mu \in \mathcal{M}(\overline{\Omega})$, the following

$$\mathcal{W}_2^2(\nu,\mu) = \inf_{\sigma^1-\sigma^2=\nu} \left(W_2^2(\sigma^1,\mu^+) + W_2^2(\sigma^2,\mu^-) \right).$$

Even this new object can be bounded from below with a distance, so that we are still able to construct a minimizing movement μ_t. Moreover, the map $\nu \mapsto \mathcal{W}_2(\nu,\mu)$ is lower semicontinuous in the weak topology of $\mathcal{M}(\overline{\Omega})$ (which is again defined by the duality with continuous and bounded functions). Then, we can reproduce the Euler-Lagrange equation, which takes the form

$$-\nabla h_{\mu_\tau^\delta} \widehat{\mu}_\tau = \frac{1}{\tau}(I - r_1)\widehat{\mu}_\tau^+ + \frac{1}{\tau}(I - r_2)\widehat{\mu}_\tau^-,$$

where r_1 and r_2 are optimal transport maps between portions of positive and negative parts of the measures. Also, we are able to obtain L^p regularity for discrete minimizers by means of an entropy argument, as in the positive case.

We can no more apply Theorem 3.4.4. Still we can pass to the limit as τ goes to zero, and we obtain an equation like (0.0.10), that is

$$\frac{d}{dt}\mu_t - \mathrm{div}(\chi_\Omega \nabla h_{\mu(t)} \varrho_t) = 0 \quad in\ \mathcal{D}'((0,+\infty)\times\mathbb{R}^2).$$

In the second term, a suitable measure $\varrho_t \geq |\mu_t|$ appears in place of $|\mu_t|$. It is not yet clear whether $\varrho_t = |\mu_t|$ or not, this is related to the problem of proving convergence in stronger topologies for this scheme. We plan to investigate this in the future, in collaboration with L. Ambrosio and S. Serfaty.

Chapter 6 deals with a shape optimization problem in $\mathbb{R}^2$, that is, a problem of the form

$$\min\left\{\mathcal{F}(\Sigma)\ :\ \Sigma \text{ closed connected subset of } \mathbb{R}^2\right\}.$$

We refer in particular to [19]. The techniques exhibited and the results obtained have in fact a little relation with the analysis of the previous chapters. Nevertheless, we decide to add this part since it has been developed during the PhD studies and since the problem is strictly related to optimal transport.

We consider the following functional

$$\mathcal{F}(\Sigma) := \int_{\mathbb{R}^2} \mathrm{dist}\,(x,\Sigma)\, d\mu(x).$$

It is obtained as a particular instance of the transport network cost (0.0.3) discussed in Section 2.4. In fact, we may think that we are concerned with a density μ of population and a transport line Σ of fixed length to be constructed such that the average cost (proportional to the distance) for citizens for reaching it is minimized. A constraint on the maximal length of Σ has to be introduced, so we will add a penalization term, proportional to $\mathcal{H}^1(\Sigma)$. In terms of Wasserstein distance, this minimization problem corresponds to

$$\inf_{\Sigma}\{W_1(\mu,\nu)+\lambda\mathcal{H}^1(\Sigma) : \operatorname{supp}\nu \in \Sigma\}.$$

Existence of a solution is achieved by standard compactness arguments. One of the most important questions here is about the regularity of minimizers. But we do not address this argument. Rather, we are interested in stationary points. First of all, since we do not have a standard differentiable structure on closed connected subsets of $\mathbb{R}^2$, we have to find a suitable definition. We show that, in the Hausdorff topology, it is not possible to satisfy the usual first order stationarity equality. On the other hand, taking variations induced by a smooth vector field X, hence without changing the topology of sets, we obtain the Euler equation

$$\int_{\mathbb{R}^2}\left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|}\right\rangle d\mu - \lambda\langle H_\Sigma, X\rangle = 0.$$

Here H_Σ denotes the distributional curvature and π^Σ is the projection map on Σ.

The rest of the analysis is concerned with this equation, with μ being the Lebesgue measure on some bounded subset of $\mathbb{R}^2$. We show that a stationary point (in the above sense) can contain closed loops (whereas it is known that minimizers can not). We show that a set with a corner between two segments can not be stationary. On the other hand, we show that irregular sets can be stationary, since we are able to give an explicit example of a stationary set Σ containing a corner point. We also show that the same set can not be stationary in a convex Ω if the angle is too large.

In the final part of the chapter, we analyze a different functional, that is,

$$\mathcal{F}(\Sigma) := \int_\Omega u_\Sigma(x)f(x)\,dx + \lambda\mathcal{H}^1(\Sigma).$$

Here $\Omega \subset \mathbb{R}^2$ is a given bounded open set, f is a given function, and u_Σ is the unique solution of an elliptic PDE with $\Sigma\cap\overline{\Omega}$ as Dirichlet region. We see how to derive a similar Euler equation for this case.

Notation

$\subset$	Subset, possibly not strict
$\Subset$	Compact subset
a.e.	Almost every, almost everywhere
l.s.c.	Lower semicontinuous
$B_r(x)$	Open ball of radius r and center x
$\lvert \cdot \rvert$	Hilbertian norm, total variation of a measure
supp	Support
X'	Dual space: linear continuous functionals on X
$\mathcal{B}(X)$	Family of Borel subsets of X
$\mathcal{M}_+(X)$	Set of nonnegative measures over X
$\mathcal{M}(X)$	Set of real measures over X
$\mathcal{M}_{\kappa,M}(X)$	Set of real measures μ over X with $\mu(X) = \kappa$ and $\lvert\mu\rvert(X) \leq M$
$\mathscr{P}(X)$	Set of probability measures over X
$\mathscr{P}_p(X)$	Measures in $\mathscr{P}(X)$ with finite p-th moment
$\mathscr{P}^r_p(X)$	Regular measures in $\mathscr{P}(X)$ with finite p-th moment (Definition 1.4.5)
$\Gamma(\cdot, \ldots, \cdot)$	Probability measures in a product space with given marginals
$\Gamma_0(\cdot, \cdot)$	Optimal transport plans between given probability measures, see (2.1.4)
$C^0(X)$	Continuous functions over X
$C^0_b(X)$	Continuous and bounded functions over X
$C^0_c(X)$	Compactly supported continuous functions over X
$C^\infty_c(X)$	Compactly supported smooth functions over X
$C^{k,\alpha}(X)$	Hölder spaces over X
$\mathrm{Cyl}(X)$	Smooth cylindrical functions over X (Definition 1.4.3)
$L^p(X, \mu)$	Real valued p-summable functions over X with respect to μ
$L^p(X, \mu; Y)$	Y-valued p-summable functions over X with respect to μ
$W^{k,p}(X)$	Sobolev spaces over X
$\mathrm{Lip}(f)$	Lipschitz constant of f

$\mathbf{t}_{\#}\mu$	Push forward of the measure μ through the map $\mathbf{t}$, see (1.2.1)
$\mathbf{t}_{\mu}^{\nu}$	Optimal transport map between measures μ and ν
$W_p(\cdot,\cdot)$	p-Wasserstein distance, see (2.2.1)
$\mathcal{W}_2(\cdot,\cdot)$	Pseudo 2-Wasserstein distance, see (5.6.15)
$\mathcal{L}^n$	n-dimensional Lebesgue measure
$\mathcal{H}^k$	k-dimensional Hausdorff measure
$\mathrm{Tan}^k(\Sigma, x)$	Approximate tangent space of the k-manifold Σ at x (Definition 1.6.4)
$\mathrm{div}^{\Sigma} f$	Tangential divergence of f with respect to the manifold Σ, see (1.6.1)
$\mathbf{H}_{\Sigma}$	Distributional vector curvature of the manifold Σ (Definition 1.6.7)
$\mu \bigotimes \nu$	Tensor product of measures μ and ν
$\overline{\mathrm{Conv}}(\Omega)$	Closed convex hull of Ω
$AC([0,T];E)$	Absolutely continuous curves in a metric space E (Definition 2.3.1)
$\lvert\mu'\rvert_t$	Metric derivative of the curve $t \mapsto \mu_t$ (Definition 2.3.2)
$D(\phi)$	Effective domain of functional ϕ, see (3.1.1)
$\lvert\partial\phi\rvert(\mu)$	Metric slope of functional ϕ at μ (Definition 3.1.5)
$\partial\phi(\mu)$	Wasserstein subdifferential of functional ϕ at μ (Definition 3.1.7)

Chapter 1

Measure theoretic setting

This first chapter is devoted to the introduction of the basic mathematical framework.

Let X be a separable metric space with distance d. Let $\mathcal{B}(X)$ denote the σ-algebra of all Borel subsets of X and $\mathscr{P}(X)$ be the set of probability measures over $(X, \mathcal{B}(X))$. The space X is called a Radon space (see for instance [14, 67]) if the following inner regularity property holds: given a measure $\mu \in \mathscr{P}(X)$,

$$\forall B \in \mathcal{B}(X), \varepsilon > 0, \quad \exists K \Subset B \text{ s.t. } \mu(B \setminus K) < \varepsilon. \tag{1.0.1}$$

In this chapter, $X, Y, X_n, n \in \mathbb{N}$ will be understood to be separable metric Radon spaces. Sometimes we will particularize the setting to Hilbert or Euclidean spaces. When X will stand for a separable Hilbert space, we will denote by $|\cdot|$ the norm and by $\{\mathbf{e}_j\}_{j\in\mathbb{N}}$ the reference orthonormal basis. Polish spaces satisfy the Radon property (1.0.1), but we will have to deal also with non complete spaces.

1.1 The weak topology of probability spaces

We need to introduce a topological structure on $\mathscr{P}(X)$.

Definition 1.1.1 (Weak convergence) *We say that a sequence $(\mu_n) \subset \mathscr{P}(X)$ is weakly (or narrowly) converging to $\mu \in \mathscr{P}(X)$ if, for any continuous and bounded function f on X, there holds*

$$\int_X f \, d\mu_n \to \int_X f \, d\mu. \tag{1.1.1}$$

In this case we write $\mu_n \rightharpoonup \mu$.

Remark 1.1.2 If X is also locally compact, in view of the Riesz representation theorem (see for instance [63, chapter 2]), the space of Radon measures over X can be identified with the dual space of $C^0_b(X)$. In such context, the convergence in duality with $C^0_b(X)$ is the weak* convergence in the dual space $(C^0_b(X))'$.

In the case of a semicontinuous function, a suitable approximation argument with Lipschitz functions allows to obtain an inequality (for the details we refer to [4, §5.1], where the fact that it is enough to check (1.1.1) on Lipschitz functions is also remarked). Let $f : X \to (-\infty, +\infty]$ be l.s.c. and bounded from below. Let $\mu_n \rightharpoonup \mu$. Then

$$\int_X f\,d\mu \le \liminf_{n\to\infty} \int_X f\,d\mu_n. \tag{1.1.2}$$

Analogously, if f is upper semicontinuous and bounded from above one obtains

$$\int_X f\,d\mu \ge \limsup_{n\to\infty} \int_X f\,d\mu_n. \tag{1.1.3}$$

Applying the last two inequalities respectively to the characteristic function of an open and a closed subset of X we find

$$\mu(A) \le \liminf_{n\to\infty} \mu_n(A) \quad \forall A \text{ open}, A \subset X, \tag{1.1.4}$$

$$\mu(A) \ge \limsup_{n\to\infty} \mu_n(A) \quad \forall A \text{ closed}, A \subset X. \tag{1.1.5}$$

Let now d be a distance on X and fix $p > 0$. In the sequel, we will say that a function f on X has p-growth at infinity if, for some $A, B \ge 0$ and $\bar{x} \in X$ there holds

$$|f(x)| \le A + Bd^p(x, \bar{x}) \quad \forall x \in X. \tag{1.1.6}$$

Let us introduce the subset $\mathscr{P}_p(X)$ of probability measures with finite p-moment.

Definition 1.1.3 (Moments) *We say that $\mu \in \mathscr{P}_p(X)$ if its p-th moment is finite, that is, for some $\bar{x} \in X$,*

$$\int_X d^p(x, \bar{x})\,d\mu < +\infty. \tag{1.1.7}$$

Moreover, we say that a set $\Xi \subset \mathscr{P}(X)$ has uniformly integrable p-moments if, for some $\bar{x} \in X$,

$$\lim_{r\to\infty} \sup_{\mu\in\Xi} \int_{X\setminus B_r(\bar{x})} d^p(x, \bar{x})\,d\mu = 0. \tag{1.1.8}$$

Equivalently, one can ask that

$$\sup_{\mu\in\Xi} \int_X f\,d\mu < +\infty, \tag{1.1.9}$$

for some positive function f whose growth at infinity is faster than p.

Concerning narrowly converging sequences of measures, we have the following characterization of uniform integrability of p-moments.

Proposition 1.1.4 *Let $\{\mu_n\}_{n\in\mathbb{N}} \subset \mathscr{P}_p(X)$ be such that $\mu_n \rightharpoonup \mu \in \mathscr{P}(X)$. The family $\{\mu_n\}_{n\in\mathbb{N}}$ has uniformly integrable p-moments if and only if* (1.1.1) *holds for any continuous function f on X with at most p-growth at infinity.*

Proof. Let $p > 0$, $\bar{x} \in X$. Suppose that the sequence (μ_n) has uniformly integrable p-moments and let $f : X \to \mathbb{R}$ be such that (1.1.6) holds. Let moreover $f_k = f \wedge k$, $k > 0$. Since f_k is continuous and bounded from above, it satisfies (1.1.3), and since $f \geq f_k$ we have

$$\begin{aligned}\limsup_{n\to\infty} \int_X f(x)\,d\mu_n(x) &= \limsup_{n\to\infty} \int_X (f(x) - f_k(x))\,d\mu_n(x) + \limsup_{n\to\infty} \int_X f_k(x)\,d\mu_n(x)\\ &\leq \sup_{n\in\mathbb{N}} \int_{\{x\in X: f(x)>k\}} (f(x) - k)\,d\mu_n(x) + \int_X f_k(x)\,d\mu(x)\\ &\leq \sup_{n\in\mathbb{N}} \int_{\{x\in X:\, d^p(\bar{x},x)>\frac{k-A}{B}\}} (A + Bd^p(\bar{x},x))\,d\mu_n(x) + \int_X f(x)\,d\mu(x).\end{aligned}$$

Taking the limit in k, since (μ_n) has uniformly integrable p-moments, we see that f satisfies (1.1.3). In order to show (1.1.2), one reasons in the same way with $f_k = (-k) \vee f$, $k > 0$, and taking advantage of (1.1.6).

On the other hand, let (1.1.1) hold for any continuous function with p-growth. Hence (1.1.1) holds in particular for the function $d^p(\bar{x},x) - d^p_k(\bar{x},x)$, where $d^p_k(\bar{x},x) := d^p(\bar{x},x) \wedge k$, $k > 0$. Moreover, (1.1.5) can be applied to the d-closed set $D_k := \{x \in X : d^p(\bar{x},x) \geq k\}$ (complement of the open ball $B_{k^{1/p}}(\bar{x})$), for any $k > 0$. Hence

$$\begin{aligned}\limsup_{n\to\infty} \int_{D_k} d^p(\bar{x},x)\,d\mu_n(x) &= \limsup_{n\to\infty} \left(\int_X (d^p(\bar{x},x) - d^p_k(\bar{x},x))\,d\mu_n(x) + k\mu_n(D_k) \right)\\ &\leq \int_X (d^p(\bar{x},x) - d^p_k(\bar{x},x))\,d\mu(x) + k\mu(D_k)\\ &= \int_{D_k} d^p(\bar{x},x)\,d\mu(x),\end{aligned}$$

yielding

$$\lim_{k\to\infty} \limsup_{n\to\infty} \int_{D_k} d^p(\bar{x},x)\,d\mu_n(x) = 0.$$

Here the limsup can be substituted with the supremum over $\mathbb{N}$, since of course any finite subset of $\{\mu_n\}_{n\in\mathbb{N}}$ has uniformly integrable p-moments, so that (1.1.8) is proved. □

It is natural to give the next definition.

Definition 1.1.5 (Convergence with moments) *We say that a sequence $(\mu_n) \subset \mathscr{P}_p(X)$ converges to μ in $\mathscr{P}_p(X)$ if $\mu_n \rightharpoonup \mu$ and* (1.1.1) *holds for any continuous function f on X with at most p-growth. After* Proposition 1.1.4, *it is equivalent to require that $\mu_n \rightharpoonup \mu$ and, for some $\bar{x} \in X$,*

$$\int_X d^p(x,\bar{x})\,d\mu_n \to \int_X d^p(x,\bar{x})\,d\mu. \qquad (1.1.10)$$

We write $\mu_n \to \mu$ in $\mathscr{P}_p(X)$ to denote this convergence (or simply $\mu_n \to \mu$ in the distinguished case $p = 2$).

We also recall the following result about narrow convergence.

Proposition 1.1.6 *Let $\mu_n \rightharpoonup \mu$. Then for any $x \in \operatorname{supp}\mu$ there exist elements $x_n \in \operatorname{supp}\mu_n$ such that $x_n \to x$ in X.*

Proof. Let $x \in \operatorname{supp}\mu$ and consider the following sequence:

$$a_k = \min\left\{n \in \mathbb{N} :\ a_{k-1} < n, \quad \operatorname{supp}\mu_m \cap B_{1/k}(x) \neq \emptyset \quad \forall m \geq n\right\},$$

with $a_0 = 0$. It is well defined since (1.1.4) gives $0 < \mu(B_{1/k}(x)) \leq \liminf_{n\to\infty} \mu_n(B_{1/k}(x))$. Let $(n_k) \subset \mathbb{N}$ be a sequence such that, $a_k \leq n_k < a_{k+1}$. For any $k \in \mathbb{N}$, let x_{n_k} be a point in $\operatorname{supp}\mu_{n_k} \cap B_{1/k}(x)$. The sequence (x_{n_k}) realizes the desired approximation. □

L^p spaces

Let $\mu \in \mathscr{P}(X)$ and let $f : X \to \mathbb{R}$ be μ-measurable. f belongs to $\mathcal{L}^p(X,\mu)$ if

$$\int_X |f|^p \, d\mu < +\infty.$$

As usual, with $L^p(X,\mu)$ we denote the corresponding quotient space with respect to the equivalence relation which identifies functions agreeing μ-a.e. on X. Now let $\boldsymbol{\rho} : X \to Y$ be a μ-measurable vector map and let d_Y denote the distance on the space Y. We say that $\boldsymbol{\rho} \in L^p(X,\mu;Y)$ if, for some $\bar{y} \in Y$

$$\int_X d_Y^p(\bar{y}, \boldsymbol{\rho}(x)) \, d\mu(x) < +\infty. \tag{1.1.11}$$

Given two functions $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2 \in L^p(X,\mu;Y)$, their L^p distance is defined by

$$\mathbf{d}_{L^p(X,\mu;Y)}(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = \left(\int_X d_Y^p(\boldsymbol{\rho}_1(x), \boldsymbol{\rho}_2(x)) \, d\mu(x)\right)^{1/p}. \tag{1.1.12}$$

Tightness and compactness

We introduce the following

Definition 1.1.7 (Tightness) *A set $\Xi \subset \mathscr{P}(X)$ is said to be tight if for any $\varepsilon > 0$ there exist compact subsets K_ε of X such that*

$$\mu(X \backslash K_\varepsilon) < \varepsilon, \quad \forall \mu \in \Xi. \tag{1.1.13}$$

An equivalent condition is given by

Proposition 1.1.8 *A set $\Xi \subset \mathscr{P}(X)$ is tight if and only if there exists a function $\varphi : X \to [0, +\infty]$ with compact sublevels such that*

$$\sup_{\mu \in \Xi} \int_X \varphi \, d\mu < +\infty. \tag{1.1.14}$$

Proof. Let $\Xi \subset \mathscr{P}(X)$ satisfy (1.1.14). Tschebyshev inequality gives

$$\sup_{\mu\in\Xi} \mu(\{x \in X : \varphi(x) > 1/\varepsilon\}) \leq \varepsilon \sup_{\mu\in\Xi} \int_X \varphi \, d\mu.$$

By (1.1.14) we deduce that, as $\varepsilon \to 0$, $\sup_{\mu\in\Xi} \mu(\{x \in X : \varphi(x) > 1/\varepsilon\}) \to 0$, and since the sublevels of φ are compact, (1.1.13) follows choosing $K_\varepsilon = \{x \in X : \varphi(x) \leq 1/\varepsilon\}$.

On the other hand, let $(a_n) \in \mathbb{R}$ be a sequence such that $\sum_{n=0}^{+\infty} a_n < +\infty$ and let K_{a_n} be a sequence of sets satisfying (1.1.13), and such that $K_{\varepsilon_n} \subset K_{\varepsilon_{n+1}}$. Then

$$\varphi(x) := \inf\{n \in \mathbb{N} : x \in K_{\varepsilon_n}\} = \sum_{n=0}^{+\infty} \chi_{X\setminus K_{\varepsilon_n}}(x)$$

satisfies (1.1.14). □

Remark 1.1.9 If X is an Euclidean space, local compactness ensures that tightness, in view of Proposition 1.1.8, is implied by (1.1.9), for any $p \geq 0$. This way, tightness can be seen to follow from uniform integrability of moments.

The next theorem is the standard characterization of compactness in probability spaces (see for instance [14, 33]).

Theorem 1.1.10 (Prokhorov) *A set $\Xi \subset \mathscr{P}(X)$ is relatively compact (with respect to the weak topology in $\mathscr{P}(X)$) if it is tight. The converse holds true in Polish (complete separable metric) spaces.*

In a Radon space, any measure $\mu \in \mathscr{P}(X)$, as a singleton, is tight, thanks to the inner approximation property with compact sets (1.0.1). Hence, a sequence $(\mu_n) \subset \mathscr{P}(X)$ is a tight set if and only if

$$\sup_{K\Subset X} \liminf_{n\to\infty} \mu_n(K) = 1.$$

Moreover, it is well known that in a Radon space every weakly converging sequence $(\mu_n) \in \mathscr{P}(X)$ is tight (see [46]).

1.2 Transport of measures and multiple plans

Push-forward

Let $\mathbf{T} : X_1 \to X_2$ be a μ-measurable map. The *push-forward* (or *transport*) of μ on X_2 through $\mathbf{T}$ is the measure $\nu \in \mathscr{P}(X_2)$ defined by

$$\nu(B) = \mu(\mathbf{T}^{-1}(B)), \quad \forall B \in \mathcal{B}(X_2). \tag{1.2.1}$$

We write $\nu = \mathbf{T}_\#\mu$ and $\mathbf{T}$ is called a *transport map* between μ and ν.

If $f : X_2 \to \mathbb{R}$ is a $\mathbf{T}_\#\mu$-measurable function, then $f \circ \mathbf{T}$ is μ-measurable and the definition can be written in terms of the change of variables formula:

$$\int_{X_1} f \circ \mathbf{T}\, d\mu = \int_{X_2} f\, d\mathbf{T}_\#\mu. \tag{1.2.2}$$

A consequence of (1.2.2) is the following composition rule: if X_3 is a third space and $\mathbf{S} : X_2 \to X_3$ is a $\mathbf{T}_\#\mu$-measurable map, then

$$\mathbf{S}_\#(\mathbf{T}_\#\mu) = (\mathbf{S} \circ \mathbf{T})_\#\mu. \tag{1.2.3}$$

In fact from (1.2.2) one immediately has, for any $\mathbf{S}_\#(\mathbf{T}_\#\mu)$-measurable $g : X_3 \to \mathbb{R}$,

$$\int_{X_3} g\, d(\mathbf{S}_\#(\mathbf{T}_\#\mu)) = \int_{X_2} g \circ \mathbf{S}\, d\mathbf{T}_\#\mu = \int_{X_1} g \circ \mathbf{S} \circ \mathbf{T}\, d\mu = \int_{X_3} g\, d(\mathbf{S} \circ \mathbf{T})_\#\mu.$$

The measure $\mathbf{T}_\#\mu$ is concentrated on the image of $\mathbf{T}$, and we have

$$\operatorname{supp} \mathbf{T}_\#\mu \subset \overline{\mathbf{T}(\operatorname{supp}\mu)}. \tag{1.2.4}$$

The inclusion is in general strict. The two sets coincide in the case of a continuous map $\mathbf{T}$.

Product spaces, projections and multiple plans

Given the separable metric Radon spaces $X_1, \ldots, X_n$, consider the product space $\mathbf{X} = \prod_{i=1}^n X_i$. The projection map $\pi^i : \mathbf{X} \to X_i$ is defined by

$$\pi^i : (x_1, \ldots, x_n) \mapsto x_i.$$

Analogously we define $\pi^{i_1,\ldots,i_h}$, with $h \le n$ and $i_1, \ldots, i_h \in \{1, \ldots, n\}$, as

$$\pi^{i_1,\ldots,i_h} : (x_1, \ldots, x_n) \mapsto (x_{i_1}, \ldots, x_{i_h}).$$

Clearly projection maps are continuous.

Given a probability measure $\boldsymbol{\gamma} \in \mathscr{P}(\mathbf{X})$, the i-th marginal of $\boldsymbol{\gamma}$ is the measure in $\mathscr{P}(X_i)$ given by $\pi^i_\#\boldsymbol{\gamma}$. Multiple marginals are defined as $\pi^{i_1,\ldots,i_h}_\#\boldsymbol{\gamma} \in \mathscr{P}(X_{i_1} \times \ldots \times X_{i_h})$.

We now introduce the set of *multiple plans* of given marginals $\mu^i \in \mathscr{P}(X_i)$, $i = 1, \ldots, n$:

$$\Gamma(\mu^1, \ldots, \mu^n) = \left\{\boldsymbol{\gamma} \in \mathscr{P}(\mathbf{X}) : \pi^i_\#\boldsymbol{\gamma} = \mu^i, i = 1, \ldots, n\right\}. \tag{1.2.5}$$

The set $\Gamma(\mu^1, \ldots, \mu^n)$ is always nonempty, since the product measure $\mu^1 \times \ldots \times \mu^n$ is clearly a multiple plan. Moreover, we have the following

Lemma 1.2.1 *The set $\Gamma(\mu^1, \ldots, \mu^n)$ is compact in $\mathscr{P}(\mathbf{X})$.*

Proof. Let $\varepsilon > 0$. For each $i \in \{1,\dots,n\}$, the set $\{\pi^i_\#\boldsymbol{\gamma} : \boldsymbol{\gamma} \in \Gamma(\mu^1,\dots,\mu^n)\}$ is the singleton $\{\mu^i\}$, so it enjoys the inner regularity property (1.0.1) since we are working in Radon spaces. Hence, we can find compact sets $K_i \Subset X_i$, $i = 1,\dots,n$, such that $\pi^i_\#\boldsymbol{\gamma}(X_i \backslash K_i) < \varepsilon/n$, for any $\boldsymbol{\gamma} \in \Gamma(\mu^1,\dots,\mu^n)$. By definition of push forward, it follows that $\boldsymbol{\gamma}(\mathbf{X}\backslash(\pi^i)^{-1}(K_i)) < \varepsilon/n$ for any $i \in \{1,\dots,n\}$ and $\boldsymbol{\gamma} \in \Gamma(\mu^1,\dots,\mu^n)$. Notice that the set

$$\bigcap_{i=1}^n (\pi^i)^{-1}(K_i)$$

is compact in $\mathbf{X}$. We have

$$\boldsymbol{\gamma}\left(\mathbf{X}\backslash\bigcap_{i=1}^n (\pi^i)^{-1}(K_i)\right) \le \sum_{i=1}^n \boldsymbol{\gamma}(\mathbf{X}\backslash(\pi^i)^{-1}(K_i)) \quad \forall \boldsymbol{\gamma} \in (\Gamma(\mu^1,\dots,\mu^n)),$$

which gives tightness of $\Gamma(\mu^1,\dots,\mu^n)$. Compactness follows from Prokhorov theorem. □

A more general proposition about compactness and product spaces is the following (the proof is analogous).

Proposition 1.2.2 *Let $\rho^i : \mathbf{X} \to X^i$ be continuous functions such that the vector map $\boldsymbol{\rho} := (\rho^1,\dots,\rho^n)$ is proper. If $\Xi \in \mathscr{P}(\mathbf{X})$ is such that $\rho^i_\#(\Xi)$ is tight in $\mathscr{P}(X_i)$ for any $i = 1,\dots,n$, then Ξ is tight in $\mathscr{P}(\mathbf{X})$.*

The most important instance will be $n = 2$: if $\mu^1 \in \mathscr{P}(X_1)$ and $\mu^2 \in \mathscr{P}(X_2)$ the set of 2-*plans* of marginals μ^1 and μ^2 is

$$\Gamma(\mu^1,\mu^2) = \left\{\gamma \in \mathscr{P}(X_1 \times X_2) : \pi^1_\#\gamma = \mu^1, \pi^2_\#\gamma = \mu^2\right\}. \tag{1.2.6}$$

A measure $\gamma \in \Gamma(X_1 \times X_2)$ is called a *transport plan* between μ^1 and μ^2. The terminology is motivated by the fact that a transport plan can be seen as a generalization of a transport map. Indeed, suppose that $\mathbf{t} : X_1 \to X_2$ is such that $\mathbf{t}_\#\mu^1 = \mu^2$. Then it is clear that

$$(\mathbf{I},\mathbf{t})_\#\mu^1 \in \Gamma(\mu^1,\mu^2).$$

In this case we say that γ is induced by $\mathbf{t}$. On the other hand, if γ is not concentrated (in $X_1 \times X_2$) on a graph of some μ^1-measurable function from X_1 to X_2, then γ is not induced by a transport map (see [2]).

When dealing with a separable Hilbert space X with norm $|\cdot|$, scalar product $\langle\cdot,\cdot\rangle$ and orthonormal basis $\{\mathbf{e}_j\}_{j\in\mathbb{N}}$, we will often use the notation X^d to denote a d-dimensional subspace, whereas $\Pi^d : X \to X^d$ will stand for the projection map on such subspace:

$$\Pi^d : x \mapsto \sum_{j=1}^d \langle x, \mathbf{e}_j\rangle\, \mathbf{e}_j.$$

Moreover, if ν is a probability measure over X, we let

$$\nu^d := \Pi^d_\#\nu. \tag{1.2.7}$$

We can consider ν^d as a probability measure on the whole space X and we can state the following

Proposition 1.2.3 *There holds $\Pi^d_\# \nu \rightharpoonup \nu$. If $\nu \in \mathscr{P}_p(X)$, then we also have $\Pi^d_\# \nu \to \nu$ in $\mathscr{P}_p(X)$.*

Proof. Let $f : X \to \mathbb{R}$ be a continuous and bounded function. Then

$$\int_X f \, d\Pi^d_\# \nu = \int_X f \circ \Pi^d \, d\nu.$$

Notice that $f \circ \Pi^d \to f$ pointwise, and since f is bounded we get the weak convergence $\nu^d \rightharpoonup \nu$. In the same way we have $|x|^p \circ \Pi^d \to |x|^p$ pointwise and monotonically, so $\nu^d \to \nu$ in $\mathscr{P}_p(X)$ if ν has finite p-moment. □

Notation 1.2.4 Latin or Greek letters will be used to denote maps. Bold characters will be used for maps taking values in infinite dimensional spaces (as $\mathbf{t}, \mathbf{v}, \mathbf{I}, \boldsymbol{\rho}, \boldsymbol{\xi}, \boldsymbol{\omega}$), but sometimes we will omit the bold notation for Hilbert spaces. Measures and 2-plans will be indicated with the Greek letters $\mu, \nu, \gamma, \beta, \sigma, \vartheta$, whereas the corresponding bold letters will denote 3-plans or n-plans.

1.3 Disintegration

Consider a mapping which associates to each $x_1 \in X_1$ a probability measure $\mu_{x_1} \in \mathscr{P}(X_2)$. We say that such a map is Borel if $x_1 \mapsto \mu_{x_1}(B)$ is a Borel map for any $B \in \mathcal{B}(X_2)$.

Proposition 1.3.1 *Let $f : X_1 \times X_2 \to \mathbb{R}$ be a Borel function. Then the map $F : X_1 \to \mathbb{R}$ defined by*

$$x_1 \mapsto \int_X f(x_1, x_2) \, d\mu_{x_1}(x_2)$$

is itself a Borel map.

Proof. See [3, Proposition 2.26]. □

Let $x_1 \in X_1 \mapsto \mu_{x_1} \in \mathscr{P}(X_2)$ be a Borel map as in the previous definition. Let $\nu \in \mathscr{P}(X_1)$. Then we can uniquely define a probability measure on the product space $X_1 \times X_2$ by

$$\gamma(\mathbf{A}) = \int_{X_1} \left(\int_{X_2} \chi_{\mathbf{A}}(x_1, x_2) \, d\mu_{x_1}(x_2) \right) d\nu(x_1),$$

for Borel sets $\mathbf{A} \subset X_1 \times X_2$. Clearly $\pi^1 \gamma = \nu$. Such measure will be denoted by

$$\int_{X_1} \mu_{x_1} \, d\nu(x_1), \tag{1.3.1}$$

and we will say that μ_{x_1} is the family of measures which disintegrates γ with respect to ν. Another notation is $\nu \bigotimes \mu_{x_1}$, since it is a generalized product between measures. The integral of a Borel function $f : X_1 \times X_2 \to \mathbb{R}$ with respect to $\int_X \mu_{x_1} \, d\nu(x_1)$ is

$$\int_{X_1} \left(\int_{X_2} f(x_1, x_2) \, d\mu_{x_1}(x_2) \right) d\nu(x_1).$$

Theorem 1.3.2 (Disintegration) *Let $\mathbf{Y}$ be a Radon metric space. Let $\boldsymbol{\gamma} \in \mathscr{P}(\mathbf{Y})$, let $\Psi : \mathbf{Y} \to X_1$ be Borel and let $\nu = \Psi_\# \boldsymbol{\gamma}$. Then, for ν-a.e. $x_1 \in X_1$, there exists a measure $\mu_{x_1} \in \mathscr{P}(\mathbf{Y})$ such that the map $x_1 \mapsto \mu_{x_1}$ is Borel, μ_{x_1} is concentrated on $\Psi^{-1}(x_1)$ and for any Borel set $\mathbf{A} \subset \mathbf{Y}$ there holds*

$$\boldsymbol{\gamma}(\mathbf{A}) = \int_{X_1} \mu_{x_1}(\mathbf{A})\, d\nu(x_1). \tag{1.3.2}$$

Finally, μ_{x_1} is uniquely determined up to ν-negligible sets.

Proof. See [3, Theorem 2.28] and [33]. □

As a consequence, for every Borel map $f : \mathbf{Y} \to \mathbb{R}$ there holds

$$\int_{\mathbf{Y}} f(\mathbf{y})\, d\boldsymbol{\gamma}(\mathbf{y}) = \int_{X_1} \left(\int_{\Psi^{-1}(x_1)} f(\mathbf{y})\, d\mu_{x_1}(\mathbf{y}) \right) d\nu(x_1). \tag{1.3.3}$$

The most interesting case of the disintegration theorem comes when $\mathbf{Y}$ is a product space and Ψ is a projection map on one of its factors: let $\mathbf{Y} = X_1 \times X_2$ and $\Psi = \pi^1$, so that $\Psi^{-1}(x_1) = \{x_1\} \times X_2$ can be identified with X_2 for any $x_1 \in X_1$. This way μ_{x_1} can be seen as concentrated in X_2, for ν-a.e $x_1 \in X_1$ and (1.3.2) becomes

$$\boldsymbol{\gamma}(\mathbf{A}) = \int_{X_1} \mu_{x_1}\{x_2 : (x_1, x_2) \in \mathbf{A}\}\, d\nu(x_1), \quad \forall \mathbf{A} \in \mathcal{B}(X_1 \times X_2). \tag{1.3.4}$$

Next we investigate the relation between weak limits of a sequence of the form $(\nu \bigotimes \mu^n_{x_1}) \subset \mathscr{P}(X_1 \times X_2)$ and weak limits of the disintegrations $\mu^n_{x_1}$. It is clear that, if $\mu^n_{x_1} \rightharpoonup \mu_{x_1}$ for ν-a.e. $x_1 \in X_1$, then $\nu \bigotimes \mu^n_{x_1} \rightharpoonup \nu \bigotimes \mu_{x_1}$ in $\mathscr{P}(X_1 \times X_2)$. On the other hand we have the following

Lemma 1.3.3 *Let $x_1 \in X_1 \mapsto \mu_{x_1} \in \mathscr{P}(X_2)$ be a Borel map. Let $\gamma^n = \nu \bigotimes \mu^n_{x_1}$ weakly converge to γ in $\mathscr{P}(X_1 \times X_2)$. Let $G_{x_1} \subset \mathscr{P}(X_2)$ be set of weak limit points of the sequence $(\mu^n_{x_1})$. If μ_{x_1} is the disintegration of γ with respect to ν, then*

$$\mu_{x_1} \subset \overline{\mathrm{Conv}} G_{x_1} \quad \textit{for } \nu\textit{-a.e. } x_1 \in X_1.$$

Proof. Let $f \in C^0_b(X_2)$. Possibly adding a constant, we can assume that f is positive. Let $A \subset X_1$ be closed, so that $\chi_A(x_1) f(x_2)$ is upper semicontinuous. By the characterization of tightness given in Proposition 1.1.8, we can find a function $\varphi : X_2 \to [0, +\infty]$ with compact sublevels such that

$$\sup_{n \in \mathbb{N}} \int_{X_1 \times X_2} \varphi(x_2)\, d\gamma^n(x_1, x_2) = C < +\infty.$$

Then, by (1.1.3) and Fatou Lemma we have

$$\begin{aligned}\int_{A\times X_2} f(x_2)\,d\gamma(x_1,x_2)+\varepsilon C &\geq \limsup_{n\to\infty}\int_{A\times X_2}(f(x_2)+\varepsilon\varphi(x_2))\,d\gamma^n(x_1,x_2)\\ &\geq \liminf_{n\to\infty}\int_A\left(\int_{X_2}(f(x_2)+\varepsilon\varphi(x_2))\,d\mu^n_{x_1}(x_2)\right)d\nu(x_1)\\ &\geq \int_A\left(\inf_{\varepsilon>0}\liminf_{n\to\infty}\int_{X_2}(f(x_2)+\varepsilon\varphi(x_2))\,d\mu^n_{x_1}(x_2)\right)d\nu(x_1).\end{aligned}$$

By arbitrariness of A we find

$$\int_{X_2} f(x_2)\,d\mu_{x_1}(x_2)\geq \inf_{\varepsilon>0}\liminf_{n\to\infty}\int_{X_2}(f(x_2)+\varepsilon\varphi(x_2))\,d\mu^n_{x_1}(x_2)\quad\text{for } \nu\text{-a.e. } x_1\in X_1. \tag{1.3.5}$$

Since φ has compact sublevels, we can write

$$\liminf_{n\to\infty}\int_{X_2}(f(x_2)+\varepsilon\varphi(x_2))\,d\mu^n_{x_1}(x_2)\geq \inf_{\sigma\in G_{x_1}}\int_{X_2} f(x_2)\,d\sigma(x_2)\quad\text{for } \nu\text{-a.e. } x_1\in X_1.$$

Hence (1.3.5) gives

$$\int_{X_2} f(x_2)\,d\mu_{x_1}(x_2)\geq \inf_{\sigma\in G_{x_1}}\int_{X_2} f(x_2)\,d\sigma(x_2)\quad\text{for } \nu\text{-a.e. } x_1\in X_1. \tag{1.3.6}$$

Choose now a sequence of bounded Lipschitz functions approximating f pointwise. There exists a ν-negligible set $N\in X_1$ such that (1.3.6) holds for any function in such sequence and any $x_1\in X_1\backslash N$, and this is enough to conclude that (1.3.6) holds for each $f\in C^0_b(X_2)$ (see [4, §5.1]). The Hahn-Banach theorem now implies that $\mu_{x_1}\subset\overline{\mathrm{Conv}G_{x_1}}$ for any $x_1\in X_1\backslash N$. □

1.4 Hilbertian framework

Suppose in this section that X is a separable Hilbert space and let $\gamma\in\mathscr{P}(X)$. Let $\mu\in\mathscr{P}(X)$ be absolutely continuous with respect to γ, and let $\rho: X\to\mathbb{R}$ be its density. Then we can apply the disintegration theorem to the function ρ, as shown with the next lemma (recall the notation (1.2.7)).

Lemma 1.4.1 *Let $\mu\ll\gamma$. There holds $\mu^d\ll\gamma^d$, and the corresponding density is given by*

$$\int_X \rho(x)\,d\gamma_{x^d}(x),$$

where γ_{x^d} is the family of measures, concentrated on $(\Pi^d)^{-1}(x^d)$, which disintegrates γ with respect to γ^d.

Proof. Let $A^d \in \mathcal{B}(X^d)$. Let $A \in \mathcal{B}(X)$ be the cylindrical set defined as $A = \{x \in X : (\langle x, \mathbf{e}_1\rangle, \ldots, \langle x, \mathbf{e}_d\rangle) \in A^d\}$. Applying Theorem 1.3.2 (here x^d will denote the variable in X^d) we find

$$\begin{aligned}\mu^d(A^d) &= \int_{X^d} \chi_{A^d}(x^d)\, d\Pi^d_\# \mu(x^d) = \int_X \chi_{A^d}(\Pi^d(x))\, d\mu(x) = \int_X \chi_A(x)\rho(x)\, d\gamma(x)\\ &= \int_{X^d} \left(\int_{\Pi^{-1}(x^d)} \chi_A(x)\rho(x)\, d\gamma_{x^d}(x)\right) d\gamma^d(x^d)\\ &= \int_{X^d} \chi_{A^d}(x^d) \left(\int_{\Pi^{-1}(x^d)} \rho(x)\, d\gamma_{x^d}(x)\right) d\gamma^d(x^d)\\ &= \int_{A^d} \left(\int_{\Pi^{-1}(x^d)} \rho(x)\, d\gamma_{x^d}(x)\right) d\gamma^d(x^d),\end{aligned}$$

which is the thesis. □

We are led to the following

Definition 1.4.2 (Cylindrical projections) *Let $\rho \in L^1(X, \gamma)$. If $\mu = \rho\gamma$, then $\mu^d \ll \gamma^d$ and its density ρ^d is explicitly given by*

$$\rho^d(x^d) = \int_X \rho(x)\, d\gamma_{x^d}(x), \tag{1.4.1}$$

as shown in Lemma 1.4.1. *We shall call ρ^d cylindrical projection of ρ in d dimension.*

In addition, using for instance (1.4.1), one can prove that if $u \in L^p(X, \gamma)$, $p \in [1, +\infty)$, then $u^d \in L^p(X, \gamma^d)$ and

$$u^d \circ \Pi^d \to u \quad \text{in } L^p(X, \gamma) \text{ as } d \to \infty. \tag{1.4.2}$$

Notice that (1.4.1) makes sense (componentwise) also for maps u taking values in X, and if $u \in L^p(X, \gamma; X)$, then $u^d \circ \Pi^d \to u$ in $L^p(X, \gamma; X)$.

Next we define the cylindrical functions.

Definition 1.4.3 (Smooth cylindrical functions) *We say that $\varphi : X \to \mathbb{R}$ is a smooth cylindrical function if $\varphi = \psi \circ \Pi^d$, where Π^d is a projection map and $\psi \in C_c^\infty(\mathbb{R}^d)$. The set of smooth cylindrical functions on X will be denoted by* Cyl(X).

Definition 1.4.4 (Gaussian measures) *A measure $\mu \in \mathscr{P}(X)$ is a nondegenerate Gaussian measure if, for any projection map Π^d, the induced measure $\mu^d = \Pi^d_\# \mu$ on X^d is the law of a Gaussian random variable, that is,*

$$\mu(A) = \frac{1}{\sqrt{(2\pi)^d \det Q}} \int_A e^{(x-m)^T Q^{-1}(x-m)/2}\, dx \tag{1.4.3}$$

for any Borel set $A \subset X^d$, for some mean vector m and positive definite symmetric covariance matrix Q.

Definition 1.4.5 (Regular measures) *We say that $\mu \in \mathscr{P}(X)$ is regular, and write $\mu \in \mathscr{P}^r(X)$, if $\mu(B) = 0$ whenever B is a Gaussian null set, that is $\nu(B) = 0$ for every nondegenerate Gaussian measure ν. Moreover, we denote by $\mathscr{P}^r_p(X)$ the set of probability measures on X which are both regular and of finite p-moment.*

Remark 1.4.6 In $\mathbb{R}^d$, if μ is a Gaussian measure then $\mu \ll \mathscr{L}^d$, the density being given itself by (1.4.3). So, if X is an Euclidean space, a measure is regular simply if it is absolutely continuous with respect to $\mathscr{L}^d$.

Next we introduce a weaker convergence for probability measure than the narrow one, suitable to have local compactness in $\mathscr{P}(X)$ when X is a separable Hilbert space. In this case, the standard weak topology of X is not metrizable. Hence, we introduce a topology on X which no more makes it a complete space (here this is not crucial), but it is induced by a norm. Recall that our theory works in general separable metric Radon spaces, even non complete.

Definition 1.4.7 (A weak topology for X) *We define, for $x \in X$, the norm*

$$\|x\|^2_\varpi = \sum_{j=1}^{\infty} \frac{1}{j^2} \langle x, \mathbf{e}_j \rangle^2. \tag{1.4.4}$$

Notice that bounded sequences of X are compact with respect to the new norm. To see this, let $\sup_n |x_n| < +\infty$ and let x_{n_k} be a subsequence weakly converging to x. Since for any $j \in \mathbb{N}$ we have $\langle x_{n_k} - x, \mathbf{e}_j \rangle \to 0$, we have

$$\lim_{k\to\infty} \|x_{n_k}\|^2_\varpi = \lim_{k\to\infty} \sum_{j=1}^{\infty} \frac{1}{j^2} \langle x_{n_k} - x, \mathbf{e}_j \rangle^2 = 0$$

by the dominated convergence theorem. The topology induced by $\|\cdot\|_\varpi$ is in fact weaker than the standard weak topology of the Hilbert space, but the two coincide on bounded sets. The space X, endowed with the $\|\cdot\|_\varpi$ norm, will be denoted by X_ϖ (the notation is borrowed from [4]). Notice also that the narrow topology of $\mathscr{P}(X_\varpi)$ is weaker than the standard narrow topology in $\mathscr{P}(X)$, since of course continuous functions in $C^0_b(X_\varpi)$ are less than the ones of $C^0_b(X)$. A set $\Xi \in \mathscr{P}(X)$ is tight with respect to the weak topology of $\mathscr{P}(X_\varpi)$ if and only if there exists a Borel function $\varphi : X \to [0, +\infty]$, with $h \to +\infty$ as $|x| \to +\infty$, such that

$$\sup_{\mu \in \Xi} \int_X \varphi(x)\, d\mu(x) < +\infty. \tag{1.4.5}$$

Notice that we no longer require the compactness of the sublevels of φ, as for the standard tightness in $\mathscr{P}(X)$ in the characterization given by Proposition 1.1.8 (here the proof is analogous).

We end the section giving the definition of barycentric projection.

Definition 1.4.8 (Barycenter) *Let $\gamma = \int_X \gamma_{x_1}\, d\mu(x_1)$ be a plan in $\mathscr{P}(X \times X)$ with first marginal $\pi^1_\# \mu$. Suppose that γ_x has finite first moment for μ-a.e. $x \in X$. The barycenter $\bar{\gamma} : X \to X$ of γ is defined, for μ-a.e. $x \in X$, by*

$$\bar{\gamma}(x_1) = \int_X x_2\, d\gamma_{x_1}(x_2).$$

1.5 Convergence results

In this section we state some convergence results involving measures, transport maps and transport plans. We begin with a result describing the convergence of measures of the form $\boldsymbol{\rho}_{n\#}\mu_n$.

Lemma 1.5.1 *Let $\boldsymbol{\rho}, \boldsymbol{\rho}_n : X \to Y$, $n \in \mathbb{N}$, be Borel maps such that $\boldsymbol{\rho}_n \to \boldsymbol{\rho}$ uniformly on compact subsets of X, and let $\boldsymbol{\rho}$ be continuous. If $(\mu_n) \subset \mathscr{P}(X)$ weakly converges to $\mu \in \mathscr{P}(X)$, then $\boldsymbol{\rho}_{n\#}\mu_n \rightharpoonup \boldsymbol{\rho}_\#\mu$. In particular, we have $\boldsymbol{\pi}_\#\mu_n \rightharpoonup \boldsymbol{\pi}_\#\mu$ if $\boldsymbol{\pi}$ is a projection map on a component of X (or on a subspace, if X is a linear space).*

Proof. Let $f : Y \to \mathbb{R}$ be bounded, Lipschitz and nonnegative. Since (μ_n) is tight, there exist a sequence (K_m) of compact subsets of X such that $\sup_{n\in\mathbb{N}} \mu_n(X \backslash K_m)$ goes to zero as m goes to infinity. Moreover, $(f \circ \boldsymbol{\rho}_n)$ is a sequence converging uniformly to $f \circ \boldsymbol{\rho}$ on each K_m. By (1.1.2) we have

$$\begin{aligned}
\liminf_{n\to\infty} \int_X f \circ \boldsymbol{\rho}_n\, d\mu_n &\geq \liminf_{n\to\infty} \int_{K_m} f \circ \boldsymbol{\rho}\, d\mu_n \\
&\geq (-\sup_{x\in X} f) \sup_{n\in\mathbb{N}} \mu_n(X \backslash K_m) + \liminf_{n\to\infty} \int_X f \circ \boldsymbol{\rho}\, d\mu_n \\
&\geq (-\sup_{x\in X} f) \sup_{n\in\mathbb{N}} \mu_n(X \backslash K_m) + \int_X f \circ \boldsymbol{\rho}\, d\mu.
\end{aligned}$$

Passing to the limit as $m \to \infty$, we get the liminf inequality. If f has sign, simply add a constant to f. The limsup inequality follows changing f with $-f$. As noticed is Section 1.1, it is enough to check (1.1.1) on Lipschitz functions. □

Lemma 1.5.2 *Let X be a separable Hilbert space. Let $p > 1$ and $q = p/(p-1)$. Let $(\gamma_n) \in \mathscr{P}(X \times X)$ be a sequence weakly converging to γ in $\mathscr{P}(X \times X_\varpi)$ and let*

$$\sup_{n\in\mathbb{N}} \int_{X\times X} \left(|x_1|^p + |x_2|^q\right) d\gamma_n(x_1, x_2) < +\infty. \tag{1.5.1}$$

If $(\pi^1_\#\gamma_n)$ has uniformly integrable p-moments or $(\pi^2_\#\gamma_n)$ has uniformly integrable q-moments, then

$$\lim_{n\to\infty} \int_{X\times X} \langle x_1, x_2 \rangle\, d\gamma_n(x_1, x_2) = \int_{X\times X} \langle x_1, x_2 \rangle\, d\gamma(x_1, x_2).$$

Proof. Suppose that $(\pi^1_\# \gamma_n)$ has uniformly integrable p-moments (the other case is analogous). Let $m, k \in \mathbb{N}$ and $C^q := \sup_{n\in\mathbb{N}} \int_X |x_2|^q \, d\gamma_n(x_1, x_2)$. Clearly if $|x_1||x_2| \geq k$ and $|x_1| \leq m$, then $|x_2| \geq k/m$, and then

$$\begin{aligned}\int_{\{|x_1||x_2|\geq k\}} |x_1||x_2| \, d\gamma_n(x_1, x_2) \leq & m \int_{\{|x_2|\geq k/m\}} |x_2| \, d\pi^2_\# \gamma_n(x_2) \\ & + C \left(\int_{\{|x_1|\geq m\}} |x_1|^p \, d\pi^1_\# \gamma_n(x_1) \right)^{1/p}.\end{aligned}$$

By (1.5.1) and (1.1.9), $(\pi^2_\# \gamma_n)$ has uniformly integrable 1-moments, hence

$$\limsup_{k\to\infty} \sup_{n\in\mathbb{N}} \int_{\{|x_1||x_2|\geq k\}} |x_1||x_2| \, d\gamma_n(x_1, x_2) \leq \sup_{n\in\mathbb{N}} C \left(\int_{\{|x_1|\geq m\}} |x_1|^p \, d\pi^1_\# \gamma_n(x_1) \right)^{1/p} \tag{1.5.2}$$

and the left hand side goes to 0 for $m \to \infty$, since $(\pi^1_\# \gamma_n)$ has uniformly integrable p-moments. Notice that the map $(x_1, x_2) \mapsto \langle x_1, x_2 \rangle$ is not continuous over $X \times X_\varpi$ (that is, in the $\|\cdot\| \times \|\cdot\|_\varpi$ topology), but it is continuous in the same topology when restricted to $X \times \overline{B_r}(\mathbf{0})$ (on $\overline{B_r}(\mathbf{0})$ the $\|\cdot\|_\varpi$ topology coincides with the standard weak one). Let $k, l > 0$ and let

$$f(x_1, x_2) = \langle x_1, x_2 \rangle, \quad f_k(x_1, x_2) = (-k) \vee \langle x_1, x_2 \rangle, \quad f_{k,l} = (-k) \vee \langle x_1, x_2 \rangle \wedge l.$$

Of course $f_{k,l}$ is itself continuous on $X \times \overline{B_r}(\mathbf{0})$ in the $X \times X_\varpi$ topology. We consider a l.s.c. extension over the whole space $X \times X$, defined as

$$g_{k,l}(x_1, x_2) = \begin{cases} f_{k,l} & \text{if } x_2 \in \overline{B_r}(\mathbf{0}), \\ l & \text{if } x_2 \in X \setminus \overline{B_r}(\mathbf{0}). \end{cases}$$

By lower semicontinuity $g_{k,l}$ satisfies (1.1.2), and since $g_{k,l} \geq f_{k,l}$, we have

$$\begin{aligned}\liminf_{n\to\infty} \int_{X\times X} f_{k,l} \, d\gamma_n &= \liminf_{n\to\infty} \left(\int_{X\times X} g_{k,l} \, d\gamma_n + \int_{(X\times X)\setminus(X\times \overline{B_r}(\mathbf{0}))} (f_{k,l} - g_{k,l}) \, d\gamma_n \right) \\ &\geq \int_{X\times X} f_{k,l} \, d\gamma - (k + l) \limsup_{n\to\infty} \gamma_n \left((X \times X) \setminus (X \times \overline{B_r}(\mathbf{0})) \right).\end{aligned}$$

We let $r \to \infty$. The last term vanishes thanks to (1.5.1), which yields in particular $\lim_{r\to\infty} \sup_{n\in\mathbb{N}} \gamma_n \left((X \times X) \setminus (X \times \overline{B_r}(\mathbf{0})) \right) = 0$. Then we let $l \to \infty$, making use of the dominated convergence theorem in the right hand side, and since for any k and l there holds $f_k \geq f_{k,l}$, we get

$$\liminf_{n\to\infty} \int_{X\times X} f_k \, d\gamma_n \geq \int_{X\times X} f_k \, d\gamma. \tag{1.5.3}$$

Next consider that

$$0 \geq \liminf_{n\to\infty} \int_{X\times X} (f - f_k) \, d\gamma_n = -\limsup_{n\to\infty} \int_{X\times X} (f_k - f) \, d\gamma_n \geq -\sup_{n\in\mathbb{N}} \int_{X\times X} (f_k - f) \, d\gamma_n,$$

and since

$$\sup_{n\in\mathbb{N}}\int_{X\times X}(f_k(x_1,x_2)-f(x_1,x_2))\,d\gamma_n(x_1,x_2)\le \sup_{n\in\mathbb{N}}\int_{\{|x_1||x_2|>k\}}|x_1||x_2|\,d\gamma_n(x_1,x_2),$$

by (1.5.2) we see that $\liminf_{n\to\infty}\int_{X\times X}(f-f_k)\,d\gamma_n$ goes to zero as k goes to infinity. Hence, passing to the limit with respect to k in

$$\liminf_{n\to\infty}\int_{X\times X} f\,d\gamma_n = \liminf_{n\to\infty}\left(\int_{X\times X} f_k\,d\gamma_n + \int_{X\times X}(f-f_k)\,d\gamma_n\right), \tag{1.5.4}$$

with (1.5.3) and with the dominated convergence theorem again, we find

$$\liminf_{n\to\infty}\int_{X\times X}\langle x_1,x_2\rangle\,d\gamma_n(x_1,x_2)\ge\int_{X\times X}\langle x_1,x_2\rangle\,d\gamma(x_1,x_2).$$

In order to obtain the corresponding limsup inequality, one changes sign in (1.5.3), obtaining the limsup inequality for $\langle x_1,x_2\rangle\wedge k$, and then the conclusion is analogous. □

In order to deal with sequences of pairs $(\boldsymbol{\rho}_n,\mu_n)$, where $\boldsymbol{\rho}_n\in L^p(X,\mu_n;X)$ and μ_n are measures on a Hilbert space X, we will need the following notion of convergence.

Definition 1.5.3 *Let X be a separable Hilbert space. Let $(\mu_n)\in\mathscr{P}(X)$ be weakly convergent to μ. Let $\boldsymbol{\rho}_n\in L^1(X,\mu_n;X)$ and $\boldsymbol{\rho}\in L^1(X,\mu;X)$. We say that $\boldsymbol{\rho}_n$ weakly converge to $\boldsymbol{\rho}$ if*

$$\lim_{n\to\infty}\int_X\zeta(x)\rho_n^j d\mu_n(x)=\int_X\zeta(x)\rho^j\,d\mu(x) \tag{1.5.5}$$

for any $\zeta\in\mathrm{Cyl}(X)$ and $j\in\mathbb{N}$, where ρ_n^j and ρ^j are respectively the components of $\boldsymbol{\rho}$ and $\boldsymbol{\rho}_n$ along the basis $(\mathbf{e}_j)$.
We say that $\boldsymbol{\rho}_n$ strongly converge to $\boldsymbol{\rho}$ in L^p, $p>1$, if in addition it holds

$$\lim_{n\to\infty}\|\boldsymbol{\rho}_n\|_{L^p(X,\mu_n;X)}=\|\boldsymbol{\rho}\|_{L^p(X,\mu;X)}. \tag{1.5.6}$$

Analogously, in the scalar case we say that $\rho_n\in L^1(X,\mu_n)$ weakly converge to $\rho\in L^1(X,\mu)$ if

$$\lim_{n\to\infty}\int_X\zeta(x)\rho_n d\mu_n(x)=\int_X\zeta(x)\rho d\mu(x)\qquad\forall\zeta\in\mathrm{Cyl}(X), \tag{1.5.7}$$

and strongly if moreover

$$\|\rho_n\|_{L^p(X,\mu_n)}\to\|\rho\|_{L^p(X,\mu)},\quad n\to+\infty.$$

We now state the related convergence lemma.

Lemma 1.5.4 *Let $p>1$. Let $(\mu_n)\subset\mathscr{P}(X)$ be weakly converging to μ in $\mathscr{P}(X_\varpi)$, and let $(\boldsymbol{\rho}_n)\subset L^p(X,\mu_n;X)$ be such that*

$$\sup_{n\in\mathbb{N}}\int_X|\boldsymbol{\rho}_n(x)|^p\,d\mu_n(x)<+\infty. \tag{1.5.8}$$

Let $\gamma_n:=(\mathbf{I},\boldsymbol{\rho}_n)_\#\mu_n$. Then *i) the sequence* $(\gamma_n) \subset \mathscr{P}(X \times X)$ *has limit points in* $\mathscr{P}(X_\varpi \times X_\varpi)$.

ii) If (γ_{n_k}) *is a subsequence converging weakly in* $\mathscr{P}(X_\varpi \times X_\varpi)$ *to* γ, *then* $\boldsymbol{\rho}_{n_k}$ *converges weakly as in* Definition 1.5.3 *to the barycenter of* γ *and*

$$\int_X g(\bar\gamma(x))\, d\mu(x) \le \liminf_{k\to\infty} \int_X g(\boldsymbol{\rho}_{n_k})\, d\mu_{n_k}(x) \tag{1.5.9}$$

for any convex and weakly l.s.c. function $g : X \to (-\infty, +\infty]$. *In particular,* $\boldsymbol{\rho}_n$ *weakly converges to* $\boldsymbol{\rho}$ *in the sense of* Definition 1.5.3 *if and only if* $\boldsymbol{\rho}$ *is the barycenter of every limit point in* $\mathscr{P}(X_\varpi \times X_\varpi)$ *of the sequence* (γ_n).

iii) If moreover $\boldsymbol{\rho}_n$ *strongly converge to* $\boldsymbol{\rho}$ *in* L^p *in the sense of* Definition 1.5.3, *then* γ_n *weakly converges to* $(\mathbf{I}, \boldsymbol{\rho})_\# \mu$ *in* $\mathscr{P}(X_\varpi \times X_\varpi)$ *and*

$$\lim_{n\to\infty} \int_{X\times X} |x_2|^p\, d\gamma_n(x_1, x_2) = \|\boldsymbol{\rho}\|^p_{L^p(X,\mu;X)}.$$

Proof. The first marginals of γ_n are weakly convergent in $\mathscr{P}(X_\varpi)$, hence tight in $\mathscr{P}(X_\varpi)$. The second ones are $\boldsymbol{\rho}_{n\#}\mu_n$, and (1.5.8) says that

$$\sup_{n\in\mathbb{N}} \int_X |x|^p\, d(\boldsymbol{\rho}_{n\#}\mu_n)(x) < +\infty.$$

So, by the characterization of tightness in $\mathscr{P}(X_\varpi)$ given by (1.4.5), $(\pi^2_\#\gamma_n)$ is also tight in $\mathscr{P}(X_\varpi)$. Then the statement of *i)* follows by Proposition 1.2.2.

We are tacitely assuming, with no loss of generality, that the reference orthonormal system in X_ϖ is the same $\{\mathbf{e}_j\}_{j\in\mathbb{N}}$ we introduced for X. This way, $\mathrm{Cyl}(X) \subset C^0(X_\varpi)$ and, for any $j \in \mathbb{N}$, the map $x \mapsto \langle \mathbf{e}_j, x\rangle$ belongs to $C^0_b(X_\varpi)$. Hence, if (γ_{n_k}) denotes a subsequence of (γ_n) which converges weakly to γ in $\mathscr{P}(X_\varpi \times X_\varpi)$, for any $j \in \mathbb{N}$ and any $\varphi \in \mathrm{Cyl}(X)$ there holds

$$\begin{aligned}\lim_{k\to\infty} \int_X \varphi(x)\langle \mathbf{e}_j, \boldsymbol{\rho}_{n_k}(x)\rangle\, d\mu_{n_k}(x) &= \lim_{k\to\infty} \int_{X\times X} \varphi(x_1)\langle \mathbf{e}_j, x_2\rangle\, d\gamma_{n_k}(x_1, x_2)\\ &= \int_{X\times X} \varphi(x_1)\langle \mathbf{e}_j, x_2\rangle\, d\gamma(x_1, x_2).\end{aligned}$$

Notice that, disintegrating γ with respect to its first marginal μ, and denoting by $\bar\gamma$ its barycenter (see Definition 1.4.8), the last term is equal to

$$\int_X \varphi(x)\langle \mathbf{e}_j, \bar\gamma(x)\rangle\, d\mu(x).$$

This shows the weak convergence of $\boldsymbol{\rho}_{n_k}$ to $\bar\gamma$. Then it is clear that if the full sequence $(\boldsymbol{\rho}_n)$ converges weakly to $\boldsymbol{\rho}$, the barycenter of any $\mathscr{P}(X_\varpi \times X_\varpi)$ limit of γ_n has to be equal to $\boldsymbol{\rho}$. In this case, let g be l.s.c. and convex (this implies that it is also l.s.c. in the weak $\mathscr{P}(X_\varpi)$

topology). By definition of barycenter, Jensen inequality, disintegration and (1.1.2), we have

$$\begin{aligned}\int_X g(\boldsymbol{\rho})\,d\mu &= \int_X g(\bar\gamma)\,d\mu = \int_X g\left(\int_X x_2\,d\gamma_{x_1}(x_2)\right)\,d\mu(x_1)\\ &\le \int_X\int_X g(x_2)\,d\gamma_{x_1}(x_2)\,d\mu(x_1) = \int_{X\times X} g(x_2)\,d\gamma(x_1,x_2)\\ &\le \liminf_{n\to\infty}\int_{X\times X} g(x_2)\,d((\mathbf{I},\boldsymbol{\rho}_n)_{\#}\mu_n)(x_1,x_2) = \liminf_{n\to\infty}\int_X g(\boldsymbol{\rho}_n)\,d\mu_n,\end{aligned} \tag{1.5.10}$$

which proves (1.5.9).

Finally we show *iii)*. For, let γ be a generic weak $\mathscr{P}(X_\varpi\times X_\varpi)$ limit point of $\gamma_n = (\mathbf{I},\boldsymbol{\rho}_n)_{\#}\mu_n$. Since $\boldsymbol{\rho}_n\to\boldsymbol{\rho}$ weakly as in Definition 1.5.3, we have just seen in point *ii)* that $\boldsymbol{\rho}=\bar\gamma$. Then, by strong convergence of $\boldsymbol{\rho}_n$ to $\boldsymbol{\rho}$ we have

$$\lim_{n\to\infty}\int_X |\boldsymbol{\rho}_n|^p\,d\mu_n = \int_X |\bar\gamma|^p\,d\mu(x). \tag{1.5.11}$$

Choosing $g(\cdot)=|\cdot|^p$ in (1.5.10), by (1.5.11) we see that the Jensen inequality in (1.5.10) is in fact an equality. Since for strictly convex functions like the p-power, $p>1$, this happens only for Dirac masses, we get $\gamma_{x_1}=\delta_{\boldsymbol{\rho}(x_1)}$ for μ-a.e. $x_1\in X$, that is, $\gamma=(\mathbf{I},\boldsymbol{\rho})_{\#}\mu$. Hence the full sequence γ_n converges to $(\mathbf{I},\boldsymbol{\rho})_{\#}\mu$. The strong convergence of $\boldsymbol{\rho}_n$ implies in particular

$$\lim_{n\to\infty}\int_X |x_2|^p\,d\gamma_n(x_1,x_2) = \lim_{n\to\infty}\int_X |\boldsymbol{\rho}_n(x)|^p\,d\mu_n(x) = \int_X |\boldsymbol{\rho}(x)|^p\,d\mu(x)$$

and the proof is completed. □

1.6 Elements of geometric measure theory

In this section we let $X=\mathbb{R}^n$.

Hausdorff measures and rectifiable sets

We recall that a function $f:\Omega\subset\mathbb{R}^n\to\mathbb{R}^m$ is Lipschitz if

$$\mathrm{Lip}(f) := \sup\left\{\frac{|f(x)-f(y)|}{|x-y|}\ : x,y\in\Omega, x\neq y\right\} < +\infty.$$

The quantity $\mathrm{Lip}(f)$ is the Lipschitz constant of f. We state the classical differentiability result (see for instance [34]).

Theorem 1.6.1 (Rademacher) *Let $f:\Omega\subset\mathbb{R}^n\to\mathbb{R}^m$ be a Lipschitz function. Then f is differentiable at x for $\mathcal{L}^n-a.e.x\in\Omega$ (in the classical sense). The a.e. derivative is a $L^\infty(\Omega,\mathcal{L}^n;\mathbb{R}^{nm})$ function which coincides with the distributional derivative of f.*

Definition 1.6.2 (Hausdorff measures) *Let $\Omega \subset \mathbb{R}^n$ and $k \in [0, +\infty)$. The k-dimensional Hausdorff measure of Ω is given by*

$$\mathcal{H}^k(\Omega) := \lim_{\delta \downarrow 0} \mathcal{H}^k_\delta(\Omega),$$

where

$$\mathcal{H}^k_\delta(\Omega) := \frac{\omega_k}{2^k} \left\{ \sum_{i \in I} \operatorname{diam}(\Omega_i)^k : \Omega \in \bigcup_{i \in I} \Omega_i, \operatorname{diam}(\Omega_i) < \delta, I \text{ countable} \right\},$$

ω_k being the volume of the k-dimensional unit ball.

Definition 1.6.3 (Rectifiable sets) *An $\mathcal{H}^k$-measurable set $\Omega \subset \mathbb{R}^n$ is said to be $\mathcal{H}^k$-rectifiable if $\mathcal{H}^k(\Omega) < +\infty$ and there exists countably many Lipschitz functions $f_i : \mathbb{R}^k \to \mathbb{R}^n$ such that*

$$\mathcal{H}^k \left(\Omega \backslash \bigcup_{i=0}^{\infty} f_i(\mathbb{R}^k) \right) = 0.$$

See [3, 35, 36] for other equivalent notions of rectifiability, and as general geometric measure theory references.

Let us introduce the notation for the restriction of measures: if Ω is μ-measurable then $\mu \llcorner \Omega$ is defined by $\mu \llcorner \Omega(A) = \nu(\Omega \cap A)$ for any μ-measurable set A.

Next we give the definition of tangent space for rectifiable sets.

Definition 1.6.4 (Approximate tangent space) *Let $\Omega \subset \mathbb{R}^n$ be an $\mathcal{H}^k$-rectifiable set. Let $k \leq n$. We say that a k-plane P is the approximate tangent space* $\operatorname{Tan}^k(\Omega, x_0)$ *of Ω at x_0 if*

$$\mathcal{H}^k \llcorner \Omega_{x_0, r} \rightharpoonup \mathcal{H}^k \llcorner P$$

as $r \to 0$, where $\Omega_{x_0,r}$ is the rescaled set $\frac{1}{r}(\Omega - x_0)$.

Tangential derivative to curves and manifolds

Let M be an $\mathcal{H}^k$-rectifiable set, $k \leq n$. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a C^1 function. Let $x \in M$ be a point where the approximate tangent space exists and let h be a vector in $\operatorname{Tan}^k(M, x)$. The directional derivative of f at x is defined by

$$\nabla_h f(x) := \langle \nabla \phi(x), h \rangle.$$

The *tangential gradient* of ∇^M of ϕ at x is defined by

$$\nabla^M \phi(x) := \sum_{j=1}^{k} \nabla_{\tau_j} \phi(x) \tau_j,$$

where $\{\tau_j\}_{j\in\{1,\dots,k\}}$ is an orthonormal basis of $\mathrm{Tan}^k(M,x)$. Hence $\nabla^M\phi$ is just the orthogonal projection of $\nabla\phi$ on the approximate tangent space. If ϕ takes values in $\mathbb{R}^m$ we have

$$\nabla_h\phi(x) = \sum_{j=1}^{m}\langle\nabla\phi_j(x), h\rangle \mathbf{e}_j.$$

Here ϕ_j is the j-th component of ϕ and $\{\mathbf{e}_j\}_{j\in\{1,\dots,m\}}$ is an orthonormal basis for $\mathbb{R}^m$.

Notice that the above derivatives are defined with respect to the usual tangent space T_xM when M is a C^1 k-manifold.

Definition 1.6.5 (Tangential divergence) *Let M be as above and $\phi \in C^1(\mathbb{R}^n;\mathbb{R}^m)$. The tangential divergence of ϕ on M is defined as*

$$\operatorname{div}^M\phi(x) = \sum_{j=1}^{m}\langle\nabla^M\phi_j(x), \mathbf{e}_j\rangle, \quad x\in M. \tag{1.6.1}$$

For the next result we refer to [3, Theorem 7.31].

Theorem 1.6.6 *Let M be an $\mathcal{H}^k$-rectifiable manifold in $\mathbb{R}^n$. Let Ω be an open set containing M and let $\xi \in C^1_c(\Omega;\mathbb{R}^n)$. Let ϕ_ε be a one parameter group of diffeomorphism on Ω defined as $\phi_\varepsilon(x) = x + \varepsilon\xi(x) + o(\varepsilon)$. Then there holds*

$$\frac{d}{d\varepsilon}\mathcal{H}^k(\phi_\varepsilon(M\cap\Omega))\bigg|_{\varepsilon=0} = \int_{M\cap\Omega}\operatorname{div}^M\xi\, d\mathcal{H}^M. \tag{1.6.2}$$

Definition 1.6.7 (Distributional vector curvature) *The distributional curvature $\mathbf{H}_M$ of an $\mathcal{H}^k$-rectifiable manifold M in $\mathbb{R}^n$ is the vector defined by*

$$\int_M\langle\xi, \mathbf{H}_M\rangle\, d\mathcal{H}^k = -\int_M \operatorname{div}^M\xi\, d\mathcal{H}^k \quad \forall\xi\in C^1_c(\mathbb{R}^n;\mathbb{R}^n).$$

In the case of a smooth manifold, $\mathbf{H}_M$ is a normal vector, directed towards the center of curvature, with modulus equal to the standard mean curvature. It can be defined at $x \in M$ by $-(\operatorname{div}^M\mathbf{n}(x))\mathbf{n}(x)$, where $\mathbf{n}$ is any C^1 normal vector field. In the case of an $\mathcal{H}^1$-rectifiable set, $\mathbf{H}_M$ is given by $\delta_A\boldsymbol{\tau}$ at any endpoint A, where $\boldsymbol{\tau}$ is the tangent vector in A (pointing to the side of the curve itself). Similarly in correspondence of corner points. For the details on this subject, we refer to [18, 38].

Chapter 2

Optimal transportation

2.1 Formulation of the problem and well-posedness issues

Let $\mu \in \mathscr{P}(X)$, $\nu \in \mathscr{P}(Y)$. Let $c : X \times Y \to [0, +\infty]$ be a Borel cost function. Given a map $\mathbf{t} : X \to Y$ transporting μ to ν, the corresponding transport cost is defined as

$$\int_X c(x, \mathbf{t}(x))\, d\mu(x), \tag{2.1.1}$$

so it is an average of the costs for transporting the point x to the point $\mathbf{t}(x)$. Monge's optimal transportation problem asks to find the map which minimizes the transport cost, so it is formulated as

$$\inf \left\{ \int_X c(x, \mathbf{t}(x))\, d\mu(x) : \mathbf{t}_{\#}\mu = \nu \right\}. \tag{2.1.2}$$

A relaxed formulation, involving transport plans rather than maps, is given by the Kantorovich problem (we refer to [44, 45]), that is

$$\inf \left\{ \int_{X\times Y} c(x, y)\, d\gamma(x, y) : \gamma \in \Gamma(\mu, \nu) \right\}. \tag{2.1.3}$$

First of all, notice that, as soon as the cost function c is l.s.c., the direct method immediately gives a solution to (2.1.3). Indeed, the set $\Gamma(\mu, \nu)$ (which has been defined in Section 1.2) is nonempty and compact in $\mathscr{P}(X \times Y)$, as shown in Lemma 1.2.1. If γ is solution to (2.1.3), we say that it is an optimal transport plan, or that $\gamma \in \Gamma_0(\mu, \nu)$, where

$$\Gamma_0(\mu, \nu) := \left\{ \gamma \in \Gamma(\mu, \nu) : \forall \tilde{\gamma} \in \Gamma(\mu, \nu), \int_{X\times Y} c(x, y)\, d\gamma(x, y) \le \int_{X\times Y} c(x, y)\, d\tilde{\gamma}(x, y) \right\}. \tag{2.1.4}$$

On the other hand, problem (2.1.2) may be ill posed. In fact, if for example μ is a Dirac mass and ν is not, there exists no transport map between μ and ν. One of the most important questions is to know when an optimal transport plan is induced by a map, so that such map solves Monge's problem.

Next we introduce some concepts which will be needed to answer the question. From now on, c will always be a proper l.s.c. cost function.

Definition 2.1.1 (c-transform) *Let $f : X \to \mathbb{R} \cup \{-\infty, +\infty\}$, $g : Y \to \mathbb{R} \cup \{-\infty, +\infty\}$. The c-transforms of f and g are defined (with the convention $+\infty - (+\infty) = +\infty$) by*

$$\begin{aligned} f^c(y) &:= \inf_{x\in X} c(x,y) - f(x), \quad y \in Y, \\ g^c(x) &:= \inf_{y\in Y} c(x,y) - g(y), \quad x \in X. \end{aligned} \tag{2.1.5}$$

Definition 2.1.2 (c-concavity) *A function $f : X \to \mathbb{R} \cup \{-\infty, +\infty\}$ is c-concave if it is the c-transform of some function $g : Y \to \mathbb{R} \cup \{-\infty, +\infty\}$. On the other hand, a function $g : Y \to \mathbb{R} \cup \{-\infty, +\infty\}$ is c-concave if it is the c-transform of some function $f : X \to \mathbb{R} \cup \{-\infty, +\infty\}$.*

Definition 2.1.3 (c-monotonicity) *Let $K \subset X \times Y$. K is said to be c-monotone if, for every $(x_1, y_1), \dots, (x_n, y_n) \in K$ and every permutation σ of $\{1, \dots, n\}$, there holds*

$$\sum_{i=1}^{n} c(x_i, y_i) \le \sum_{i=1}^{n} c(x_i, y_{\sigma(i)}). \tag{2.1.6}$$

The following two results constitute the main characterization of solutions to (2.1.3).

Theorem 2.1.4 (Dual problem) *The minimum of (2.1.3) is equal to*

$$\sup\left\{\int_X \phi(x)\,d\mu(x) + \int_Y \psi(y)\,d\nu(y) : \phi(x) + \psi(y) \le c(x,y), \phi \in L^1(X,\mu), \psi \in L^1(Y,\nu)\right\}. \tag{2.1.7}$$

Definition 2.1.5 (Kantorovich potential) *A Borel function $\phi \in L^1(X, \mu)$ is said to be a maximal Kantorovich potential if the couple (ϕ, ϕ^c) is a maximizing pair for (2.1.7).*

Theorem 2.1.6 (Optimality conditions) *About the dual problem just introduced, we have the following propositions:*

i) If $\gamma \in \Gamma_0(\mu, \nu)$ and $\int_{X\times Y} c(x,y)\,d\gamma(x,y) < +\infty$, then γ is concentrated on a c-monotone Borel set in $X \times Y$ (and $\operatorname{supp}\gamma$ is c-monotone if c is continuous).

ii) Suppose that $c \neq +\infty$, that

$$\mu\left(\left\{x \in X : \int_Y c(x,y)\,d\nu(y) < +\infty\right\}\right), \quad \nu\left(\left\{y \in Y : \int_X c(x,y)\,d\mu(x) < +\infty\right\}\right) \tag{2.1.8}$$

are greater than zero and that $\gamma \in \Gamma(\mu, \nu)$ is concentrated on a c-monotone Borel set in $X \times Y$. Then there exists a maximal Kantorovich potential φ, and it is c-concave. Moreover γ is optimal and $\int_{X\times Y} c(x,y)\,d\gamma(x,y) < +\infty$.

iii) Under the assumptions of ii), if the supremum in (2.1.3) *is finite, then*

$$\varphi(x) + \varphi^c(y) = c(x,y) \quad \gamma\text{-a.e. in } X \times Y. \tag{2.1.9}$$

iv) Under the assumptions of ii), if φ is a Borel potential in $L^1(X,\mu)$ satisfying (2.1.9), *then $\gamma \in \Gamma_0(\mu,\nu)$.*

Although the results we are going to present hold in more general settings, here and in the next sections, in view of the applications of the next chapters, frequently we will assume that X is a separable Hilbert space and that $Y = X$. In this framework, often we will work with the p-cost function $c(x,y) = |x-y|^p$. For a more complete discussion, and for the proofs of Theorem 2.1.4 and Theorem 2.1.6, we refer to [4, 62, 72, 73].

The next result is crucial, since it gives sufficient conditions for the existence of optimal transport maps.

Theorem 2.1.7 (Existence of the optimal transport map) *Suppose that X is a separable Hilbert space. Let $\mu \in \mathscr{P}^r(X)$, $\nu \in \mathscr{P}(X)$. Then problem* (2.1.3) *possesses a unique solution $\gamma \in \Gamma(\mu,\nu)$, and there exists a Borel map $\mathbf{t} \in L^p(X,\mu;X)$ such that*

$$\gamma = (\mathbf{I}, \mathbf{t})_{\#}\mu.$$

If moreover $\operatorname{supp}\nu$ *is bounded, there exists a locally Lipschitz, c-concave maximal Kantorovich potential φ, with Gateaux differential $\nabla\varphi$, such that*

$$\mathbf{t}(x) = x - p^{1-q}|\nabla\varphi(x)|^{q-2}\nabla\varphi(x) \quad \text{for } \mu\text{-a.e. } x \in X,$$

where $q = p/(p-1)$.

Proof. Let $\gamma \in \Gamma_0(\mu,\nu)$. Let us begin with the case of bounded $\operatorname{supp}\nu$. The strategy is the following: first one makes use of the optimality conditions given by Theorem 2.1.6 to find a locally Lipschitz maximal Kantorovich potential. Second, one has to differentiate it, and so some Rademacher like result has to be invoked to guarantee existence of the derivative outside of sets where μ is null. This way one will show that γ is concentrated on a graph.

Fix a maximal Kantorovich potential ϕ, which exists thanks to Theorem 2.1.6. Let

$$\varphi(x) := \inf\{c(x,y) - \phi^c(y) : y \in \operatorname{supp}\nu\},$$

so that φ is Lipschitz on bounded subsets of X, and since ϕ is a maximal Kantorovich potential, the infimum is achieved and $\varphi = \phi$ μ-a.e. in X. Let

$$\psi(y) := \begin{cases} \phi^c(y) & \text{if } y \in \operatorname{supp}\nu, \\ -\infty & \text{otherwise,} \end{cases}$$

and notice that φ is the c-transform of ψ, hence $\varphi^c = (\psi^c)^c \geq \psi = \phi^c$ on $\operatorname{supp}\nu$. This shows that φ is itself a maximal Kantorovich potential. So, for μ-a.e. $x \in X$, there exists

y such that $\varphi(x) + \varphi^c(y) = |x-y|^p$. Since φ is locally Lipschitz, an extended version of Rademacher theorem (see [13, Theorem 5.11.1]) shows that it is Gateaux differentiable outside a Gaussian null set (by regularity of μ, outside a μ-negligible set). The function $z \mapsto |z-y|^p - \varphi(z)$ attains its minimum for $z = x$ (because $\varphi(x)+\varphi^c(y) = |x-y|^p$ γ-a.e. in $X \times Y$ as φ is a maximal Kantorovich potential). We can Gateaux differentiate it obtaining

$$\nabla(|z-y|^p - \varphi(z)) = p|z-y|^{p-2}(z-y) - \nabla\varphi(z).$$

The Gateaux differential is zero at the minimum, so that

$$p|x-y|^{p-2}(x-y) = \nabla\varphi(x).$$

Inverting this relation we find

$$y = x - p^{1-q}|\nabla\varphi(x)|^{q-2}\nabla\varphi(x), \qquad q := \frac{p}{p-1}.$$

Hence y is uniquely determined by x, for μ-a.e. $x \in X$, but $\varphi(x) + \varphi^c(y) = |x-y|^p$ γ-a.e. in $X \times Y$, so that γ is concentrated on a graph.

If $\operatorname{supp}\nu$ is unbounded, let $\gamma_n = \gamma \llcorner B_n(\mathbf{0})$, $n \in \mathbb{N}$. Notice that $\pi^1_\#\gamma_n$ is regular and that $\pi^2_\#\gamma_n$ is bounded. Since $\operatorname{supp}\gamma$ is $|\cdot|^p$-monotone (see statement *i)* in Theorem 2.1.6), the same is true for $\operatorname{supp}\gamma_n$, therefore γ_n is optimal itself. So, we can apply the previous part of the proof to infer that γ_n is induced by an optimal map $\boldsymbol{\rho}_n$, and clearly $(\mathbf{I}, \boldsymbol{\rho}_n)_\#(\pi^1_\#\gamma_n) \le (\mathbf{I}, \boldsymbol{\rho}_m)_\#(\pi^1_\#\gamma_m)$ if $n < m$, yielding $\boldsymbol{\rho}_n(x) = \boldsymbol{\rho}_m(x)$ for $\pi^1_\#\gamma_n$-a.e. $x \in X$ if $n < m$. So the definition $\boldsymbol{\rho} := \boldsymbol{\rho}_n$ μ_n-a.e., for any n, is consistent. We find

$$\gamma \leftharpoonup \gamma_n = (\mathbf{I}, \boldsymbol{\rho})_\#(\pi^1_\#\gamma_n) \rightharpoonup (\mathbf{I}, \boldsymbol{\rho})_\#\mu,$$

so γ is induced by $\boldsymbol{\rho}$.

In order to show uniqueness of the map, suppose by contradiction that $\boldsymbol{\rho}_1$, $\boldsymbol{\rho}_2$ are distinct optimal maps. Then $(\mathbf{I}, \boldsymbol{\rho}_1)_\#\mu$ is an optimal plan, which can be disintegrated with respect to its first marginal μ, so that it has the form $\int_X \delta_{\boldsymbol{\rho}_1(x)}\,d\mu(x)$, where $\delta_{\boldsymbol{\rho}_1(x)}$ is the Dirac delta on the graph of $\boldsymbol{\rho}_1$. Similarly for $(\mathbf{I}, \boldsymbol{\rho}_2)_\#\mu$. But then $\frac{1}{2}\int_X(\delta_{\boldsymbol{\rho}_1(x)} + \delta_{\boldsymbol{\rho}_2(x)})\,d\mu(x)$ is itself optimal, but not concentrated on a graph, since the union of the graphs of $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$ is of course no more a graph. We have found an optimal plan not induced by a map, a contradiction. □

In order to prove an injectivity result for the optimal transport map, we need the following

Definition 2.1.8 (Inverse plan) *We denote by $\varsigma : X \times Y \to Y \times X$ the map which inverts the components, so $\varsigma(x, y) = (y, x)$. Let $\gamma \in \mathscr{P}(X \times Y)$. We define by $\gamma^{-1} := \varsigma_\#\gamma \in \mathscr{P}(Y \times X)$ the inverse plan of γ, that is, for any Borel bounded function $f : Y \times X \to R$,*

$$\int_{Y\times X} f(y, x)\,d\gamma^{-1}(y, x) = \int_{X\times Y} f(y, x)\,d\gamma(x, y).$$

Lemma 2.1.9 *Let $\mu, \nu \in \mathscr{P}_p^r(X)$, where X is a separable Hilbert space. Then the unique solution to problem* (2.1.3) *is induced by a μ-essentially injective optimal transport map $\mathbf{t}$ between μ and ν (that is, there exists a μ-negligible set $N \subset X$ such that $\mathbf{t}$ is injective when restricted to $X \backslash N$).*

Proof. The existence of a unique optimal transport plan γ solving (2.1.3), and induced by a transport map $\mathbf{t} : X \to X$, is a consequence of Theorem 2.1.7. On the other hand, the inverse plan γ^{-1} (defined in (2.1.8)) belongs to $\Gamma_0(\nu, \mu)$, since the cost function is symmetric. As ν is regular, invoking again Theorem 2.1.7, we deduce that γ^{-1} is the unique element of $\Gamma_0(\nu, \mu)$, and it is induced by a map $\mathbf{s} : X \to X$. We get $(\mathbf{I}, \mathbf{t})_\# \mu = ((\mathbf{I}, \mathbf{s})_\# \nu)^{-1}$, hence $\mathbf{s} \circ \mathbf{t} = \mathbf{I}$ μ-a.e. in X.

□

2.2 The Wasserstein distance

Definition and basic results

Let X be a separable metric Radon space with distance d. The Wasserstein distance between $\mu, \nu \in \mathscr{P}_p(X)$ is defined as the Kantorovich optimal transport cost from μ to ν, that is

$$W_p^p(\mu, \nu) = \inf \left\{ \int_{X \times X} d^p(x_1, x_2)\, d\gamma(x_1, x_2) : \gamma \in \Gamma(\mu, \nu) \right\}. \tag{2.2.1}$$

It is also called the optimal transportation distance. It is easily seen that W_p is symmetric and vanishes only when its arguments coincide. In order to verify the triangle inequality, we need the following

Lemma 2.2.1 (Dudley's lemma) *Let X_1, X_2, X_3 be separable metric Radon spaces. Let $\gamma^{12} \in \mathscr{P}(X_1 \times X_2)$ and $\gamma^{13} \in \mathscr{P}(X_1 \times X_3)$ be two transport plans with same first marginal μ^1. Then there exists $\boldsymbol{\gamma} \in \mathscr{P}(X_1 \times X_2 \times X_3)$ such that*

$$\pi^{1,2}_\# \boldsymbol{\gamma} = \gamma^{12}, \quad \pi^{1,3}_\# \boldsymbol{\gamma} = \gamma^{13}. \tag{2.2.2}$$

If $\gamma^{12}, \gamma^{13}, \boldsymbol{\gamma}$ are disintegrated with respect to μ^1 respectively by the Borel families $\gamma^{12}_{x_1}, \gamma^{13}_{x_1}, \boldsymbol{\gamma}_{x_1}$, (2.2.2) *is equivalent to*

$$\boldsymbol{\gamma}_{x_1} \in \Gamma(\gamma^{12}_{x_1}, \gamma^{13}_{x_1}) \subset \mathscr{P}(X_2 \times X_3) \quad \textit{for } \mu^1\textit{-a.e. } x_1 \in X_1.$$

Proof. Consider the family of product measures $\gamma^{12}_{x_1} \times \gamma^{13}_{x_1} \in \mathscr{P}(X_2 \times X_3)$. It is clear that, defining $\boldsymbol{\gamma}$ as

$$\int_{X_1} \gamma^{12}_{x_1} \times \gamma^{13}_{x_1}\, d\mu^1(x_1),$$

we have

$$\pi^{1,2}_\#\boldsymbol{\gamma} = \int_{X_1} \gamma^{12}_{x_1}\,d\mu^1(x_1), \quad \pi^{1,3}_\#\boldsymbol{\gamma} = \int_{X_1} \gamma^{13}_{x_1}\,d\mu^1(x_1),$$

hence the thesis is achieved. □

Remark 2.2.2 (Composition of plans) Let $\gamma^{12} \in \mathscr{P}(X_1 \times X_2)$ and $\gamma^{23} \in \mathscr{P}(X_2 \times X_3)$ be such that $\pi^2_\#\gamma^{12} = \pi^1_\#\gamma^{23} = \mu^2$. Let $\gamma^{12}_{x_2}$ and $\gamma^{23}_{x_2}$ be their disintegration with respect to μ^2. Then, after Lemma 2.2.1, we know that the plan

$$\boldsymbol{\gamma} := \int_{X_2} \gamma^{12}_{x_2} \times \gamma^{23}_{x_2}\,d\mu^2(x_2)$$

satisfies $\pi^{1,2}_\#\boldsymbol{\gamma} = \gamma^{12}$ and $\pi^{2,3}_\#\boldsymbol{\gamma} = \gamma^{23}$. We define the composed plan $\gamma^{23}\circ\gamma^{12} \in \mathscr{P}(X_1 \times X_3)$ as

$$\gamma^{23} \circ \gamma^{12} := \pi^{1,3}_\#\boldsymbol{\gamma}. \tag{2.2.3}$$

Hence, for any bounded Borel function $f : X_1 \times X_3 \to \mathbb{R}$, we have

$$\int_{X_1\times X_3} f(x_1,x_3)\,d(\gamma^{23}\circ\gamma^{12})(x_1,x_3) = \int_{X_2}\left(\int_{X_1\times X_3} f(x_1,x_3)\,d(\gamma^{12}_{x_2}\times\gamma^{23}_{x_2})(x_1,x_3)\right)d\mu^2(x_2).$$

Notice that this definition is consistent with the one of inverse plan (see Definition 2.1.8), because if $X_3 \equiv X_1$ and $\pi^1_\#\gamma^{12} = \pi^2_\#\gamma^{23} = \mu^1$ we immediately get $\gamma^{23}\circ\gamma^{12} = (\mathbf{I},\mathbf{I})_\#\mu^1$.

Lemma 2.2.3 *The application $(\mu,\nu) \in X \times X \mapsto W_p(\mu,\nu) \in [0,+\infty)$ is a distance on $\mathscr{P}_p(X)$.*

Proof. We show the triangle inequality in the following way. Let $\mu^1,\mu^2,\mu^3 \in \mathscr{P}_p(X)$, let $\gamma^{12} \in \Gamma_0(\mu^1,\mu^2)$ and $\gamma^{23} \in \Gamma_0(\mu^2,\mu^3)$. By Lemma 2.2.1 there exists a transport plan $\boldsymbol{\gamma} \in \mathscr{P}(X\times X\times X)$ such that $\pi^{1,2}_\#\boldsymbol{\gamma} = \gamma^{12}$ and $\pi^{2,3}_\#\boldsymbol{\gamma} = \gamma^{23}$ and $\pi^{1,3}_\#\boldsymbol{\gamma} \in \Gamma(\mu^1,\mu^3)$. In order to give more clarity when dealing with three-plans, we let X_1, X_2, X_3 be three copies of X, and we think to μ^1,μ^2,μ^3 as probability measures on $\mathscr{P}(X_1), \mathscr{P}(X_2), \mathscr{P}(X_3)$ respectively. This way we have $\pi^{1,3}_\#\boldsymbol{\gamma} \in \mathscr{P}(X_1\times X_3)$. Changing variables we get

$$\begin{aligned}
W_p^p(\mu^1,\mu^2) &= \int_{X_1\times X_2} d^p(x_1,x_2)\,d\gamma^{12}(x_1,x_2) = \int_{X_1\times X_2\times X_3} d^p(x_1,x_2)\,d\boldsymbol{\gamma}(x_1,x_2,x_3),\\
W_p^p(\mu^2,\mu^3) &= \int_{X_2\times X_3} d^p(x_2,x_3)\,d\gamma^{23}(x_2,x_3) = \int_{X_1\times X_2\times X_3} d^p(x_2,x_3)\,d\boldsymbol{\gamma}(x_1,x_2,x_3),\\
W_p^p(\mu^1,\mu^3) &\le \int_{X_1\times X_3} d^p(x_1,x_3)\,d\pi^{1,3}_\#\boldsymbol{\gamma}(x_1,x_3) = \int_{X_1\times X_2\times X_3} d^p(x_1,x_3)\,d\boldsymbol{\gamma}(x_1,x_2,x_3).
\end{aligned}$$

Hence

$$\begin{aligned}
W_p(\mu^1,\mu^2) &= d_{L^p(X_1\times X_2\times X_3,\boldsymbol{\gamma};X)}(x_1,x_2),\\
W_p(\mu^2,\mu^3) &= d_{L^p(X_1\times X_2\times X_3,\boldsymbol{\gamma};X)}(x_2,x_3),\\
W_p(\mu^1,\mu^3) &\le d_{L^p(X_1\times X_2\times X_3,\boldsymbol{\gamma};X)}(x_1,x_3),
\end{aligned}$$

and the thesis follow applying the triangle inequality of the $L^p(X_1\times X_2\times X_3,\boldsymbol{\gamma};X)$ distance. □

Remark 2.2.4 Let $\bar{x} \in X$. Notice that $\Gamma(\mu, \delta_{\bar{x}})$ is a singleton for any $\mu \in \mathscr{P}_p(X)$. Indeed, its only element is $(\mathbf{I}, \bar{x})_\# \mu$. Hence, if $\mu, \nu \in \mathscr{P}_p(X)$, there holds

$$\begin{aligned} W_p(\mu,\nu) \le W_p(\mu,\delta_{\bar{x}}) + W_p(\nu,\delta_{\bar{x}}) &= \left(\int_{X\times X} d^p(x_1,x_2)\, d(\mathbf{I},\bar{x})_\#\mu(x_1,x_2) \right)^{1/p} \\ &\quad + \left(\int_{X\times X} d^p(x_1,x_2)\, d(\mathbf{I},\bar{x})_\#\nu(x_1,x_2) \right)^{1/p} \\ &= d_{L^p(X,\mu;X)}(x,\bar{x}) + d_{L^p(X,\nu;X)}(x,\bar{x}) \end{aligned}$$

and the last quantity is finite since $\mu, \nu \in \mathscr{P}_p(X)$. Hence $W_p \neq +\infty$. Endowing $\mathscr{P}_p(X)$ with the W_p distance, we have that a set in $\Xi \subset \mathscr{P}_p(X)$ is bounded if and only if, for some $\bar{x} \in X$, there exists $R > 0$ such that $\int_X d^p(\bar{x}, x)\, d\mu(x) \le R$ for any $\mu \in \Xi$, that is, if and only if the p-moment is uniformly bounded in Ξ.

A first estimate on the Wasserstein distance is given by the following

Proposition 2.2.5 *Let $\nu \in \mathscr{P}(Y)$, let $r, s : Y \to X$ be ν-measurable maps. Then*

$$W_p(r_\#\nu, s_\#\nu) \le d_{L^p(Y,\nu;X)}(r,s). \tag{2.2.4}$$

Proof. Notice that $(r,s)_\#\nu \in \Gamma(r_\#\nu, s_\#\nu)$, hence

$$W_p^p(r_\#\nu, s_\#\nu) \le \int_{X\times X} d^p(x,y)\, d((r,s)_\#\nu)(x,y) = \int_Y d^p(r(\zeta) - s(\zeta))\, d\nu(\zeta),$$

which is the desired inequality. □

Next we consider a convergence result of optimal transport plans.

Lemma 2.2.6 *Let $(\mu_n), (\nu_n) \subset \mathscr{P}_p(X)$ be weakly converging to μ and ν respectively. Let $\gamma_n \in \Gamma_0(\mu_n, \nu_n)$, $n \in \mathbb{N}$, be such that $\int_{X\times X} d^p(x_1,x_2)\, d\gamma_n$ is bounded with respect to n. Then the sequence (γ_n) is weakly relatively compact in $\mathscr{P}(X \times X)$ and if γ is one of its weak limit points, then $\gamma \in \Gamma_0(\mu,\nu)$. Moreover*

$$W_p(\mu,\nu) \le \liminf_{n\to\infty} W_p(\mu_n,\nu_n). \tag{2.2.5}$$

Proof. The relative compactness of the sequence follows from the tightness Proposition 1.2.2. Since $(x_1,x_2) \mapsto d(x_1,x_2)$ is bounded below, we apply (1.1.2), with $f = d$, on $X \times X$ to obtain (2.2.5). By optimality and continuity of the cost, $\operatorname{supp} \gamma_n$ is $d^p(\cdot,\cdot)$-monotone for every n, and so is $\operatorname{supp} \gamma$ applying Proposition 1.1.6. □

The following result shows the relation between the convergence induced by W_p and the notions of convergence already introduced in the space of probability measures.

Lemma 2.2.7 (**W_p metrizes the convergence with moments in $\mathscr{P}_p(X)$**) *Let (μ_n) be a sequence in $\mathscr{P}_p(X)$. Let $\mu \in \mathscr{P}_p(X)$. Then $W_p(\mu_n, \mu) \to 0$ if and only if $\mu_n \to \mu$ in $\mathscr{P}_p(X)$.*

Proof. We limit ourselves to the complete case, which allows to apply a simpler argument. For the general case, see [4, Lemma 7.1.5]. Let $W_p(\mu_n, \mu) \to 0$. Assume that X is complete. Then, so is $\mathscr{P}_p(X)$ endowed with the p-Wasserstein distance (see [72, Lemma 6.12]) hence $\mu_n \rightharpoonup \mu$. Let $\gamma_n \in \Gamma_0(\mu_n, \mu)$. Consider the elementary inequality $d^p(x,z) \le C_\varepsilon d^p(x,y) + (1+\varepsilon) d^p(z,y)$, valid for any $x, y, z \in X$, where C_ε is a suitable constant. Integrating we get

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

Since $W_p(\mu_n, \mu) \to 0$ as $n \to \infty$, passing to the limit we have

$$\limsup_{n\to\infty} \int_X d^p(x,z)\, d\mu_n(x) \le (1+\varepsilon) \int_X d^p(x,z)\, d\mu(x)$$

This shows that (μ_n) has uniformly integrable p-moments, then we conclude by Proposition 1.1.4.

Conversely, let $\mu_n \to \mu$ in $\mathscr{P}_p(X)$. Since weakly converging sequences are tight, letting $\gamma_n \in \Gamma_0(\mu_n, \mu)$, the sequence (γ_n) is tight too, thanks to Proposition 1.2.1. We extract, without relabeling it, a subsequence weakly converging to $\gamma \in \mathscr{P}(X \times X)$. By Lemma 2.2.6, $\gamma \in \Gamma_0(\mu, \mu)$, so that $\gamma = (\mathbf{I}, \mathbf{I})_{\#}\mu$. Now let $z \in X$ and $R > 0$. Notice that $\{(x,y) : d(x,y) \ge R\} \subset \{(x,y) : d(x,z) \ge d(x,y)/2 \ge R/2\} \cup \{(x,y) : d(y,z) \ge d(x,y)/2 \ge R/2\}$. We have

$$\begin{aligned} W_p^p(\mu_n, \mu) &= \int_{X\times X} d^p(x,y)\, d\gamma_n(x,y) = \int_{X\times X} (d(x,y) \wedge R)^p \; d\gamma_n(x,y) \\ &\quad + \int_{X\times X} (d^p(x,y) - R^p)^+ \; d\gamma_n(x,y) \\ &\le \int_{X\times X} (d(x,y) \wedge R)^p \; d\gamma_n(x,y) + 2^p \int_{d(x,z)\ge R/2} d^p(x,z)\, d\mu_n(x) \\ &\quad + 2^p \int_{d(y,z)\ge R/2} d^p(z,y)\, d\mu(y) \end{aligned}$$

Since $d(x,y) \wedge R$ is continuous, bounded, null on the diagonal $x = y$, passing to the limit in n we see that the first term in the right hand side goes to zero. Hence

$$\limsup_{n\to\infty} W_p^p(\mu_n, \mu) \le 2^p \limsup_{n\to\infty} \left(\int_{d(x,z)\ge R/2} d^p(x,z)\, d\mu_n(x) + \int_{d(y,z)\ge R/2} d^p(z,y)\, d\mu(y) \right).$$

Taking the limit as $R \to +\infty$, since (μ_n) has uniformly integrable p-moments, we get the thesis. □

Now we are able to show a convergence result, taking advantage of Lemma 2.2.7.

Lemma 2.2.8 *Let $\mu_n \rightharpoonup \mu$ in $\mathscr{P}(X_\varpi)$ and let $\int_X |x|^2\, d\mu_n(x) \to \int_X |x|^2\, d\mu(x)$. Then μ_n weakly converge to μ also in $\mathscr{P}(X)$.*

Proof. By hypothesis, the sequence (μ_n) has uniformly integrable 2-moments (with respect to the distance induced by $|\cdot|$), and since $\|\cdot\|_\varpi$, the norm in X_ϖ, satisfies $\|\cdot\|_\varpi \le |\cdot|$, we have

$$\lim_{r\to\infty} \sup_{n\in\mathbb{N}} \int_{X\setminus B_r(0)} \|x\|_\varpi^2\, d\mu_n(x) \le \lim_{r\to\infty} \sup_{n\in\mathbb{N}} \int_{X\setminus B_r(0)} |x|^2\, d\mu_n(x) = 0,$$

so that (μ_n) has uniformly integrable 2-moments also with respect to the distance induced by $\|\cdot\|_\varpi$. By applying Proposition 1.1.4 in X_ϖ, we find

$$\lim_{n\to\infty} \int_X \|x\|_\varpi^2\, d\mu_n(x) = \int_X \|x\|_\varpi^2\, d\mu(x).$$

We conclude that μ_n converges to μ with 2-moments in $\mathscr{P}(X_\varpi)$ (so, this is the convergence with moments, but with a different underlying structure on X). Now let $\gamma_n \in \Gamma(\mu_n, \mu)$, and let each γ_n be optimal in $\mathscr{P}_2(X_\varpi)$ (that is, minimizing $\int_{X\times X} \|x-y\|_\varpi^2\, d\tilde{\gamma}_n$ among all $\tilde{\gamma}_n \in \Gamma(\mu, \mu_n)$), so that Lemma 2.2.7 (which works also in non complete sapces) yields $\gamma_n \to (\mathbf{I}, \mathbf{I})_\#\mu$ in $\mathscr{P}(X_\varpi \times X_\varpi)$. The convergence is in fact in $\mathscr{P}(X \times X_\varpi)$, since the first marginal of γ_n is μ, so it does not depend on n. For the same reason, $(\pi^1_\# \gamma_n)$ has uniformly integrable 2-moments, and thanks to the hypothesis of convergence of moments of μ_n, we see that the assumptions of Lemma 1.5.2 are all satisfied. Then there holds

$$\lim_{n\to\infty} \int_{X\times X} \langle x_1, x_2\rangle\, d\gamma_n(x_1, x_2) = \int_{X\times X} \langle x_1, x_2\rangle\, d((\mathbf{I}, \mathbf{I})_\#\mu)(x_1, x_2) = \int_X |x|^2\, d\mu(x). \quad (2.2.6)$$

Now consider the elementary equality

$$|x_1 - x_2|^2 = |x_2|^2 - |x_1|^2 - 2\langle x_1, x_2 - x_1\rangle \quad \forall x_1, x_2 \in X.$$

Let us integrate it with respect to γ_n. We obtain

$$\begin{aligned} \int_{X\times X} |x_1 - x_2|^2\, d\gamma_n(x_1, x_2) &= \int_{X\times X} \big(|x_2|^2 - |x_1|^2 - 2\langle x_1, x_2 - x_1\rangle\big)\, d\gamma_n(x_1, x_2) \\ &= \int_X |x|^2\, d\mu_n(x) - \int_X |x|^2\, d\mu(x) \\ &\quad + 2\int_{X\times X} |x_1|^2\, d\gamma_n(x_1, x_2) - 2\int_{X\times X} \langle x_1, x_2\rangle\, d\gamma_n(x_1, x_2). \end{aligned}$$

The left hand side is the squared Wasserstein distance, and the second one, making use of (2.2.6), tends to 0. We get $W_2(\mu_n, \mu) \to 0$ as $n \to \infty$. Now, Lemma 2.2.7 allows to conclude. □

An important consequence of this result is the next strengthening of Lemma 1.5.4.

Lemma 2.2.9 *Let $\mu_n \to \mu$ in $\mathscr{P}_2(X)$. If $\boldsymbol{\rho}_n$ strongly converge to $\boldsymbol{\rho} \in L^2(X, \mu; X)$ in the sense of* Definition 1.5.3, *then*

$$(\mathbf{I}, \boldsymbol{\rho}_n)_\# \mu_n \rightharpoonup (\mathbf{I}, \boldsymbol{\rho})_\# \mu \tag{2.2.7}$$

as $n \to \infty$ and

$$\lim_{n\to\infty} \int_X f(x, \boldsymbol{\rho}_n(x)) d\mu_n(x) = \int_X f(x, \boldsymbol{\rho}(x)) d\mu \tag{2.2.8}$$

for every continuous function $f : X \times X \to \mathbb{R}$ with at most 2-growth, that is

$$|f(x,y)| \le A + B(\|x\|^2 + \|y\|^2) \quad \forall (x,y) \in X \times X \tag{2.2.9}$$

for some A, $B \in \mathbb{R}$. More generally, (2.2.8) *holds also if strong L^2 convergence is replaced by*

$$\lim_{n\to\infty} \int_X g(\boldsymbol{\rho}_n(x)) d\mu_n = \int_X g(\boldsymbol{\rho}(x)) d\mu \tag{2.2.10}$$

for some strictly convex function $g : X \to \mathbb{R}$ with at least 2-growth at infinity.

Proof. In the hypotheses of Lemma 1.5.4, point *iii)*, we proved that $\boldsymbol{\rho}_{n\#}\mu_n \rightharpoonup \boldsymbol{\rho}_\#\mu$ in $\mathscr{P}(X_\varpi)$ and $(\mathbf{I}, \boldsymbol{\rho}_n)_\#\mu_n \rightharpoonup (\mathbf{I}, \boldsymbol{\rho})_\#\mu$ in $\mathscr{P}(X_\varpi \times X_\varpi)$. But the second marginals of these plans are strongly converging, that is, they satisfy (1.5.6) with $p = 2$, hence we can invoke Lemma 2.2.8 and infer that actually the weak convergence is in $\mathscr{P}(X_\varpi \times X)$. So, here we conclude that $(\mathbf{I}, \boldsymbol{\rho}_n)_\#\mu_n \rightharpoonup (\mathbf{I}, \boldsymbol{\rho})_\#\mu$ in $\mathscr{P}_2(X \times X)$, and (2.2.8) readily follows. The last statement is proved as point *iii)* of Lemma 1.5.4. In fact, it is enough to substitute the p-power therein with a generic strictly convex function with the required growth. □

Wasserstein geodesics

Let X be a separable Hilbert space. We define constant speed geodesics in $\mathscr{P}_p(X)$ with respect to W_p (or Wasserstein geodesics).

Definition 2.2.10 (Constant speed geodesic) *We say that a curve $t \in [0,1] \mapsto \mu_t \in \mathscr{P}_p(X)$ is a constant speed geodesic connecting μ_0 and μ_1 if for any s, t such that $0 \le s \le t \le 1$ there holds*

$$W_p(\mu_s, \mu_t) = (t-s) W_p(\mu_0, \mu_1). \tag{2.2.11}$$

In the next theorem we are showing that a constant speed geodesic between μ_0 and μ_1 can be characterized as a convex interpolation of the marginals of some optimal transport plan $\gamma \in \Gamma(\mu_0, \mu_1)$.

In the sequel we make use of the following notation, for $t \in [0,1]$:

$$\pi_t^{1,\,1\to 2} := (1-t)\pi^{1,1} + t\pi^{1,2}, \qquad \pi_t^{1\to 2,\,2} := (1-t)\pi^{1,2} + t\pi^{2,2}.$$

Theorem 2.2.11 (Geodesical interpolation) *Let $p > 1$. The following properties hold:*

i) Let $t \in [0,1] \mapsto \mu_t \in \mathscr{P}_p(X)$ be a constant speed geodesic. If $0 < t < 1$, then there exists a unique optimal transport plan β_t^1 between μ_t and μ_1, and such plan is induced by a map $\mathbf{t}_t^1$. On the other hand, there exists a unique optimal transport plan β_0^t between μ_0 and μ_t, and its inverse is induced by a map $\mathbf{t}_t^0$. Letting $\beta = \beta_t^1 \circ \beta_0^t$, we have

$$\beta_0^t = (\pi_t^{1,1\to 2})_\# \beta, \qquad \beta_t^1 = (\pi_t^{1\to 2,2})_\# \beta. \tag{2.2.12}$$

ii) A curve $t \in [0,1] \mapsto \mu_t \in \mathscr{P}_p(X)$ is a constant speed geodesic connecting μ_1 and μ_2 if and only if it can be written as

$$t \mapsto \left((1-t)\pi^1 + t\pi^2\right)_\# \gamma \tag{2.2.13}$$

for some $\gamma \in \Gamma_0(\mu_1, \mu_2)$.

Proof. *i)* Fix $t \in (0,1)$ and let $\gamma \in \Gamma_0(\mu_0, \mu_t), \gamma' \in \Gamma_0(\mu_t, \mu_1)$. Since we are going to deal with plans composition, we let X_1, X_2, X_3 be copies of the space X and we consider $\mu_0 \in \mathscr{P}(X_1)$, $\mu_t \in \mathscr{P}(X_2)$, $\mu_1 \in \mathscr{P}(X_3)$. Let us disintegrate γ and γ' with respect to the common marginal μ_t, so that we obtain two families of Borel measure valued maps $x_2 \in X_1 \mapsto \gamma_{x_2} \in \mathscr{P}(X_1)$ and $x_2 \in X_2 \mapsto \gamma'_{x_2} \in \mathscr{P}(X_3)$, and we have

$$\gamma = \int_{X_2} \gamma_{x_2} \, d\mu_t(x_2), \quad \gamma' = \int_{X_2} \gamma'_{x_2} \, d\mu_t(x_2).$$

As in Remark 2.2.2, we define the measure $\boldsymbol{\gamma} := \int_{X_2} \gamma_{x_2} \times \gamma'_{x_2} \, d\mu_t(x_2)$, which satisfies

$$\pi^{1,2}_\# \boldsymbol{\gamma} = \gamma, \quad \pi^{2,3}_\# \boldsymbol{\gamma} = \gamma'.$$

Hence, $\pi^{1,3}_\# \boldsymbol{\gamma}$ is the composition $\gamma' \circ \gamma \in \Gamma(\mu_0, \mu_1)$. Let us show that $\gamma' \circ \gamma$ is optimal. For, consider that $t \mapsto \mu_t$ is a geodesic, so by (2.2.11) we have

$$W_p(\mu_0, \mu_1) = W_p(\mu_0, \mu_t) + W_p(\mu_t, \mu_1),$$

and since γ and γ' are optimal,

$$\begin{aligned} W_p(\mu_0, \mu_1) &= \left(\int_{X_1 \times X_2} |x_1 - x_2|^p \, d\gamma\right)^{1/p} + \left(\int_{X_2 \times X_3} |x_2 - x_3|^p \, d\gamma'\right)^{1/p} \\ &= \left(\int_{X_1 \times X_2} |x_1 - x_2|^p \, d\pi^{1,2}_\# \boldsymbol{\gamma}\right)^{1/p} + \left(\int_{X_2 \times X_3} |x_2 - x_3|^p \, d\pi^{2,3}_\# \boldsymbol{\gamma}\right)^{1/p} \\ &= \left(\int_{X_1 \times X_2 \times X_3} |x_1 - x_2|^p \, d\boldsymbol{\gamma}\right)^{1/p} + \left(\int_{X_1 \times X_2 \times X_3} |x_2 - x_3|^p \, d\boldsymbol{\gamma}\right)^{1/p}. \end{aligned}$$

Hence, by the triangle inequality of the L^p norm, and recalling the rule for composing plans, we find

$$\begin{aligned} W_p(\mu_0,\mu_1) &\geq \left(\int_{X_1\times X_2\times X_3} |x_1-x_3|^p\, d\boldsymbol{\gamma}\right)^{1/p} \\ &= \left(\int_{X_2}\left(\int_{X_1\times X_3} |x_1-x_3|^p\, d(\gamma_{x_2}\times\gamma'_{x_2})(x_1,x_3)\right)\, d\mu_t(x_2)\right)^{1/p} \\ &= \left(\int_{X_1\times X_3} |x_1-x_3|^p\, d\gamma'\circ\gamma(x_1,x_3)\right)^{1/p}. \end{aligned}$$

This shows that $\gamma'\circ\gamma$ is indeed optimal between μ_0 and μ_1. Moreover, the triangle inequality above is an equality, and since the L^p norm is strictly convex, this implies that the three vectors involved are parallel, that is, there exists $\alpha > 0$ such that $x_2 - x_1 = \alpha(x_3 - x_1)$ for $\boldsymbol{\gamma}$-a.e. $(x_1, x_2, x_3) \in X_1 \times X_2 \times X_3$. Since $W_p(\mu_0, \mu_t) = tW_p(\mu_0, \mu_1)$, we see that $\alpha = t$. We integrate such $\boldsymbol{\gamma}$-a.e. equality on X_1 w.r.t. γ_{x_2} and, letting $z(x_2)$ be its barycenter (see Definition 1.4.8), we find $x_2 - z(x_2) = t(x_3 - z(x_2))$ for γ'-a.e. $(x_2, x_3) \in X_2 \times X_3$. This means that x_3 is uniquely determined by x_2, for μ_t-a.e. $x_2 \in X_2$, and so γ' is concentrated on a graph and induced by a map. γ' is unique because once γ is fixed, z is fixed too. Then, γ' corresponds to the unique optimal plan β^t_1 given in the statement of the theorem.

The uniqueness of the plan in $\Gamma_0(\mu_0, \mu_t)$, and the fact that its inverse is induced by a map, is obtained with the same argument, starting from μ_1 and going to μ_0. The relations (2.2.12) are readily seen to follow.

ii) Let the curve $t \in [0, 1] \mapsto \mu_t \in \mathscr{P}_p(X)$ enjoy the representation (2.2.13). By (2.2.4) we have, for $0 \leq s \leq t \leq 1$,

$$W_p(\mu_t, \mu_s) \leq (t - s)W_p(\mu_0, \mu_1).$$

By the triangle inequality, it is readily seen that this has to be an equality for any s, t.

On the other hand, let $r \in [0, 1] \mapsto \mu_r \in \mathscr{P}_p(X)$ be a constant speed geodesic. Fix $t \in (0, 1)$ and let β^t_0, β^1_t be the unique optimal plans of point *i)* and let $\beta = \beta^1_t \circ \beta^t_0$. Now $s \in [0, 1] \mapsto \mu_{ts}$ is a constant speed geodesic between μ_0 and μ_t, and since β^t_0 is the unique element of $\Gamma_0(\mu_0, \mu_t)$, by the converse implication we see that $\mu_{ts} = ((1 - s)\pi^1 + s\pi^2)_\# \beta^t_0$. But

$$((1-s)\pi^1 + s\pi^2)_\#\beta^t_0 = \left(((1-s)\pi^1 + s\pi^2)\circ\pi^{1,1\to 2}_t\right)_\# \beta = ((1-ts)\pi^1 + ts\pi^2)_\#\beta.$$

We have the desired representation of μ_r for r varying from 0 to t, and we get to 1 inverting μ_0 with μ_1 and repeating the same argument. □

2.3 The continuity equation in $(\mathscr{P}_2(X), W_2)$

In this section we discuss a first instance through which a relation with PDEs arises. Absolutely continuous curves in $\mathscr{P}_p(X)$ are in fact strictly connected with the continuity equation, as we shall see.

We begin giving the basic definitions.

Definition 2.3.1 *Let $t \in [a,b] \mapsto \mu_t \in \mathscr{P}_p(X)$ be a curve. We say that it is absolutely continuous, and write $\mu_t \in AC([a,b];\mathscr{P}_p(X))$, if*

$$W_p(\mu_{t_1},\mu_{t_2}) \leq \int_{t_1}^{t_2} U(t)\,dt \quad \forall a < t_1 \leq t_2 < b, \tag{2.3.1}$$

for some $U \in L^1((a,b))$. If moreover there is some $U \in L^p((a,b))$, $p>1$, such that (2.3.1) *holds, then we say that $\mu_t \in AC^p([a,b];\mathscr{P}_p(X))$.*

Definition 2.3.2 *Let $t \in [a,b] \mapsto \mu_t \in \mathscr{P}_p(X)$ be a curve. Its metric derivative at the point t is defined as*

$$|\mu'|_t := \lim_{s\to t} \frac{W_p(\mu_s,\mu_t)}{|s-t|}, \tag{2.3.2}$$

if the limit exists.

The metric derivative plays the role of the modulus of the standard derivative of the fundamental theorem of calculus. Indeed, it can be shown that, if $\mu_t \in AC^p([a,b];\mathscr{P}_p(X))$, then the limit in (2.3.2) exists $\mathcal{L}^1$-a.e. in (a,b) and metric the derivative $|\mu'|_t$ so defined belongs to $L^p((a,b))$ and satisfies (2.3.1). Moreover, if $U \in L^p((a,b))$ satisfies (2.3.1) then $|\mu'|_t \leq U(t)$ $\mathcal{L}^1$-a.e. in (a,b) (see for instance [10]).

We are going to show that each absolutely continuous curve $\mu_t \in AC([a,b];\mathscr{P}_2(X))$ satisfies the continuity equation

$$\partial_t\mu_t + \operatorname{div}(\mathbf{v}_t\mu_t) = 0, \tag{2.3.3}$$

for some vector field $\mathbf{v}_t(x) : (a,b)\times X \to X$ such that the map $x \in X \mapsto \mathbf{v}_t(x)$ belongs to $L^2(X,\mu_t;X)$ for $\mathcal{L}^1$-a.e. $t \in (a,b)$. Here, equation (2.3.3) has to be intended in the sense of distributions, in duality with smooth cylindrical functions (introduced in Definition 1.4.3), that is

$$\int_a^b \int_X \left(\partial_t\varphi(x,t) + \langle v_t(x), \nabla_x\varphi(x,t)\rangle\right) d\mu_t(x)\,dt = 0 \quad \forall \varphi \in \mathrm{Cyl}((a,b)\times X). \tag{2.3.4}$$

In particular, among all the vector fields $\mathbf{v}_t$ which, coupled to the curve μ_t, satisfy (2.3.3), the minimal $L^2(X,\mu_t;X)$ norm is given by $|\mu'|_t$. Moreover, $\mathbf{v}_t$ has minimal norm if

$$\mathbf{v}_t \in \overline{\{\nabla\varphi : \varphi \in \mathrm{Cyl}(X)\}}^{L^2(X,\mu_t;X)}.$$

This set can also be seen as the tangent space to $\mathscr{P}_2(X)$ at μ (see Remark 2.3.4 below). For a complete discussion about the geometrical properties of $\mathscr{P}_2(X)$, see [40].

Now we state and prove the result, letting for simplicity $(a,b) = (0,1)$.

Theorem 2.3.3 (Continuity equation) *Let $\mu_t \in AC([0,1];\mathscr{P}_2(X))$ be an absolutely continuous curve with metric derivative $|\mu'|_t$. Then there exists a Borel vector field $v : (x,t) \mapsto \mathbf{v}_t(x)$ such that, for $\mathcal{L}^1$-a.e. $t \in (0,1)$, $\mathbf{v}_t \in L^2(X,\mu_t;X)$ and*

$$\|\mathbf{v}_t\|_{L^2(X,\mu_t;X)} \le |\mu'|_t, \tag{2.3.5}$$

and such that (2.3.3) *holds.*

Proof. Let $\varphi \in \mathrm{Cyl}(X)$, and $\gamma_{s,t} \in \Gamma_0(\mu_s,\mu_t)$. Let moreover

$$H(x,y) := \begin{cases} |\nabla\varphi(x)| & \text{if } x = y, \\ \dfrac{|\varphi(x)-\varphi(y)|}{|x-y|} & \text{if } x \neq y. \end{cases}$$

There holds

$$\begin{aligned} \left|\int_X \varphi(x)\,d\mu_{t+h}(x) - \int_X \varphi(x)\,d\mu_t(x)\right| &\le \int_{X\times X} |x-y| H(x,y)\,d\gamma_{t+h,\,t}(x,y) \\ &\le W_2(\mu_{t+h},\mu_t)\left(\int_{X\times X} H^2(x,y)\,d\gamma_{t+h,\,t}(x,y)\right)^{1/2}. \end{aligned} \tag{2.3.6}$$

Notice that, as $h \downarrow 0$, $\gamma_{t+h,\,t}$ converges weakly in $\mathscr{P}(X\times X)$ to $(\mathbf{I},\mathbf{I})_\#\mu_t$. In fact, $t \mapsto \mu_t$ is continuous in $\mathscr{P}_2(X)$, so that the marginals of $\gamma_{t+h,\,t}$ are both weakly converging to μ_t and we can invoke Lemma 2.2.7. Suppose now that t is a point where the metric derivative of μ_t exists. Therefore, passing to the limsup in (2.3.6) we find

$$\limsup_{h\downarrow 0} \frac{\int_X \varphi(x)\,d(\mu_{t+h}-\mu_t)(x)}{|h|} \le |\mu'|_t \left(\int_X H^2(x,x)\,d\mu_t(x)\right)^{1/2} = |\mu'|_t \|\nabla\varphi\|_{L^2(X,\mu_t;X)}. \tag{2.3.7}$$

Now define the space time measure $\mu = \int_X \mu_t\,dt \in \mathscr{P}(X\times[0,1])$. So μ_t is the family of measures which disintegrates μ with respect to its second marginal $\mathcal{L}^1\llcorner(0,1)$. Let $(x,t) \mapsto \zeta(x,t) \in \mathrm{Cyl}(X\times[0,1])$. Disintegrating μ, (2.3.7) yields (with Fatou's lemma and thanks to the boundedness of ζ)

$$\begin{aligned} \left|\int_{X\times[0,1]} \partial_t\zeta(x,t)\,d\mu(x,t)\right| &= \left|\int_{X\times[0,1]} \lim_{h\downarrow 0} \frac{\zeta(x,t+h)-\zeta(x,t)}{h}\,d\mu(x,t)\right| \\ &= \left|\int_0^1 \limsup_{h\downarrow 0} \frac{1}{h}\left(\int_X \zeta(x,t)\,d\mu_{t+h}(x) - \int_X \zeta(x,t)\,d\mu_t(x)\right)dt\right| \\ &\le \int_0^1 |\mu'|_t \left(\int_X |\nabla_x\zeta(x,t)|^2\,d\mu_t(x)\right)^{1/2} dt \\ &\le \left(\int_0^1 |\mu'|_t^2\,dt\right)^{1/2} \left(\int_{X\times[0,1]} |\nabla_x\zeta(x,t)|^2\,d\mu_t(x)\right)^{1/2}. \end{aligned}$$

Let $\Omega := \{\nabla_x \zeta : \zeta \in \mathrm{Cyl}(X \times [0,1])\}$, let $\mathfrak{L} : \Omega \to \mathbb{R}$ be the linear functional defined by

$$\mathfrak{L}(\nabla_x \zeta) := -\int_{X\times[0,1]} \partial_t \zeta(x,t)\, d\mu(x,t). \tag{2.3.8}$$

We just proved that the operator $\mathfrak{L}$, defined on Ω, is bounded with respect to the $L^2(X \times [0,1], \mu; X)$ norm, so that we can extend it uniquely to a bounded functional on $\overline{\Omega}$, the closure being in $L^2(X \times [0,1], \mu; X)$. Hence, there is a unique $v = v(x,t) \in \overline{\Omega}$ satisfying

$$\mathfrak{L}(\nabla_x \zeta) = \int_{X\times[0,1]} \langle v(x,t), \nabla_x \zeta(x,t)\rangle\, d\mu(x,t) \quad \forall \zeta \in \mathrm{Cyl}(X \times [0,1]). \tag{2.3.9}$$

The element v is just the velocity for the continuity equation. In fact, let $\mathbf{v}_t := v(x,t)$, and combining (2.3.8) and (2.3.9) we get (2.3.4). In order to show the metric derivative inequality (2.3.5), we introduce a cutoff function $\eta \in C_c^\infty(I)$, where I is an interval contained in $[0,1]$, such that $0 \le \eta \le 1$. We also introduce a sequence $(\nabla_x \zeta_n) \subset \Omega$ converging to v in $L^2(X \times [0,1], \mu; X)$. We have

$$\begin{aligned}
\int_{X\times[0,1]} \eta(t)|v(x,t)|^2\, d\mu(x,t) &= \lim_{n\to\infty} \int_{X\times[0,1]} \eta(t)\langle v(x,t), \nabla_x \zeta_n(x,t)\rangle\, d\mu(x,t)\\
&= \lim_{n\to\infty} \mathfrak{L}(\nabla_x(\eta\zeta_n))\\
&\le \left(\int_I |\mu'|_t^2\, dt\right)^{1/2} \lim_{n\to\infty} \left(\int_{X\times I} |\nabla_x \zeta_n(x,t)|^2\, d\mu(x,t)\right)^{1/2}\\
&= \left(\int_I |\mu'|_t^2\, dt\right)^{1/2} \left(\int_{X\times I} |v(x,t)|^2\, d\mu(x,t)\right)^{1/2}.
\end{aligned}$$

Now we let η approximate χ_I. We get

$$\int_I \int_X |\mathbf{v}_t(x)|^2\, d\mu_t(x)\, dt \le \int_I |\mu'|_t^2\, dt$$

and (2.3.5) follows. □

Remark 2.3.4 (Tangent vector) Let μ_t be an absolutely continuous curve and $\mathbf{v}_t$ be a vector field such that the continuity equation with the couple $(\mu_t, \mathbf{v}_t)$ holds. Then $\mathbf{v}_t$ satisfies (2.3.5) if and only if

$$\mathbf{v}_t \in \overline{\{\nabla\varphi : \varphi \in \mathrm{Cyl}(X)\}}^{L^2(X,\mu_t;X)} \quad \text{for } \mathcal{L}^1\text{-a.e. } t \in (0,1). \tag{2.3.10}$$

In fact, if this condition holds together with the continuity equation, then $\mathbf{v}_t$ is unique, as the continuity equation is linear in $\mathbf{v}_t$ and the L^2 norm is strictly convex. On the other hand, the vector field satisfying (2.3.4) and (2.3.5), explicitly exhibited in the proof of Theorem 2.3.3, does satisfy (2.3.10) by construction. We express condition (2.3.10) by saying that

$$\mathbf{v}_t \in \mathrm{Tan}_{\mu_t}\mathscr{P}_2(X).$$

For a discussion about the tangent bundle of $\mathscr{P}_2(X)$, we refer to [4, Chapter 8] and to [40, 41, 59].

We also have the following useful formula for computing the derivative of W_2 along an absolutely continuous curve (for the proof we refer to [4, Theorem 8.4.7]).

Lemma 2.3.5 *Let $t \in [0,1] \mapsto \mu_t \in \mathscr{P}_2(X)$ be an absolutely continuous curve, with tangent vector $\mathbf{v}_t$. Let $\nu \in \mathscr{P}_2(X)$. Then there holds*

$$\frac{d}{dt} W_2^2(\mu_t, \nu) = 2 \int_{X \times X} \langle x_1 - x_2, \mathbf{v}_t(x_1) \rangle \, d\gamma_t \quad \forall \gamma_t \in \Gamma_0(\mu_t, \nu), \tag{2.3.11}$$

for $\mathcal{L}^1$-a.e. $t \in (0,1)$.

2.4 Transport network problems

In this section we briefly describe a family of minimization problems whose formulation is strictly related to optimal transport. Here, let X be an Euclidean space, and let $\Omega \subset X$ be an open bounded set.

The optimal transport problem, in its basic form (2.1.2) or the generalized one (2.1.3), requires to find the best way to transport a given amount of material to a given site. The source and the target are represented by two finite measures μ, ν with the same total mass.

An urban planning problem can be formulated in a very similar way (see for instance [21] or [23]). We are given again two measures μ, ν, the first representing the population density in a region Ω, the second indicating the density of workplaces to be reached. We consider a Borel set $\Sigma \subset \Omega$ of finite $\mathcal{H}^1$ measure, representing an urban transport network, which has to be constructed minimizing the cost for transporting μ to ν. Here we consider the cost

$$I_\Sigma = \int_{\Omega \times \Omega} d_\Sigma(x,y) \, d\gamma(x,y), \tag{2.4.1}$$

with $\gamma \in \Gamma(\mu, \nu)$. The distance d_Σ accounts for the cost for citizens to move from x to y. In order to allow them optimize their cost by choosing to move on the network Σ, or without it (say on foot), we can define

$$d_\Sigma(x,y) = f(\mathcal{H}^1(\Lambda \backslash \Sigma)) + g(\mathcal{H}^1(\Lambda \cap \Sigma)).$$

Citizens will move along curves Λ connecting x to y, f and g are functions accounting for the different cost of the part of Λ on Σ and out of Σ. Let us assume that citizens prefer not to move by own means, so that the cost for travelling by the network is negligible. Let also consider the cost as proportional to the covered distances. In this case, the simplest choice is

$$d_\Sigma(x,y) = \min\{d(x,y), \operatorname{dist}(x, \Sigma) + \operatorname{dist}(y, \Sigma)\},$$

where d stands for the geodesic distance in Ω and $\operatorname{dist}(x, \Sigma) := \inf\{d(x,z) : z \in \Sigma\}$. Hence, given a network Σ, the citizen will try to optimize his cost by choosing a transport plan γ which is optimal with respect to d_Σ.

From the point of view of the planner, the aim will be to minimize the cost I_Σ among the closed connected sets Σ. A constraint on the cost of construction of the network has to be considered. For instance we can suppose that it is proportional to its length so that the problem will be

$$\min\{I_\Sigma + \lambda\mathcal{H}^1(\Sigma) : \Sigma \text{ closed connected subset of } \Omega\},$$

where $\lambda > 0$.

Next we consider a particular instance of this problem. Let μ, ν be probabilities for simplicity. Suppose that the goal is simply to transport the source to Σ. In this case ν is not fixed, but optimally chosen among probability measures with support in Σ. In this case the distance $d_\Sigma(x, y)$ reduces to $\operatorname{dist}(x, \Sigma)$ and

$$I_\Sigma = \int_\Omega \operatorname{dist}(x, \Sigma)\, d\mu(x).$$

We are led to the so-called *irrigation problem*, or *average distance problem*:

$$\min\left\{\int_\Omega \operatorname{dist}(x, \Sigma)\, d\mu(x) + \lambda\mathcal{H}^1(\Sigma) \;:\; \Sigma \text{ closed connected subset of } \Omega\right\}.$$

We notice that the problem can be equivalently formulated in terms of Wasserstein distance as

$$\min\{W_1(\mu, \nu) + \lambda\mathcal{H}^1(\Sigma) : \operatorname{supp}\nu \in \Sigma, \Sigma \text{ closed connected subset of } \Omega\}.$$

Chapter 3

Minimizing movements and gradient flows in $(\mathscr{P}_2(X), W_2)$

3.1 Convexity and subdifferential in $(\mathscr{P}_2(X), W_2)$

In this section, let X be a separable Hilbert space, and let $\phi : \mathscr{P}_2(X) \to (-\infty, +\infty]$ be a functional. The domain of ϕ is defined as

$$D(\phi) := \{\mu \in \mathscr{P}_2(X) : \phi(\mu) < +\infty\} \tag{3.1.1}$$

and we say that ϕ is proper if $D(\phi) \neq \emptyset$. The basic assumptions on ϕ, which will be understood to hold during the rest of this chapter, are now presented.

Assumption 3.1.1 *ϕ is proper, l.s.c. (with respect to the convergence with moments) and bounded from below.*

We will need different notions of convexity.

Definition 3.1.2 (λ-convexity along a curve) *Let $\lambda \in \mathbb{R}$. The functional ϕ is said to be λ-convex along the curve $t \in [0,1] \mapsto \mu_t \in \mathscr{P}_2(X)$ if*

$$\phi(\mu_t) \leq (1-t)\phi(\mu_0) + t\phi(\mu_1) - \frac{1}{2}\lambda t(1-t)W_2^2(\mu_0, \mu_1). \tag{3.1.2}$$

Definition 3.1.3 (Geodesical convexity) *Let $\lambda \in \mathbb{R}$. We say that ϕ is λ-geodesically convex, or convex along (Wasserstein) geodesics, if, for any $\mu^1, \mu^2 \in \mathscr{P}_2(X)$, there exists $\gamma \in \Gamma_0(\mu^1, \mu^2)$ such that the inequality*

$$\phi\left(((1-t)\pi^1 + t\pi^2)_\# \gamma\right) \leq (1-t)\phi(\mu^1) + t\phi(\mu^2) - \frac{1}{2}\lambda t(1-t)W_2^2(\mu^1, \mu^2) \tag{3.1.3}$$

holds for any $t \in [0,1]$.
In the case $\lambda = 0$, we simply say that ϕ is geodesically convex (or convex along geodesics).

Definition 3.1.4 (Strong geodesical convexity) *We say that $\phi : \mathscr{P}_2(X) \to [0, +\infty]$ is strongly geodesically convex (or simply strongly convex) if it is convex along geodesics and for any $\mu, \nu, \sigma \in D(\phi)$ there exists a continuous curve $\mu_t : [0,1] \to \mathscr{P}_2(X)$, with $\mu_0 = \mu$ and $\mu_1 = \nu$, such that*

$$\begin{cases} W_2^2(\mu_t, \sigma) \le (1-t)W_2^2(\mu, \sigma) + tW_2^2(\nu, \sigma) - t(1-t)W_2^2(\mu, \nu) \\ \phi(\mu_t) \le (1-t)\phi(\mu) + t\phi(\nu) \end{cases} \qquad \forall t \in [0,1]. \tag{3.1.4}$$

Sometimes it will be important for ϕ to be convex also along some curve in $\mathscr{P}_2(X)$ on which the Wasserstein distance itself behaves like a convex functional. This motivates the introduction of the strong convexity. In fact, it can be shown with a counterexample that for any $\lambda \in \mathbb{R}$ the application $\mu^2 \mapsto W_2(\mu^1, \mu^2)$, at least if the dimension of X is not 1, is not λ-geodesically convex (for more details see [4, §7.3])

Now we introduce the notions needed to develop a differential calculus in $(\mathscr{P}_2(X), W_2)$.

Definition 3.1.5 (Metric slope) *The metric slope of ϕ at the point μ is given by*

$$|\partial\phi|(\mu) = \limsup_{\nu \to \mu \ \text{in} \mathscr{P}_2(X)} \frac{(\phi(\nu) - \phi(\mu))^+}{W_2(\nu, \mu)}. \tag{3.1.5}$$

The application $\mu \mapsto |\partial\phi|(\mu)$ is l.s.c. if ϕ is λ-geodesically convex, as we shall see later.

In the sequel the following assumption on ϕ will be made.

Assumption 3.1.6 *For each $\mu \in D(|\partial\phi|)$ and for each $\nu \in \mathscr{P}_2(X)$, there exists a unique optimal transport plan in $\Gamma(\mu, \nu)$, and such plan is induced by a map, which will be denoted by $\mathbf{t}_\mu^\nu$.*

For example, a condition guaranteeing these facts is $D(|\partial\phi|) \subset \mathscr{P}_2^r(X)$, as a consequence of Theorem 2.1.7.

Definition 3.1.7 (Wasserstein subdifferential) *Let $\mu \in D(|\partial\phi|)$. The Wasserstein subdifferential $\partial\phi(\mu)$ of ϕ at μ is the set of vectors $\boldsymbol{\xi} \in L^2(X, \mu; X)$ such that*

$$\phi(\nu) - \phi(\mu) \ge \int_X \langle \boldsymbol{\xi}, \mathbf{t}_\mu^\nu(x) - x \rangle \, d\mu(x) + o(W_2(\mu, \nu)). \tag{3.1.6}$$

In the case of λ-convex functionals along Wasserstein geodesics, we have the following

Proposition 3.1.8 *Let ϕ be λ-geodesically convex. Then $\boldsymbol{\xi} \in \partial\phi(\mu)$ if and only if*

$$\phi(\nu) - \phi(\mu) \ge \int_X \langle \boldsymbol{\xi}(x), \mathbf{t}_\mu^\nu(x) - x \rangle \, d\mu(x) + \frac{1}{2}\lambda W_2^2(\mu, \nu) \quad \forall \nu \in \mathscr{P}_2(X). \tag{3.1.7}$$

Proof. If (3.1.7) holds for some $\lambda \in \mathbb{R}$, then clearly (3.1.6) holds too. On the other hand, suppose that $\boldsymbol{\xi} \in \partial\phi(\mu)$ and fix $\nu \in \mathscr{P}_2(X)$. Let μ_t be the Wasserstein geodesic connecting μ to ν, given by $\mu_t := \big((1-t)\mathbf{I} + t\mathbf{t}_\mu^\nu\big)_\# \mu$ (and this is an optimal transport, see Theorem 2.2.11). By definition of geodesic, $W_2(\mu, \mu_t) = tW_2(\mu,\nu)$, and since $\boldsymbol{\xi} \in \partial\phi(\mu)$,

$$\liminf_{t\to 0} \frac{1}{t}\left(\phi(\mu_t) - \phi(\mu)\right) \geq \liminf_{t\to 0} \frac{1}{t}\int_X \langle \boldsymbol{\xi}(x), \mathbf{t}_\mu^{\mu_t}(x) - x\rangle\, d\mu(x) = \int_X \langle \boldsymbol{\xi}(x), \mathbf{t}_\mu^{\nu}(x) - x\rangle\, d\mu(x).$$

On the other hand, in this case the λ-geodesical convexity relation (3.1.3) reads

$$\frac{\phi(\mu_t) - \phi(\mu)}{t} \leq \phi(\nu) - \phi(\mu) - \frac{1}{2}\lambda(1-t)W_2^2(\mu,\nu).$$

Combining the latter two inequalities, we get the thesis. □

3.2 The minimizing movements scheme

The minimizing movements scheme was proposed by De Giorgi (see for instance [32]), in order to deal with steepest descent curves for functionals defined in generic metric spaces. The idea of recursively descending along the gradient, with a Euler implicit discretization scheme, is standard in the Euclidean framework, whereas it can be thought as a way to define optimal curves in more general contexts. Although the setting could be much wider, to keep a more coherent exposition we are going to present it in probability spaces.

Let X be a separable Hilbert space, let $\phi : \mathscr{P}_2(X) \to \mathbb{R}$ be a l.s.c. and bounded from below functional. Consider a uniform partition of the time interval $[0,T]$, let τ be the corresponding step, and fix $\mu_0 \in \mathscr{P}_2(X)$. Consider the perturbed functional $\Phi_\tau(\cdot, \mu^0) : \mathscr{P}_2(X) \to \mathbb{R}$, defined as

$$\Phi_\tau(\nu, \mu^0) := \phi(\nu) + \frac{1}{2\tau}W_2^2(\nu, \mu^0). \tag{3.2.1}$$

The recursive scheme to be exploited is the following: given $\mu_\tau^0, \ldots, \mu_\tau^k$, find μ_τ^{k+1} solving

$$\min_{\nu \in \mathscr{P}_2(X)} \Phi_\tau(\nu, \mu_\tau^k). \tag{3.2.2}$$

Here, μ_τ^0 is an approximation of μ^0, that is:

$$\mu_\tau^0 \to \mu^0 \text{ in } \mathscr{P}_2(X) \quad \text{and} \quad \phi(\mu_\tau^0) \to \phi(\mu^0) \quad \text{as } \tau \downarrow 0. \tag{3.2.3}$$

If minimizers exist, this procedure will produce a sequence $\{\mu_\tau^k\}_{k\in N} \subset \mathscr{P}_2(X)$ (of course not a unique one in general). Once such a sequence is constructed, we define a corresponding piecewise constant interpolating curve $\bar{\mu}_t$ in $\mathscr{P}_2(X)$ as follows:

$$\bar{\mu}_\tau(t) := \begin{cases} \mu_\tau^0 & \text{if } t = 0, \\ \mu_\tau^k & \text{if } t \in ((k-1)\tau, k\tau],\ k > 0. \end{cases} \tag{3.2.4}$$

Equivalently,

$$\bar{\mu}_\tau(t) = \mu_\tau^{\lceil t/\tau \rceil} \quad \forall t > 0. \tag{3.2.5}$$

We are ready for the definition of minimizing movement.

Definition 3.2.1 (Minimizing movement) *Let $\tau > 0$ and suppose that a sequence $(\mu_\tau^k) \subset \mathscr{P}_2(X)$ solving (3.2.2) exists. Let $\bar{\mu}_t$ be the corresponding piecewise constant discrete solution (defined by (3.2.4)). We say that a curve $t \in [0,T] \mapsto \mu_t \in \mathscr{P}_2(X)$ is a (generalized) minimizing movement for functional Φ (defined in (3.2.1)), starting from μ^0, if there exists a sequence $\tau_n \downarrow 0$ such that*

$$\lim_{n\to\infty} \phi(\mu_{\tau_n}^0) = \phi(\mu^0) \quad \textit{and} \quad \bar{\mu}_{\tau_n}(t) \to \mu_t \textit{ in } \mathscr{P}_2(X),\ \forall t \in [0,T]. \tag{3.2.6}$$

Within this construction, it is natural to give the following

Definition 3.2.2 (Discrete velocity) *We define the discrete velocity of the minimizing movements scheme as*

$$\mathbf{V}_\tau^k := \frac{\mathbf{I} - \mathbf{t}_\tau^k}{\tau}, \tag{3.2.7}$$

where $\mathbf{t}_\tau^k$ is the optimal transport map between μ_τ^k and μ_τ^{k-1}. Also define the corresponding piecewise constant interpolation

$$\overline{\mathbf{V}}_\tau(t) := \mathbf{V}_\tau^k \quad \textit{for } t \in ((k-1)\tau, k\tau]. \tag{3.2.8}$$

With the next lemma we introduce some basic estimates. Consider the function

$$\phi_\tau(\mu) := \inf_{\nu \in \mathscr{P}_2(X)} \Phi_\tau(\nu, \mu). \tag{3.2.9}$$

It is called the Moreau-Yosida approximation of ϕ. Being ϕ a proper functional, $\phi_\tau(\mu)$ is finite for any $\mu \in \mathscr{P}_2(X)$. First of all, notice that, letting μ_τ be a minimizer of $\Phi_\tau(\cdot, \mu)$, we have

$$\phi_\tau(\mu) = \phi(\mu_\tau) + \frac{1}{2\tau} W_2^2(\mu_\tau, \mu) \le \phi(\nu) + \frac{1}{2\tau} W_2^2(\nu, \mu), \quad \forall \nu \in \mathscr{P}_2(X),$$

and choosing $\nu = \mu$ we find $\phi_\tau(\mu) \le \phi(\mu)$ and $\phi(\mu_\tau) \le \phi(\mu)$ for any $\tau > 0$, and also

$$\frac{W_2^2(\mu_\tau, \mu)}{2\tau} \le \phi(\mu) - \phi(\mu_\tau), \tag{3.2.10}$$

yielding $\mu_\tau \to \mu$ as $\tau \to 0$.

Lemma 3.2.3 *There hold:*

i) $(\tau, \mu) \mapsto \phi_\tau(\mu)$ *is continuous in* $(0, T) \times \mathscr{P}_2(X)$.

ii) If μ_τ is a minimizer for $\Phi_\tau(\cdot,\mu)$, there holds

$$|\partial\phi|(\mu_\tau) \le W_2^2(\mu_\tau,\mu)/\tau, \tag{3.2.11}$$

so that $D(|\partial\phi|)$ is W_2-dense in $D(\phi)$.

iii) If μ_τ is a minimizer for $\Phi_\tau(\cdot,\mu)$, then there exists a vanishing sequence (τ_n) such that $\phi(\mu_{\tau_n}) \to \phi(\mu)$ and

$$|\partial\phi|^2(\mu) = \lim_{n\to\infty} \frac{W_2^2(\mu_{\tau_n},\mu)}{\tau_n^2} = \lim_{n\to\infty} \frac{\phi(\mu)-\phi(\mu_{\tau_n})}{\tau_n} \ge \liminf_{\tau\downarrow 0} |\partial\phi|^2(\mu_\tau). \tag{3.2.12}$$

Proof. For proving *i)*, let $\tau_n \to \tau$ and $\mu_n \to \mu$ (with τ_n distant from zero). Clearly we can find a sequence $(\nu_n) \subset D(\phi)$ such that $(\Phi_{\tau_n}(\nu_n,\mu_n) - \phi_{\tau_n}(\mu_n)) \to 0$ as $n\to\infty$. Since $\phi_{\tau_n}(\mu_n)$ is bounded, so is $\Phi_{\tau_n}(\nu_n,\mu_n)$ and

$$\limsup_{n\to\infty} \Phi_{\tau_n}(\nu_n,\mu_n) = \limsup_{n\to\infty} \phi_{\tau_n}(\mu_n) \le \Phi_\tau(\nu,\mu)$$

for any $\nu \in \mathscr{P}_2(X)$. Taking the infimum w.r.t. ν, by definition of the Moreau-Yosida function we find $\limsup_n \phi_{\tau_n}(\mu_n) \le \phi_\tau(\mu)$. In order to prove the corresponding liminf inequality, notice that $W_2(\mu,\nu_n)$ is bounded. In fact, we have seen that $\Phi_{\tau_n}(\nu_n,\mu_n)$ is bounded, and since ϕ is a bounded from below functional and τ_n is distant from zero, we argue that $W_2(\mu,\nu_n)$ is bounded too. Hence

$$\begin{aligned}\liminf_{n\to\infty} \phi_{\tau_n}(\mu_n) = \liminf_{n\to\infty} \Phi_{\tau_n}(\nu_n,\mu_n) &\ge \liminf_{n\to\infty}\left(\frac{(W_2(\mu,\nu_n)-W_2(\mu_n,\mu))^2}{2\tau_n} + \phi(\nu_n)\right)\\ &\ge \liminf_{n\to\infty}\left(\frac{1}{2\tau_n}W_2^2(\nu_n,\mu) - \frac{1}{\tau_n}W_2(\nu_n,\mu)W_2(\mu_n,\mu) + \phi(\nu_n)\right) \ge \phi_\tau(\mu).\end{aligned}$$

Concerning *ii)*, suppose that μ_τ is a minimizer of $\Phi_\tau(\cdot,\mu)$. It is easily seen that, for any $\nu \in D(\phi)$,

$$\phi(\mu_\tau) - \phi(\nu) \le \frac{1}{2\tau}W_2^2(\nu,\mu) - \frac{1}{2\tau}W_2^2(\mu,\mu_\tau) \le \frac{1}{2\tau}W_2(\nu,\mu_\tau)\left(W_2(\nu,\mu)+W_2(\mu,\mu_\tau)\right),$$

and dividing by $W_2(\nu,\mu_\tau)$ we find

$$\limsup_{\nu\to\mu_\tau} \frac{(\phi(\mu_\tau)-\phi(\nu))^+}{W_2(\nu,\mu_\tau)} \le \limsup_{\nu\to\mu_\tau} \frac{1}{2\tau}\left(W_2(\nu,\mu)+W_2(\mu,\mu_\tau)\right) = \frac{W_2(\nu,\mu)}{\tau}. \tag{3.2.13}$$

This proves *ii)*. In order to show *iii)*, we start letting $|\partial\phi|(\mu) > 0$ and $0 < a < |\partial\phi|(\mu)$. Consider, for $b > 0$, the identity

$$\frac{1}{2}b^2 = \sup_{z\in\mathbb{R}}\left(bz - \frac{1}{2}z^2\right) = \sup_{z>b}\left(bz - \frac{1}{2}z^2\right).$$

Since $a < |\partial\phi|(\mu)$, if $W_2(\mu,\nu)$ is small enough we have $a < \frac{(\phi(\mu)-\phi(\nu))^+}{W_2(\mu,\nu)}$, hence making use of the previous identity we have

$$\begin{aligned}\frac{1}{2}|\partial\phi|^2(\mu) &= \limsup_{\nu\to\mu \text{ in}\mathscr{P}_2(X)} \frac{1}{2}\left(\frac{(\phi(\mu)-\phi(\nu))^+}{W_2(\mu,\nu)}\right)^2 = \limsup_{\nu\to\mu \text{ in}\mathscr{P}_2(X)} \sup_{z>a}\left(\frac{z(\phi(\mu)-\phi(\nu))^+}{W_2(\mu,\nu)} - \frac{z^2}{2}\right)\\ &= \limsup_{\nu\to\mu \text{ in}\mathscr{P}_2(X)} \sup_{0<\tau<W_2(\mu,\nu)/a}\left(\frac{(\phi(\mu)-\phi(\nu))^+}{W_2(\mu,\nu)}\frac{W_2(\mu,\nu)}{\tau} - \frac{W_2^2(\mu,\nu)}{2\tau^2}\right)\\ &\geq \limsup_{\tau\to 0}\left(\frac{\phi(\mu)-\phi(\mu_\tau)}{\tau} - \frac{W_2^2(\mu,\mu_\tau)}{2\tau^2}\right) = \limsup_{\tau\to 0}\frac{\phi(\mu)-\phi_\tau(\mu)}{\tau},\end{aligned} \tag{3.2.14}$$

where we passed to computing the limit on the particular sequence μ_τ, $\tau\to 0$, and we made use of the definition of $\phi_\tau(\mu)$. The latter also implies

$$\sup_{\nu\neq\mu}\left((\phi(\mu)-\phi(\nu))^+ - \frac{W_2^2(\mu,\nu)}{2\tau}\right) = \phi(\mu)-\phi_\tau(\mu),$$

so that for small ε

$$\begin{aligned}\frac{1}{2}|\partial\phi|^2(\mu) &= \limsup_{\nu\to\mu \text{ in}\mathscr{P}_2(X)} \sup_{0<\tau<W_2(\mu,\nu)/a}\left(\frac{(\phi(\mu)-\phi(\nu))^+}{W_2(\mu,\nu)}\frac{W_2(\mu,\nu)}{\tau} - \frac{W_2^2(\mu,\nu)}{2\tau^2}\right)\\ &\leq \sup_{\tau<\varepsilon}\sup_{\nu\neq\mu}\left(\frac{(\phi(\mu)-\phi(\nu))^+}{W_2(\mu,\nu)}\frac{W_2(\mu,\nu)}{\tau} - \frac{W_2^2(\mu,\nu)}{2\tau^2}\right) = \sup_{\tau<\varepsilon}\frac{\phi(\mu)-\phi_\tau(\mu)}{\tau}.\end{aligned}$$

Taking the limit as $\varepsilon\to 0$, we get $\frac{1}{2}|\partial\phi|^2(\mu) \leq \limsup_{\tau\to 0}(\phi(\mu)-\phi(\mu_\tau))/\tau$, and by virtue of (3.2.14) we see that such inequality is in fact an equality. Therefore, passing to a suitable vanishing subsequence τ_n, and then taking advantage once more of the definition of $\phi_\tau(\mu)$ and of metric slope, we get

$$\begin{aligned}\frac{1}{2}|\partial\phi|^2(\mu) &= \lim_{n\to\infty}\frac{\phi(\mu)-\phi_{\tau_n}(\mu)}{\tau_n} = \lim_{n\to\infty}\left(\frac{\phi(\mu)-\phi(\mu_{\tau_n})}{\tau_n} - \frac{W_2^2(\mu,\mu_{\tau_n})}{2\tau_n^2}\right)\\ &\leq \limsup_{n\to\infty}\left(|\partial\phi|(\mu)\frac{W_2(\nu,\mu_{\tau_n})}{\tau_n} - \frac{W_2^2(\mu,\mu_{\tau_n})}{2\tau_n^2}\right).\end{aligned} \tag{3.2.15}$$

In view of (3.2.15), from the equalities

$$\begin{aligned}\liminf_{n\to\infty}\left||\partial\phi|(\mu) - \frac{W_2(\mu,\mu_{\tau_n})}{\tau_n}\right|^2 &= \liminf_{n\to\infty}\left(|\partial\phi|^2(\mu) - 2|\partial\phi|(\mu)\frac{W_2(\mu,\mu_{\tau_n})}{\tau_n} + \frac{W_2^2(\mu,\mu_{\tau_n})}{\tau_n^2}\right)\\ &= |\partial\phi|^2(\mu) - \limsup_{n\to\infty}\left(2|\partial\phi|(\mu)\frac{W_2(\mu,\mu_{\tau_n})}{\tau_n} - \frac{W_2^2(\mu,\mu_{\tau_n})}{\tau_n^2}\right)\end{aligned}$$

we learn that

$$\liminf_{n\to\infty}\left||\partial\phi|(\mu) - \frac{W_2(\mu,\mu_{\tau_n})}{\tau_n}\right| = 0. \tag{3.2.16}$$

As a consequence, making use of (3.2.11) we have

$$\liminf_{n\to\infty}\left(|\partial\phi|(\mu_{\tau_n})-|\partial\phi|(\mu)\right)\le\liminf_{n\to\infty}\left(\frac{W_2(\mu,\mu_{\tau_n})}{\tau_n}-|\partial\phi|(\mu)\right)\le 0,$$

which yields the inequality $\liminf_{\tau\to 0}|\partial\phi|^2(\mu_\tau)\le|\partial\phi|^2(\mu)$ of (3.2.12). From (3.2.16) we also infer that there exists a further subsequence (that we don't relabel) such that $W_2^2(\mu,\mu_{\tau_n})/\tau_n^2\to|\partial\phi|^2(\mu)$, proving the first equality in (3.2.12). Then, from (3.2.15) we readily see that also $(\phi(\mu)-\phi(\mu_{\tau_n}))/\tau_n\to|\partial\phi|^2(\mu)$, and (3.2.12) is proven. The case $|\partial\phi|(\mu)=0$ is much simpler, we omit the details. □

3.3 The gradient flow equation

After the minimizing movements, we introduce a second issue in the study of trajectories in $\mathscr{P}_2(X)$ which are thought as steepest descent curves of a functional ϕ on $\mathscr{P}_2(X)$. We give the following

Definition 3.3.1 (Gradient flow) *Let $\mu_t\in AC([0,T];\mathscr{P}_2(X))$. After the results of* §2.3, *we know that μ_t satisfies the continuity equation*

$$\partial_t\mu_t+\operatorname{div}(\mathbf{v}_t\mu_t)=0 \tag{3.3.1}$$

in correspondence of the optimal $L^2(X,\mu_t;X)$ vector field, (optimal in the sense that $\mathbf{v}_t\in Tan_{\mu_t}\mathscr{P}_2(X)$, for $\mathscr{L}^1$-a.e. $t\in(0,T)$).

We say that μ_t solves the gradient flow equation if

$$\mathbf{v}_t\in-\partial\phi(\mu_t)\quad \textit{for } \mathscr{L}^1\textit{-a.e. } t\in(0,T). \tag{3.3.2}$$

In the λ-geodesically convex case, using Proposition 3.1.8 it is readily seen that if μ_t satisfies the gradient flow equation, for $\mathbf{v}_t\in\mathrm{Tan}_{\mu_t}\mathscr{P}_2(X)$ there hold

$$\begin{cases}\partial_t\mu_t+\operatorname{div}(\mathbf{v}_t\mu_t)=0,\\ \displaystyle\int_X\langle\mathbf{v}_t,\mathbf{t}_{\mu_t}^\nu-\mathbf{I}\rangle\,d\mu_t\le\phi(\nu)-\phi(\mu_t)-\frac12\lambda W_2^2(\mu_t,\nu)\quad\forall\nu\in D(\phi).\end{cases} \tag{3.3.3}$$

Here the family of inequalities holds for $\mathscr{L}^1$-a.e. $t\in(0,T)$, and thanks to (2.3.11) it is equivalent to

$$\frac12\frac{d}{dt}W_2^2(\mu_t,\nu)\le\phi(\nu)-\phi(\mu_t)-\frac12\lambda W_2^2(\mu_t\nu)\quad\forall\nu\in D(\phi). \tag{3.3.4}$$

The following result is strictly related to the formulation just given, and generalizations to the case $p\neq 2$ are not known.

Theorem 3.3.2 *Let $\mu^0 \in \mathscr{P}_2(X)$ and let μ_t satisfy $\mu_t \to \mu^0$ in $\mathscr{P}_2(X)$, as $t \downarrow 0$. Then μ_t is a gradient flow for functional ϕ if and only if satisfies the system of* Evolution Variational Inequalities

$$\frac{1}{2}\frac{d}{dt}W_2^2(\mu_t,\nu) + \frac{1}{2}\lambda W_2^2(\mu_t,\nu) \le \phi(\nu) - \phi(\mu_t) \quad \textit{for } \mathcal{L}^1\textit{-a.e. } t \in (0,T), \tag{3.3.5}$$

for any $\nu \in D(\phi)$.

Proof. First suppose that μ_t is a gradient flow, with $\mu_t \to \mu^0$ in $\mathscr{P}_2(X)$ as $t \downarrow 0$. This implications has already been shown through (3.3.3) and (3.3.4). On the other hand, suppose that $\mu_t \in AC([0,T];\mathscr{P}_2(X))$ satisfies (3.3.5). Then it is clear that, if $\Xi \in D(\phi)$ is a countabel set, there exists a $\mathcal{L}^1$-negligible set N_Ξ such that the inequality in (3.3.3) holds for any $(t,\nu) \in ([0,T]\backslash N_\Xi) \times \Xi$, making use of (2.3.11). Choose a countable set Ξ which is dense in $D(\phi)$ (recall that $(\mathscr{P}_2(X), W_2)$ is separable). The density can be asked also with respect to the stronger $W_2(\mu^1,\mu^2) + |\phi(\mu^1) - \phi(\mu^2)|$ metric. Now fix $t \in [0,T]\backslash N_\Xi$ and suppose that $(\nu_n) \subset \Xi$ is a sequence converging in $\mathscr{P}_2(X)$ to $\nu \in D(\phi)$. For any n, the inequality in (3.3.3) holds:

$$\int_X \langle \mathbf{v}_t, \mathbf{t}_{\mu_t}^{\nu_n} - \mathbf{I}\rangle \, d\mu_t \le \phi(\nu_n) - \phi(\mu_t) - \frac{1}{2}\lambda W_2^2(\mu_t,\nu_n)$$

The second member goes to $\phi(\nu) - \phi(\mu_t) - \frac{1}{2}\lambda W_2^2(\mu_t,\nu)$ by density. The first one, letting $\gamma_t^n = (\mathbf{I}, \mathbf{t}_{\mu_t}^{\nu_n})_\# \mu_t$, is equal to

$$\int_{X\times X} \langle \mathbf{v}_t(x_1), x_2 - x_1\rangle \, d\gamma_t^n,$$

and by virtue of Lemma 2.2.6 $\gamma_t^n \rightharpoonup \gamma_t$, where $\gamma_t \in \Gamma_0(\mu_t,\nu)$ is an optimal plan (the unique one, thanks to the regularity of μ_t). This convergence is also with moments since $W_2(\nu_n,\mu_t) \to W_2(\nu,\mu_t)$ as $n \to \infty$. Hence the integral passes to the limit if $\mathbf{v}_t$ is continuous and bounded, and as well if it is not, by another density argument, using the fact that the first marginal of γ_t^n does not depend on n. □

We conclude this section with a uniqueness result, which also depends strictly on the choice $p = 2$ of the moment exponent and on the convexity assumptions. Later in Chapter 5, with the particular choice of ϕ therein, we will be able to show a uniqueness result for the gradient flow in a non λ-geodesically convex context, with a proof that makes use of some of the ideas we are going to give in the next theorem. We need a preliminary

Lemma 3.3.3 *Let $f(s,t) : [0,T] \times [0,T] \to \mathbb{R}$ satisfy*

$$|f(s_1,t_1) - f(s_2,t_1)| \le |g(s_1) - g(s_2)| \quad \textit{and} \quad |f(s_1,t_1) - f(s_1,t_2)| \le |g(t_1) - g(t_2)|$$

for any $s_1, s_2, t_1, t_2 \in [0,T]$, where $g : [0,T] \to \mathbb{R}$ is some absolutely continuous function. Then $t \mapsto f(t,t)$ is absolutely continuous on $[0,T]$ and, for $\mathcal{L}^1$-a.e. $t \in (0,T)$, there holds

$$\frac{d}{dt}f(t,t) \le \limsup_{h\downarrow 0} \frac{f(t,t) - f(t-h,t)}{h} + \limsup_{h\downarrow 0} \frac{f(t,t+h) - f(t,t)}{h}.$$

Proof. The hypotheses yield $|f(s,s) - f(t,t)| \le 2|g(s) - g(t)|$, therefore we get absolute continuity. Now fix $h_* > 0$ and $\varphi \in C_c^\infty((0,T))$ such that $\{x \in \mathbb{R} : x = t \pm h, t \in \operatorname{supp}\varphi, 0 < h < h_*\} \subset (0,T)$. Hence, the changes of variables $t \mapsto t - h$ and then $t \mapsto t + h$ entail

$$\begin{aligned}
&\int_0^T -\frac{\varphi(t+h) - \varphi(t)}{h} f(t,t)\,dt = \int_0^T \frac{f(t,t) - f(t-h,t-h)}{h}\varphi(t)\,dt \\
=\ &\int_0^T \frac{f(t,t) - f(t-h,t)}{h}\varphi(t)\,dt + \int_0^T \frac{f(t-h,t) - f(t-h,t-h)}{h}\varphi(t)\,dt \\
=\ &\int_0^T \frac{f(t,t) - f(t-h,t)}{h}\varphi(t)\,dt + \int_0^T \frac{f(t,t+h) - f(t,t)}{h}\varphi(t+h)\,dt.
\end{aligned} \tag{3.3.6}$$

By Fatou's lemma we also have

$$\begin{aligned}
-\int_0^T f(t,t)\varphi'(t)\,dt &= \int_0^T -\lim_{h\downarrow 0}\frac{\varphi(t+h) - \varphi(t)}{h} f(t,t)\,dt \\
&\le \limsup_{h\downarrow 0}\int_0^T -\frac{\varphi(t+h) - \varphi(t)}{h} f(t,t)\,dt.
\end{aligned} \tag{3.3.7}$$

Since g is absolutely continuous, as $h \downarrow 0$ there hold

$$\begin{aligned}
\frac{1}{h}(f(t,t) - f(t-h,t)) &\le \frac{1}{h}|g(t) - g(t-h)| \to |g'(t)| \quad \text{in } L^1(0,T), \\
\frac{1}{h}(f(t,t+h) - f(t,t)) &\le \frac{1}{h}|g(t+h) - g(t)| \to |g'(t)| \quad \text{in } L^1(0,T).
\end{aligned} \tag{3.3.8}$$

Then, we combine (3.3.7) and (3.3.6), and we apply a generalized form of the Fatou's lemma to the left hand sides in (3.3.8) (indeed, the limsup inequality holds for a sequence of functions ψ_n if $\psi_n \le \zeta_n$ and $\zeta_n \to \zeta$ in L^1). We get

$$\begin{aligned}
&-\int_0^T f(t,t)\varphi'(t)\,dt \\
&\le \limsup_{h\downarrow 0}\int_0^T \frac{f(t,t) - f(t-h,t)}{h}\varphi(t)\,dt + \limsup_{h\downarrow 0}\int_0^T \frac{f(t-h,t) - f(t-h,t-h)}{h}\varphi(t)\,dt, \\
&\le \int_0^T \limsup_{h\downarrow 0}\frac{f(t,t) - f(t-h,t)}{h}\varphi(t)\,dt + \int_0^T \limsup_{h\downarrow 0}\frac{f(t-h,t) - f(t-h,t-h)}{h}\varphi(t)\,dt.
\end{aligned}$$

The thesis follows integrating by parts and by the arbitrariness of φ. □

Theorem 3.3.4 (Uniqueness of the gradient flow) *Let ϕ be a λ-geodesically convex functional and let μ_t^1, μ_t^2 be two gradient flows (in the sense of* (3.3.3)*) satisfying respectively the initial conditions $\mu_t^1 \to \mu^1$ and $\mu_t^2 \to \mu^2$ in $\mathscr{P}_2(X)$, as $t \downarrow 0$. Then*

$$W_2(\mu_t^1, \mu_t^2) \le e^{-\lambda t} W_2(\mu^1, \mu^2), \quad t \ge 0.$$

Hence, $\mu_t^1 = \mu_t^2$ if $\mu^1 = \mu^2$.

Proof. Notice that, by (3.3.5),

$$\begin{aligned}\limsup_{h\downarrow 0}\frac{W_2^2(\mu_t^1,\mu_t^2)-W_2^2(\mu_{t-h}^1,\mu_t^2)}{h} &= \frac{d}{ds}W_2^2(\mu_s^1,\mu_t^2)\Big|_{s=t}\\ &\le -\lambda W_2^2(\mu_t^1,\mu_t^2)+2\phi(\mu_t^2)-2\phi(\mu_t^1)\end{aligned}$$

and

$$\begin{aligned}\limsup_{h\downarrow 0}\frac{W_2^2(\mu_t^1,\mu_{t+h}^2)-W_2^2(\mu_t^1,\mu_t^2)}{h} &= \frac{d}{ds}W_2^2(\mu_t^1,\mu_s^2)\Big|_{s=t}\\ &\le -\lambda W_2^2(\mu_t^1,\mu_t^2)+2\phi(\mu_t^1)-2\phi(\mu_t^2).\end{aligned}$$

Therefore, applying Lemma 3.3.3, with $W_2^2(\mu_s^1,\mu_t^2)$ in the role of $f(s,t)$, we get

$$\frac{d}{dt}W_2^2(\mu_t^1,\mu_t^2)\le -2\lambda W_2^2(\mu_t^1,\mu_t^2),$$

so that Gronwall's lemma immediately gives the thesis. □

3.4 Existence of solutions

Next we discuss about the convergence of the minimizing movements scheme and the solution of the gradient flow equation. As usual, here $\phi:\mathscr{P}_2(X)\to(-\infty,+\infty]$ will be a proper, l.s.c. and bounded from below functional satisfying Assumption 3.1.6.

The first result we need is about the subdifferential of the perturbed functional characterizing the minimizing movements scheme, defined by (3.2.1).

Lemma 3.4.1 *Let μ_τ be a minimizer of $\Phi_\tau(\cdot,\mu^0)$, the functional defined by* (3.2.1)*, for a given $\mu^0\in\mathscr{P}_2(X)$. Then $\mu_\tau\in D(|\partial\phi|)$ and*

$$\frac{\mathbf{t}_{\mu_\tau}^{\mu^0}-\mathbf{I}}{\tau}\in\partial\phi(\mu_\tau). \tag{3.4.1}$$

Proof. We have seen that $\mu_\tau\in D(|\partial\phi|)$ in Lemma 3.2.3. Since μ_τ minimizes $\Phi_\tau(\cdot,\mu^0)$ we readily get, for any $\nu\in\mathscr{P}_2(X)$

$$\frac{1}{2\tau}\left(W_2^2(\mu_\tau,\mu^0)-W_2^2(\nu,\mu^0)\right)\le\phi(\nu)-\phi(\mu_\tau). \tag{3.4.2}$$

If $\mathbf{t}$ is a transport map between μ_τ and ν, applying (3.4.2) and the elementary identity $\frac{1}{2}|a|^2-\frac{1}{2}|b|^2=\langle a,a-b\rangle-\frac{1}{2}|a-b|^2$, $a,b\in X$, we get

$$\begin{aligned}\phi(\nu)-\phi(\mu_\tau) &\ge \frac{1}{2\tau}\int_X|\mathbf{t}_{\mu_t}^{\mu^0}(x)-x|^2\,d\mu_\tau(x)-\frac{1}{2\tau}\int_X|\mathbf{t}_{\mu_t}^{\mu^0}(x)-\mathbf{t}(x)|^2\,d\mu_\tau(x)\\ &=\frac{1}{\tau}\int_X\langle\mathbf{t}_{\mu_\tau}^{\mu^0}(x)-x,\mathbf{t}(x)-x\rangle\,d\mu_\tau(x)-\frac{1}{2\tau}\int_X|\mathbf{t}(x)-x|^2\,d\mu_\tau(x)\\ &=\frac{1}{\tau}\int_X\langle\mathbf{t}_{\mu_\tau}^{\mu^0}(x)-x,\mathbf{t}(x)-x\rangle\,d\mu_\tau(x)-\frac{1}{2\tau}\|\mathbf{t}-\mathbf{I}\|^2_{L^2(X,\mu_\tau;X)}.\end{aligned}$$

The thesis follows, and notice that the inequality obtained is even stronger then (3.1.6), since here $\mathbf{t}$ is not necessarily optimal. □

Corollary 3.4.2 *The discrete velocity of the minimizing movements scheme, introduced in* Definition 3.2.2*, satisfies*

$$\overline{\mathbf{V}}_\tau^k \in -\partial\phi(\mu_\tau^k). \tag{3.4.3}$$

In this section and in the subsequent one we are going to construct a curve $t \in [0,T] \mapsto \mu_t \in \mathscr{P}_2(X)$ which is both a minimizing movement and a gradient flow for functional ϕ. So, it will be clear that the two concepts are strictly related, in some sense the two approaches lead to the same result. The minimizing movements scheme (which itself is defined in general metric spaces) has indeed to be understood as the time discretization of equation (3.3.2). Such a point of view, in the context of optimal transportation, was introduced in the seminal paper [43], as we will remark better in Chapter 4.

We will present two main results, the first one needing the hypothesis of compactness in $\mathscr{P}_2(X)$ for the sublevels of ϕ, but without convexity assumptions, the second one concerning λ-geodesically convex functionals. In the first case, we are able to construct a minimizing movement and to show that it satisfies the gradient flow equation in a relaxed form that we are going to specify. For, we need a preliminary

Definition 3.4.3 (Limiting subdifferential) *Let $\mu \in D(\phi)$, let $\boldsymbol{\xi} \in L^2(X,\mu;X)$. We say that $\boldsymbol{\xi}$ belongs to the limiting subdifferential $\partial_\ell\phi(\mu)$ of ϕ at μ if there exist two sequences $\mu_k \subset D(|\partial\phi|)$, $(\boldsymbol{\xi}_k) \subset L^2(X,\mu_k;X)$, such that $\boldsymbol{\xi}_k \in \partial\phi(\mu_k)$, $\mu_k \rightharpoonup \mu$, $\boldsymbol{\xi}_k \to \boldsymbol{\xi}$ weakly as in* Definition 1.5.3 *and*

$$\sup_{k\in\mathbb{N}} \int_X |\boldsymbol{\xi}_k(x)|^2 \, d\mu_k(x) < +\infty.$$

Theorem 3.4.4 (Limit curve) *Let ϕ be such that if $\Xi \subset \mathscr{P}_2(X)$ is bounded with respect to W_2 and contained in a sublevel of ϕ, then Ξ is compact in $(\mathscr{P}_2(X), W_2)$. Let $\mu^0 \in D(\phi)$. Let (τ_n) be a vanishing sequence of time steps, and let $\bar{\mu}_{\tau_n}$ be a discrete solution, defined by (3.2.4), in correspondence of the initial datum $\mu^0_{\tau_n}$. Then there exists a subsequence (still denoted by τ_n) of time steps such that $\bar{\mu}_{\tau_n}$ converges, identifying a limiting curve $t \in [0,T] \mapsto \mu_t \in \mathscr{P}_2(X)$, that is*

$$\bar{\mu}_{\tau_n} \to \mu_t \text{ in } \mathscr{P}_2(X), \quad \forall t \in [0,T]. \tag{3.4.4}$$

The limit curve μ_t is also absolutely continuous. Moreover, there exists a vector map $(x,t) \mapsto \mathbf{v}_t(x)$ belonging to $L^2(X \times [0,T], \mu; X)$, where $\mu = \frac{1}{T}\int_0^T \mu_t \, dt$, such that

$$\overline{\mathbf{V}}_{\tau_n} \to \mathbf{v} \quad \text{weakly in } L^2 \text{ as in Definition 1.5.3.}$$

Such convergence holds with respect to the product measures $\bar{\mu}_{\tau_n} := \frac{1}{T}\int_0^T \bar{\mu}_{\tau_n}(t) \, dt$, which weakly converge in $\mathscr{P}(X \times [0,T])$ to μ. Finally, letting $\mathbf{v}_t(x) = \mathbf{v}(x,t)$, the couple $(\mu_t, \mathbf{v}_t)$ satisfies the continuity equation (2.3.3) *and the relaxed gradient flow inclusion*

$$\mathbf{v}_t \in -\overline{\mathrm{Conv}}\, \partial_\ell\phi(\mu_t) \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T). \tag{3.4.5}$$

Proof. **Step 1.** Let us denote the finite infimum of ϕ by $\imath$. Consider a sequence μ_τ^k of discrete minimizers for (3.2.1). Since

$$\Phi_\tau(\mu_\tau^k, \mu_\tau^{k-1}) = \min_{\nu \in \mathscr{P}_2(X)} \Phi_\tau(\nu, \mu_\tau^{k-1}),$$

we readily get

$$\frac{1}{2\tau} W_2^2(\mu_\tau^k, \mu_\tau^{k-1}) + \phi(\mu_\tau^k) \le \phi(\mu_\tau^{k-1}), \tag{3.4.6}$$

hence

$$\phi(\mu_\tau^k) \le \phi(\mu_\tau^0) \quad \forall k \in \mathbb{N}.$$

This means that

$$\sup_{t \in [0,T]} \phi(\bar\mu_\tau(t)) \le \phi(\mu_\tau^0) \quad \forall \tau > 0. \tag{3.4.7}$$

Moreover, summing up in (3.4.6), we have

$$\frac{1}{2\tau} \sum_{k=1}^{n} W_2^2(\mu_\tau^k, \mu_\tau^{k-1}) \le \phi(\mu_\tau^0) - \phi(\mu_\tau^n) \le \phi(\mu_\tau^0) - \imath \le C, \tag{3.4.8}$$

where C is a suitable positive constant which does not depend on τ (the choice of such a constant is possible thanks to (3.2.3)). Let now $n > m$. The triangle inequality entails

$$W_2(\mu_\tau^m, \mu_\tau^n) \le \sum_{k=m+1}^{n} W_2(\mu_\tau^k, \mu_\tau^{k+1}),$$

and, by means of the elementary inequality $\left(\sum_{i=1}^N a_i\right)^2 \le N \sum_{i=1}^N a_i^2$, we get

$$W_2(\mu_\tau^m, \mu_\tau^n) \le \left(\frac{1}{\tau} \sum_{k=m+1}^{n} W_2^2(\mu_\tau^k, \mu_\tau^{k-1}) \right)^{1/2} ((n-m)\tau)^{1/2}.$$

Exploiting (3.4.8), we obtain

$$W_2(\mu_\tau^m, \mu_\tau^n) \le \sqrt{2(n-m)C\tau}. \tag{3.4.9}$$

Considering (3.4.7) and (3.2.3), we learn that the family $\{\bar\mu_\tau(t)\}_{t\in[0,T],\tau>0}$ is a bounded set contained in a sublevel of ϕ, hence it is relatively compact in $\mathscr{P}_2(X)$ by hypothesis. Moreover, connecting the consecutive values μ_τ^k, μ_τ^{k+1} for instance with Wasserstein geodesics, we obtain a family, parametrized by τ, of continuous curves in $\mathscr{P}_2(X)$. Let us denote such family by $\{\hat\mu_\tau(\cdot)\}_{\tau>0}$. By virtue of (3.4.9), each curve in this family has $C^{0,1/2}$ regularity:

$$W_2^2(\hat\mu_\tau(t), \hat\mu_\tau(s)) \le C|t-s|,$$

where the constant C does not depend on τ. This uniform equicontinuity, together with the compactness of the set of values assumed in $\mathscr{P}_2(X)$ by the curves, allows to apply the Ascoli-Arzelà theorem and extract a sequence uniformly converging in $[0,T]$, that is, a sequence of functions $(\hat{\mu}_{\tau_h}(\cdot))$ converging to some μ_t in $C^0([0,T];\mathscr{P}_2(X))$. On the other hand, taking advantage again of (3.4.9) we get

$$W_2(\bar{\mu}_\tau(t),\mu_t) \le W_2(\hat{\mu}_\tau(t),\bar{\mu}_\tau(t)) + W_2(\hat{\mu}_\tau(t),\mu_t) \le C\sqrt{\tau} + W_2(\hat{\mu}_\tau(t),\mu_t) \quad \forall t\in[0,T],$$

and since $\hat{\mu}_{\tau_h}$ goes to μ_t uniformly as $h\to\infty$, we have $W_2(\bar{\mu}_{\tau_h}(t),\mu_t)\to 0$, hence

$$\bar{\mu}_{\tau_h}(t)\to\mu_t \text{ in } \mathscr{P}_2(X) \quad \forall t\in[0,T]. \tag{3.4.10}$$

Now we have to show that $\mu_t\in AC([0,T];\mathscr{P}_2(X))$. For, define

$$|\bar{\mu}'_\tau|(t) := \frac{W_2(\mu_\tau^k,\mu_\tau^{k-1})}{\tau} \quad \text{for } t\in((k-1)\tau,k\tau] \tag{3.4.11}$$

and notice that, for any $k>0$,

$$\int_{(k-1)\tau}^{k\tau} |\bar{\mu}'_\tau|(t)\,dt = W_2(\mu_\tau^k,\mu_\tau^{k-1}) \quad \text{and} \quad \int_{(k-1)\tau}^{k\tau} |\bar{\mu}'_\tau|^2(t)\,dt = \frac{W_2^2(\mu_\tau^k,\mu_\tau^{k-1})}{\tau}. \tag{3.4.12}$$

Fixing $n>T/\tau$ and taking advantage of (3.4.8), we find

$$\int_0^T |\bar{\mu}'_\tau|^2(t)\,dt \le \frac{1}{\tau}\sum_{k=1}^n W_2^2(\mu_\tau^k,\mu_\tau^{k-1}) \le 2C.$$

This shows that the family of functions $\{|\bar{\mu}'_\tau|(\cdot)\}_{\tau>0}$ is bounded in $L^2((0,T))$, so that it possesses a subsequence weakly converging to some $U(\cdot)$ in $L^2((0,T))$. Let $0\le s_1<s_2\le T$, and making use of (3.2.5) and of the triangle inequality we get

$$W_2(\bar{\mu}_{\tau_h}(s_1),\bar{\mu}_{\tau_h}(s_2)) = W_2(\mu_\tau^{\lceil s_1/\tau\rceil},\mu_\tau^{\lceil s_2/\tau\rceil}) \le \int_{(\lceil s_1/\tau_h\rceil)\tau_h}^{(\lceil s_2/\tau_h\rceil)\tau_h} |\bar{\mu}'_{\tau_h}|(t)\,dt. \tag{3.4.13}$$

As $\bar{\mu}_{\tau_h}(t)\to\mu_t$ for any $t\in[0,T]$, we can invoke the semicontinuity property of W_2, proved in Lemma 2.2.6. On the other hand, we have the just stated weak compactness of $\{|\bar{\mu}'_\tau|(\cdot)\}_{\tau>0}$ in $L^2((0,T))$. Hence, possibly extracting one more subsequence, from (3.4.13) we get

$$W_2(\mu_{s_1},\mu_{s_2}) \le \limsup_{h\to\infty} W_2(\bar{\mu}_{\tau_h}(s_1),\bar{\mu}_{\tau_h}(s_2)) \le \int_{s_1}^{s_2} U(t)\,dt,$$

which shows the absolute continuity of μ_t.

Step 2. Let, as usual, $\mathbf{t}_\tau^{k+1}$ be the optimal transport map between μ_τ^{k+1} and μ_τ^k and recall the definition of discrete velocity given by (3.2.7)-(3.2.8). Let $\gamma_\tau^k = (\mathbf{I},\mathbf{V}_\tau^k)_\#\mu_\tau^k$ and

$$\bar{\gamma}_\tau(t) := (\mathbf{I},\overline{\mathbf{V}}_\tau(t))_\#\bar{\mu}_\tau(t) \quad \text{if } t\in((k-1)\tau,k\tau).$$

As in the proof of Theorem 2.3.3, let us introduce space time probability measures over $X \times [0,T]$, in correspondence of time dependent families of probability measures over X: given $t \in [0,T] \mapsto \nu_t \in \mathscr{P}(X)$, let $\nu := \nu_t \bigotimes \frac{1}{T}\mathcal{L}^1 \llcorner (0,T) \in \mathscr{P}(X \times [0,T])$. So, in correspondence of $\bar{\mu}_\tau(t)$ and μ_t, we have the measures $\bar{\mu}_\tau, \mu \in \mathscr{P}(X \times [0,T])$. In particular, since by (3.4.10) $\bar{\mu}_{\tau_h}(t) \rightharpoonup \mu_t$ for any $t \in [0,T]$, there holds

$$\bar{\mu}_{\tau_h} \rightharpoonup \mu \text{ in } \mathscr{P}(X \times [0,T]). \tag{3.4.14}$$

Moreover, the piecewise constant map $t \mapsto \overline{\mathbf{V}}_\tau(t) \in L^2(X, \mu_t; X)$ can be seen as an element of $L^2(X \times [0,T], \bar{\mu}_\tau; X)$. On the other hand, even the family of measures $\bar{\gamma}_\tau(t)$ has its counterpart $\bar{\gamma}_\tau \in \mathscr{P}((X \times [0,T]) \times X)$, and with the notation just introduced there holds

$$\bar{\gamma}_\tau = (\mathbf{I}_T, \overline{\mathbf{V}}_\tau)_\# \bar{\mu}_\tau,$$

where $\mathbf{I}_T : X \times [0,T] \to X \times [0,T]$ is the identity in $X \times [0,T]$. Let x_T, x_2 denote respectively the variable in the first and in the second factor of the product space $(X \times [0,T]) \times X$. If $n > T/\tau$, we have, by (3.4.8),

$$\int_{X \times [0,T]} |\overline{\mathbf{V}}_\tau(x_T)|^2 \, d\bar{\mu}_\tau(x_T) \leq \frac{1}{T} \sum_{k=1}^{n} \int_X \frac{|x - \mathbf{t}_\tau^k(x)|^2}{\tau} \, d\mu_\tau^k(x) \leq \frac{2C}{T}. \tag{3.4.15}$$

Together with (3.4.14), this shows that the hypotheses of point *i)* of Lemma 1.5.4 are satisfied, yielding tightness for the family $\bar{\gamma}_\tau$ in the $\mathscr{P}((X_\varpi \times [0,T]) \times X_\varpi)$ topology. Then there exists a subsequence (that we don't relabel) of τ_h such that $\bar{\gamma}_{\tau_h} \rightharpoonup \gamma$ in $\mathscr{P}((X_\varpi \times [0,T]) \times X_\varpi)$, and since $\pi^1_\# \bar{\gamma}_{\tau_h} = \bar{\mu}_{\tau_h}$ for any h, at the limit we find $\pi^1_\# \gamma = \mu$ (here π^1 denotes the projection on the first factor $X \times [0,T]$). We define the limit velocity as

$$\mathbf{v}(x_T) := \int_X x_2 \gamma_{x_T}(x_2), \tag{3.4.16}$$

that is, $\mathbf{v}$ is the barycenter (see Definition 1.4.8) of the family $x_T \in (X \times [0,T]) \mapsto \gamma_{x_T} \in \mathscr{P}(X)$ which disintegrates γ w.r.t. its first marginal μ. Since $\mathbf{v}$ is the barycenter, by Lemma 1.5.4 we also have that $\overline{\mathbf{V}}_{\tau_h}$ weakly converges to $\mathbf{v} \in L^2(X \times [0,T], \mu; X)$ in the sense of Definition 1.5.3 and

$$\int_{X_T} |\mathbf{v}|^2 \, d\mu \leq \liminf_{h \to \infty} \int_{X_T} |\overline{\mathbf{V}}_{\tau_h}|^2 \, d\bar{\mu}_{\tau_h}. \tag{3.4.17}$$

Step 3. We have to check that the vector $\mathbf{v}$ in (3.4.16) satisfies the continuity equation, coupled with μ. Let $\varphi \in \mathrm{Cyl}(X)$. Then

$$\begin{aligned} \int_X \varphi(x) \, d\mu_\tau^{k+1}(x) - \int_X \varphi(x) \, d\mu_\tau^k(x) &= \int_X (\varphi(x) - \varphi(\mathbf{t}_\tau^{k+1}(x))) \, d\mu_\tau^{k+1}(x) \\ &= \int_X \langle \nabla\varphi(x), x - \mathbf{t}_\tau^{k+1}(x) \rangle \, d\mu_\tau^{k+1}(x) + \mathcal{R}_\tau^{k+1} \\ &= \tau \int_{X \times X} \langle \nabla\varphi(x_1), x_2 \rangle \, d\gamma_\tau^{k+1}(x_1, x_2) + \mathcal{R}_\tau^{k+1}, \end{aligned}$$

where $\mathcal{R}_\tau^{k+1}$ is the remainder of the Taylor expansion. Writing it in integral form we see that

$$\begin{aligned}\mathcal{R}_\tau^{k+1} &= \frac{1}{2}\int_0^1 \int_X \langle \nabla^2\varphi(x)(x-\mathbf{t}_\tau^{k+1}(x)), x-\mathbf{t}_\tau^{k+1}(x)\rangle \, d\mu_\tau^{k+1}(x) \\ &\le \tau^2 \sup|\nabla^2\varphi| \int_X |\mathbf{V}_\tau^{k+1}(x)|^2 \, d\mu_\tau^{k+1}(x).\end{aligned} \tag{3.4.18}$$

Let $\psi \in \mathrm{Cyl}(X_T)$ and choose $\varphi(\cdot) = \psi(\cdot, t)$. Recall that $\bar\mu_\tau$ is a measure on $X \times [0,T]$ and that its disintegration with respect to $\frac{1}{T}\mathcal{L}^1 \llcorner([0,T])$ is $\bar\mu_\tau(t)$, which is given by μ_τ^k for $t \in ((k-1)\tau, k\tau]$. Let also $\bar{\mathcal{R}}_\tau(t) = \mathcal{R}_\tau^k$ for $t \in ((k-1)\tau, k\tau]$. Hence, by the previous estimate,

$$\begin{aligned}\int_{X_T} \partial_t\psi(x,t)\, d\mu(x,t) &= \lim_{h\to\infty}\int_{X_T} \partial_t\psi(x,t)\, d\bar\mu_{\tau_h}(x,t) \\ &= \lim_{h\to\infty}\int_{X_T} \frac{\psi(x,t+\tau_h)-\psi(x,t)}{\tau_h}\, d\bar\mu_{\tau_h}(x,t) \\ &= -\lim_{h\to\infty}\int_{X_T\times X} \langle \nabla\psi(x_T), x_2\rangle\, \bar\gamma_{\tau_h}(x_T,x_2) - \lim_{h\to\infty}\frac{1}{\tau_h}\int_0^T \bar{\mathcal{R}}_{\tau_h}(t)\, dt.\end{aligned}$$

But (3.4.18) and (3.4.15) show that the last limit is zero. We are left with

$$\int_{X_T} \partial_t\psi(x,t)\, d\mu(x,t) = -\int_{X\times[0,T]} \langle \nabla\varphi(x,t), \mathbf{v}(x,t)\rangle\, d\mu(x,t),$$

where $\mathbf{v}$ is the limiting velocity defined by (3.4.16). This is the continuity equation.

Finally, we are left to prove that the limiting velocity $\mathbf{v}(x,t) = \mathbf{v}_t(x)$, defined by (3.4.16), satisfies (3.4.5). Splitting the integral in (3.4.15) and making use of Fatou Lemma, we see that we can find a $\mathcal{L}^1$-negligible set $N \subset (0,T)$ such that

$$\liminf_{h\to\infty}\int_X |\overline{\mathbf{V}}_{\tau_h}(t)|^2\, d\bar\mu_{\tau_h}(t) < +\infty \quad \forall t \in (0,T)\setminus N. \tag{3.4.19}$$

Here $\overline{\mathbf{V}}_{\tau_h}(t)$ is seen, for any fixed t, as a $L^2(X,\mu_t;X)$ vector field. This implies that, for any $t \notin N$, $\overline{\mathbf{V}}_{\tau_h}(t)$ has weak limit points in the sense of definition 1.5.3 (by Lemma 1.5.4). Since $-\overline{\mathbf{V}}_{\tau_h}(t)$ are subdifferentials of ϕ at $\bar\mu_{\tau_h}(t)$, from Definition 3.4.3 we see that $\partial_\ell\phi(\mu_t) \neq \emptyset$ for any $t \notin N$. Recall that the limiting plan γ is given by $(\mathbf{I}_T, \mathbf{v})_\#\mu \in \mathscr{P}(X_T \times X)$. It is the limit of $(\mathbf{I}_T, \overline{\mathbf{V}}_{\tau_h})_\#\bar\mu_{\tau_h}$ in the $\mathscr{P}(X_\varpi \times [0,T] \times X_\varpi)$ topology (on a subsequence, that here we are not relabeling). Disintegrating w.r.t. t we get the family of plans $t \mapsto \bar\gamma_{\tau_h}(t) \in \mathscr{P}(X \times X)$, where $\bar\gamma_{\tau_h}(t) = (\mathbf{I}, \overline{\mathbf{V}}_{\tau_h}(t))_\#\bar\mu_{\tau_h}(t)$. But (3.4.19) implies, by virtue of Lemma 1.5.4, that $\bar\gamma_{\tau_h}(t)$ has limit points in the $\mathscr{P}(X_\varpi \times X_\varpi)$ topology, for any $t \notin N$. Let G_t denote the corresponding sets of accumulation points. On the other hand, the disintegration of γ w.r.t. t is $(\mathbf{I}, \mathbf{v}_t)_\#\mu_t$. Invoking Lemma 1.3.3 we get

$$(\mathbf{I}, \mathbf{v}_t)_\#\mu_t \in \overline{\mathrm{Conv}} G_t \quad \text{for } \mathcal{L}^1\text{-a.e. } t \in [0,T]. \tag{3.4.20}$$

But limit points of $\bar{\gamma}_{\tau_h}(t)$ correspond to weak (in the sense of Definition 1.5.3) limits of $\overline{\mathbf{V}}_{\tau_h}(t)$. Indeed, by Lemma 1.5.4, the latter are barycenters of the elements of G_t. Moreover, the opposites of the limits of $\overline{\mathbf{V}}_{\tau_h}(t)$ are the elements of the limiting subdifferential of ϕ at μ_t. By (3.4.20), $(\mathbf{I}, \mathbf{v}_t)_{\#}\mu_t$ is the limit (in $\mathscr{P}(X_\varpi \times X_\varpi)$, for fixed t) of a convex combination of elements of G_t. By the linearity of the barycentric projection operation, we get

$$-\mathbf{v}_t \in \overline{\mathrm{Conv}} \partial_\ell \phi(\mu_t) \quad \text{for } \mathcal{L}^1\text{-a.e. } t \in [0, T].$$

Here the closure is the strong $L^2(\mu_t)$ closure (strong and weak closure are the same for convex sets). The proof is concluded. □

3.5 Existence of solutions: the convex case

In the convex case the minimization scheme inherits very nice properties. Before stating the main convergence theorem, we will present some intermediate results which will be also useful, in particular in Chapter 4. So, the first part of this section is meant to lead us to a deeper insight on the consequences of λ-geodesical convexity.

Theorem 3.5.1 (Slope and minimal selection in the subdifferential) *Let ϕ be a λ-geodesically convex functional. Then $\mu \in D(|\partial\phi|)$ if and only if $\partial\phi(\mu)$ is not empty and*

$$|\partial\phi|(\mu) = \min\left\{ \|\boldsymbol{\xi}\|_{L^2(X,\mu;X)} : \boldsymbol{\xi} \in \partial\phi(\mu) \right\}. \tag{3.5.1}$$

The minimum is realized in correspondence of a unique vector $\boldsymbol{\xi} \in \partial\phi(\mu)$, which we denote by $\partial^0\phi(\mu)$.

Proof. By the definition of metric slope and Wasserstein subdifferential, we immediately see that

$$|\partial\phi|(\mu) \le \|\boldsymbol{\xi}\|_{L^2(X,\mu;X)} \quad \forall \boldsymbol{\xi} \in \partial\phi(\mu).$$

Let $\mu \in D(|\partial\phi|)$ and let μ_τ be a minimizer of the perturbed functional $\Phi_\tau(\cdot, \mu)$, defined by (3.2.1). Lemma 3.4.1 shows that the rescaled vector $\boldsymbol{\xi}_\tau := (\mathbf{t}_{\mu_\tau}^{\mu} - \mathbf{I})/\tau$ belongs to $\partial\phi(\mu_\tau)$, hence

$$\int_X |\boldsymbol{\xi}_\tau(x)|^2 \, d\mu_\tau(x) = \frac{W_2^2(\mu_\tau, \mu)}{\tau^2}. \tag{3.5.2}$$

By Lemma 3.2.3, we know that there exists a vanishing subsequence (τ_n) such that $\mu_{\tau_n} \to \mu$ and

$$\lim_{n\to\infty} \int_X |\boldsymbol{\xi}_{\tau_n}(x)|^2 \, d\mu_{\tau_n}(x) = |\partial\phi|^2(\mu). \tag{3.5.3}$$

It follows from Lemma 1.5.4 that $\boldsymbol{\xi}_{\tau_n}$ has some weak limit point $\boldsymbol{\xi} \in L^2(X, \mu; X)$, in the sense of Definition 1.5.3, as $\tau \to 0$. We denote by μ_n and $\boldsymbol{\xi}_n$ the corresponding sequences. We have to show that $\boldsymbol{\xi} \in \partial\phi(\mu)$ and $|\partial\phi|(\mu) \ge \|\boldsymbol{\xi}\|_{L^2(X,\mu;X)}$. Let $\nu \in D(\phi)$, with $\mathbf{t}_{\mu_n}^{\nu}$ being the optimal transport map between μ_n and ν as usual. Consider the 3-plan $\boldsymbol{\gamma}_n = (\mathbf{I}, \boldsymbol{\xi}_n, \mathbf{t}_{\mu_n}^{\nu})_{\#}\mu_n$. It is relatively compact in $\mathscr{P}(X \times X_\varpi \times X)$. Indeed, the first and the third marginals are tight in $\mathscr{P}(X)$, whereas the second one is tight in $\mathscr{P}(X_\varpi)$, thanks to (3.5.3) and (1.4.5), and we apply Proposition 1.2.1. By (3.1.7) we have

$$\phi(\nu) \geq \phi(\mu_n) + \int_{X\times X\times X} \langle x_2, x_3 - x_1\rangle \, d\boldsymbol{\gamma}_n(x_1,x_2,x_3) + \frac{1}{2}\lambda W_2^2(\mu_n,\nu). \tag{3.5.4}$$

Let $\boldsymbol{\gamma}$ be a limit point of $\boldsymbol{\gamma}_n$. Since the first and the third marginal of $\boldsymbol{\gamma}$ are converging in $\mathscr{P}_2(X)$ (with moments), there holds

$$\lim_{r\to\infty} \sup_{n\in\mathbb{N}} \int_{(X\times X\times X)\setminus \mathbf{B}_r(\mathbf{0})} |x_i|^2 \, d\boldsymbol{\gamma}_n(x_1,x_2,x_3) = 0$$

for $i=1$ and $i=3$, where $\mathbf{B}_r(\mathbf{0})$ is the centered ball in $X\times X\times X$. Hence we can apply Lemma 1.5.2 and the lower semicontinuity of ϕ to get

$$\phi(\nu) \geq \phi(\mu) + \int_{X\times X\times X} \langle x_2, x_3 - x_1\rangle \, d\boldsymbol{\gamma}(x_1,x_2,x_3) + \frac{1}{2}\lambda W_2^2(\mu,\nu). \tag{3.5.5}$$

Since we are working under Assumption 3.1.6, μ is regular and there exists the optimal transport map $\mathbf{t}_\mu^\nu$. Then, letting $\gamma = \pi^{1,2}_\#\boldsymbol{\gamma}$, Lemma 2.2.1 gives

$$\int_{X\times X\times X} \langle x_2, x_3 - x_1\rangle \, d\boldsymbol{\gamma}(x_1,x_2,x_3) = \int_{X\times X} \langle x_2, \mathbf{t}_\mu^\nu(x_1) - x_1\rangle \, d\gamma(x_1,x_2),$$

hence

$$\begin{aligned}\phi(\nu) - \phi(\mu) &\geq \int_{X\times X} \langle x_2, \mathbf{t}_\mu^\nu(x_1) - x_1\rangle \, d\gamma(x_1,x_2) + \frac{1}{2}\lambda W_2^2(\mu,\nu)\\ &\geq \int_X \langle \bar\gamma(x), \mathbf{t}_\mu^\nu(x_1) - x_1\rangle \, d\mu(x) + \frac{1}{2}\lambda W_2^2(\mu,\nu),\end{aligned}$$

where $\bar\gamma$ is the barycenter of γ, which corresponds to the limit $\boldsymbol{\xi}$ of $\boldsymbol{\xi}_n$, by virtue of Lemma 1.5.4. We conclude that $\boldsymbol{\xi} \in \partial\phi(\mu)$ and then (3.5.3) and (1.5.9) imply $|\partial\phi|(\mu) \geq \|\boldsymbol{\xi}\|_{L^2(X,\mu;X)}$. □

Proposition 3.5.2 (Convexity of Φ) *Let ϕ be λ-geodesically convex, let $\tau > 0$. Let moreover $\mu^0, \mu^1 \in D(\phi)$ and γ_t be a constant speed geodesic connecting μ^0 and μ^1. Then there holds*

$$\Phi_\tau(\gamma_t,\mu^0) \leq (1-t)\Phi_\tau(\mu^0,\mu^0) + t\Phi_\tau(\mu^1,\mu^0) - \frac{1}{2}\left(\frac{1}{\tau}+\lambda\right) t(1-t) W_2^2(\mu^0,\mu^1). \tag{3.5.6}$$

Proof. A simple computation yields

$$\begin{aligned}\Phi_\tau(\gamma_t,\mu^0) &= \frac{1}{2\tau}W_2^2(\mu^0,\gamma_t) + \phi(\gamma_t)\\ &= \frac{1}{2\tau}t^2 W_2^2(\mu^0,\mu^1) + \phi(\gamma_t)\\ &\leq \frac{1}{2\tau}t^2 W_2^2(\mu^0,\mu^1) + (1-t)\phi(\mu^0) + t\phi(\mu^1) - \frac{1}{2}\lambda t(1-t) W_2^2(\mu^0,\mu^1)\\ &= (1-t)\phi(\mu^0) + t\left(\phi(\mu^1) + \frac{1}{2\tau}W_2^2(\mu^0,\mu^1)\right) - \frac{1}{2}\left(\frac{1}{\tau}+\lambda\right) t(1-t) W_2^2(\mu^0,\mu^1),\end{aligned}$$

where we made use of the definition of λ-geodesical convexity. □

Remark 3.5.3 By the previous proposition we can not infer that $\Phi_\tau(\cdot,\mu^0)$ is $(\lambda+1/\tau)$-geodesically convex. We have convexity only along geodesics starting from the base point μ^0, and not for each couple of points in $\mathscr{P}_2(X)$. By the way, when needed we will ask strong geodesical convexity (see Definition 3.1.4). It is immediate to see that under such assumption, $\Phi_\tau(\cdot,\mu^0)$, with μ^0 given in $\mathscr{P}_2(X)$, is 1-convex along a suitable curve (not necessarily a Wasserstein geodesic) with respect to an arbitrary base point.

Lemma 3.5.4 (Uniqueness of minimizers for strictly convex functionals) *Let $\lambda > 0$. Let ϕ be λ-geodesically convex. Then there exists a unique minimizer $\tilde\mu$ for ϕ in $\mathscr{P}_2(X)$.*

Proof. Let ν_n be a minimizing sequence, so that there exists $(\omega_n)\in\mathbb{R}$, $\omega_n\to 0$, such that

$$\phi(\nu_n) \le \inf_{\nu\in\mathscr{P}_2(X)} \phi(\nu) + \omega_n. \tag{3.5.7}$$

ϕ is λ-geodesically convex, so that, by definition, there exists a geodesic $\nu_t^{n,m}$ connecting ν_n to ν_m such that there holds

$$\frac{\phi(\nu_t^{n,m})-\phi(\nu_n)}{t} \le \phi(\nu_m)-\phi(\nu_n)-\frac{1}{2}\lambda(1-t)W_2^2(\nu_n,\nu_m).$$

Choosing $t=1/2$ in this relation, we find

$$2\phi(\nu_{1/2}^{n,m})-2\phi(\nu_n)\le \phi(\nu_m)-\phi(\nu_n)-\frac{1}{4}\lambda W_2^2(\nu_n,\nu_m),$$

hence,

$$\frac{1}{8}\lambda W_2^2(\nu_n,\nu_m)\le \frac{1}{2}\,\phi(\nu_m)+\frac{1}{2}\,\phi(\nu_n)-\phi(\nu_{1/2}^{n,m})\le \frac{1}{2}\,\omega_m+\frac{1}{2}\,\omega_n,$$

where we made use of (3.5.7). Since $\lambda>0$ and $\omega_n\to 0$, we learn that ν_n is a Cauchy sequence in $\mathscr{P}_2(X)$, which is complete since X is a Hilbert space (see [72, Lemma 6.12]). Then there exists ν such that $\nu_n\to\nu$ in $\mathscr{P}_2(X)$. Since ϕ is l.s.c., we conclude that ν is a minimizer for ϕ. Moreover, it is the unique minimizer. Indeed, suppose ν_1 and ν_2 and are distinct minimum points. Let ν_t be a Wasserstein geodesic connecting them. Since ϕ is λ-geodesically convex, ν_t can be chosen such that (3.1.3) holds, and such relation immediately gives a contradiction. □

In the λ-geodesical convex case, we can give the following representation for the metric slope.

Lemma 3.5.5 (Metric slope in the convex case) *Let $\lambda\in\mathbb{R}$. If ϕ is λ-geodesically convex, then there holds*

$$|\partial\phi|(\nu)=\sup_{\mu\ne\nu}\left(\frac{\phi(\nu)-\phi(\mu)}{W_2(\nu,\mu)}+\frac{1}{2}\lambda W_2(\nu,\mu)\right)^+ \quad \forall\nu\in D(\phi). \tag{3.5.8}$$

Moreover, the application $\nu\mapsto|\partial\phi|(\nu)$ is l.s.c. in $\mathscr{P}_2(X)$.

Proof. It is clear that

$$|\partial\phi|(\nu) = \limsup_{\mu\to\nu \text{ in}\mathscr{P}_2(X)} \frac{(\phi(\nu)-\phi(\mu))^+}{W_2(\nu,\mu)} \le \sup_{\mu\neq\nu}\left(\frac{\phi(\nu)-\phi(\mu)}{W_2(\nu,\mu)} + \frac{1}{2}\lambda W_2(\nu,\mu)\right)^+ .$$

On the other hand, consider that, by λ-geodesical convexity of ϕ, there exists a Wasserstein geodesic ν_t connecting ν and μ such that

$$\frac{\phi(\nu)-\phi(\nu_t)}{W_2(\nu,\nu_t)} \ge \frac{tW_2(\nu,\mu)}{W_2(\nu,\nu_t)}\left(\frac{\phi(\nu)-\phi(\mu)}{W_2(\nu,\mu)} + \frac{1}{2}\lambda(1-t)W_2(\nu,\mu)\right) \tag{3.5.9}$$

Clearly there holds

$$|\partial\phi(\nu)| \ge \limsup_{t\downarrow 0} \frac{\phi(\nu)-\phi(\nu_t)}{W_2(\nu,\nu_t)},$$

and using $W_2(\nu,\nu_t) = tW_2(\nu,\mu)$, from (3.5.9) we get

$$|\partial\phi(\nu)| \ge \left(\frac{\phi(\nu)-\phi(\mu)}{W_2(\nu,\mu)} + \frac{1}{2}\lambda W_2(\nu,\mu)\right)^+ .$$

Taking the supremum over $\mu\neq\nu$, we get (3.5.8). In order to show lower semicontinuity, consider a sequence (ν_n) converging to ν in $\mathscr{P}_2(X)$, and let $\mu\neq\nu$, so that, for n large enough, $\nu_n\neq\mu$. Then, thanks to (3.5.8) and to the lower semicontinuity of ϕ,

$$\begin{aligned}\liminf_{n\to\infty}|\partial\phi(\nu)| &\ge \liminf_{n\to\infty}\left(\frac{\phi(\nu_n)-\phi(\mu)}{W_2(\nu_n,\mu)} + \frac{1}{2}\lambda W_2(\nu_n,\mu)\right)^+ \\ &\ge \left(\frac{\phi(\nu)-\phi(\mu)}{W_2(\nu,\mu)} + \frac{1}{2}\lambda W_2(\nu,\mu)\right)^+ .\end{aligned}$$

Again, the conclusion follows taking the supremum over $\mu\neq\nu$. □

We also state the following result, in the same spirit of Lemma 3.5.4.

Proposition 3.5.6 *Let ϕ be a strongly convex functional, let $\mu\in\overline{D(\phi)}$ and $\tau>0$. Then $\Phi_\tau(\cdot,\mu)$ admits a unique minimizer $\mu_\tau\in\mathscr{P}_2(X)$.*

Proof. Let $\phi_\tau(\mu)$ be the Moreau-Yosida approximation of ϕ, (see (3.2.9)). ϕ_τ depends continuously on μ (with respect to the W_2 convergence), as shown in point *i)* of Lemma 3.2.3. Suppose now that (ν_n) is a minimizing sequence for $\Phi_\tau(\cdot,\mu)$. Then, since $\mu\in\overline{D(\phi)}$, there exists a sequence $(\mu_n)\subset D(\phi)$ converging to μ in $\mathscr{P}_2(X)$ such that

$$\limsup_{n\to\infty}\Phi_\tau(\nu_n,\mu_n) = \limsup_{n\to\infty}\Phi_\tau(\nu_n,\mu) \le \phi_\tau(\mu). \tag{3.5.10}$$

Let us now take advantage of (3.1.4), choosing a continuous curve μ_t connecting ν_n to ν_m, with μ_n as a base point. We obtain

$$\begin{aligned}\Phi_\tau(\mu_{1/2},\mu_n) &= \phi(\mu_{1/2}) + \frac{1}{2\tau}W_2^2(\mu_{1/2},\mu_n) \\ &\le \frac{1}{2}\left(\phi(\nu_n) + \frac{1}{2\tau}W_2^2(\nu_n,\mu_n)\right) + \frac{1}{2}\left(\phi(\nu_m) + \frac{1}{2\tau}W_2^2(\nu_m,\mu_n)\right) - \frac{W_2^2(\nu_n,\nu_m)}{8\tau}.\end{aligned} \tag{3.5.11}$$

Now notice that the left hand side can be bounded from below with $\phi_\tau(\mu_n)$, while the first two terms in the right hand side are asymptotically smaller than $\frac{1}{2}\phi_\tau(\mu)$, by (3.5.10). We conclude that $W_2^2(\nu_n, \nu_m) \to 0$ as n, $m \to \infty$, whence $\nu_n \to \nu$ in $\mathscr{P}_2(X)$. Then, it follows easily that ν is a minimizer. Since the minimizing sequence was chosen arbitrarily, we also conclude that ν is the unique minimizer. □

Proposition 3.5.7 *Let $\phi : \mathscr{P}_2(X) \to (-\infty, +\infty]$ be a strongly convex functional, let $\mu^0 \in \mathscr{P}_2(X)$ and let μ_τ be a minimizer of $\Phi_\tau(\cdot, \mu^0)$. Then*

$$W_2^2(\mu_\tau, \nu) - W_2^2(\mu^0, \nu) \;\leq\; 2\,\tau\,[\phi(\nu) - \phi(\mu_\tau)] \qquad \forall \nu \in D(\phi). \tag{3.5.12}$$

Proof. Let $\nu_0 = \mu_\tau$, $\nu_1 = \nu$ and consider the interpolating curve $\nu_t : [0,1] \to \mathscr{P}_2(X)$ along which (3.1.4) holds. The minimality of μ_τ and (3.1.4) give

$$\begin{aligned}
&\phi(\mu_\tau) + \frac{1}{2\tau} W_2^2(\mu_\tau, \mu^0) \leq \phi(\nu_t) + \frac{1}{2\tau} W_2^2(\nu_t, \mu^0) \\
&\leq (1-t)\left[\phi(\mu_\tau) + \frac{1}{2\tau} W_2^2(\mu_\tau, \mu^0)\right] + t\left[\phi(\nu) + \frac{1}{2\tau} W_2^2(\nu, \mu^0)\right] - \,t(1-t)\,\frac{1}{2\tau} W_2^2(\mu_\tau, \nu).
\end{aligned}$$

Subtracting $\phi(\mu_\tau) + W_2^2(\mu_\tau, \mu^0)/2\tau$ and dividing by $t > 0$ we obtain:

$$\begin{aligned}
\phi(\nu) - \phi(\mu_\tau) \;&\geq\; \frac{1}{2\tau} W_2^2(\mu_\tau, \mu^0) \;-\; \frac{1}{2\tau} W_2^2(\nu, \mu^0) \;+\; \frac{1-t}{2\tau} W_2^2(\mu_\tau, \nu) \\
&\geq\; \frac{1}{2\tau}\left[(1-t)\, W_2^2(\mu_\tau, \nu) \;-\; W_2^2(\nu, \mu^0)\right].
\end{aligned}$$

Letting $t \downarrow 0$ we have

$$W_2^2(\mu_\tau, \nu) - W_2^2(\nu, \mu^0) + W_2^2(\mu_\tau, \mu^0) \leq 2\tau\Big(\phi(\nu) - \phi(\mu_\tau)\Big), \tag{3.5.13}$$

which yields (3.5.12). □

Now we are going to discuss the convergence results for the convex case. We will see that stronger conclusions then the ones of Theorem 3.4.4 hold, even omitting one of the hypotheses therein, that is, the compactness for sublevels of ϕ. For simplicity, and since it is enough for our subsequent purposes, we will restrict the analysis to the case $\lambda = 0$, and we address the reader to [4] for the general case. The main assumption here is the strong geodesical convexity of functional ϕ, as introduced in Definition 3.1.4.

Theorem 3.5.8 (Existence and uniqueness of gradient flows) *Let ϕ be a strongly convex functional. Then, for all $\mu^0 \in \overline{D(\phi)}$, there exists a limit curve μ_t for the minimizing movements scheme, starting from μ^0, which is solution to* (3.3.5) *and the unique gradient flow. μ_t is given by a contraction semigroup $S(t)$ on $\overline{D(\phi)}$ and*

$$W_2(S(t)(\mu^0), S(s)(\mu^0)) \leq \sqrt{2C}\sqrt{|t-s|}, \qquad t,\, s \geq 0,\ \mu^0 \in \mathscr{P}_2(X), \tag{3.5.14}$$

where $C = \phi(\mu^0) - \inf \phi$. The solution here provided satisfies the following additional properties.

i) Let $\bar{\mu}_\tau(t)$ be defined by (3.2.4). There holds

$$W_2(S(t)(\mu^0), \bar{\mu}_\tau(t)) \leq (3\sqrt{2}+4)\sqrt{\tau C}$$

if $\mu^0 \in D(\phi)$;

ii) the following regularizing effect holds:

$$\phi(S(t)(\mu^0)) \leq \inf_{\nu \in D(\phi)} \frac{1}{2t} W_2^2(\mu^0, \nu) + \phi(\nu) < +\infty \tag{3.5.15}$$

for all $t > 0$, $\mu^0 \in \overline{D(\phi)}$. Moreover, $S(t)\mu^0 \in D(|\partial\phi|)$ for any $t > 0$.

iii) If μ_0 is a minimum point for ϕ and $t > 0$, then

$$\phi(S(t)(\mu^0)) - \phi(\mu_0) \leq \frac{W_2^2(\mu^0, \mu_0)}{2t}$$

and the map $t \mapsto W_2(S(t)(\mu^0), \mu_0)$ is nonincreasing.

Proof. Concerning uniqueness, we refer to Theorem 3.3.4.

In order to show existence, we consider first the case when $\mu^0 \in D(\phi)$. Let $\mu_\tau^0 = \mu^0$. μ_τ^{k+1} satisfies

$$\phi(\mu_\tau^{k+1}) + \frac{1}{2\tau} W_2^2(\mu_\tau^{k+1}, \mu_\tau^k) \leq \phi(\nu) + \frac{1}{2\tau} W_2^2(\nu, \mu_\tau^k) \qquad \forall \nu \in D(\phi) \tag{3.5.16}$$

and reasoning as in the first part of the proof of Theorem 3.4.4, we are led again to the discrete $C^{1/2}$ estimate

$$W_2(\bar{\mu}_\tau(t), \bar{\mu}_\tau(s)) \leq \sqrt{2C}\sqrt{|t-s|+\tau}, \tag{3.5.17}$$

where $t, s > 0$ and $C > 0$ is given by $\phi(\mu^0) - \inf \phi$ (mind that ϕ is bounded from below). We do not have compactness of sublevels, nonetheless we can make use of the convexity estimate (3.5.12). Let $\nu^0 \in D(\phi)$ and denote as usual by $\bar{\nu}_\tau(t)$ the piecewise constant interpolation of discrete minimizers starting from ν^0. We start proving that

$$W_2^2(\bar{\mu}_\tau(t), \bar{\nu}_{\tau/2}(t)) - W_2^2(\mu^0, \nu^0) \leq 2\tau(\phi(\nu^0) - \imath), \tag{3.5.18}$$

for all $\tau > 0$ and all times t that are integer multiples of τ, where $\imath = \inf \phi$. To this aim, notice that (3.5.12) implies, since $\bar{\mu}_\tau((k+1)\tau) = \mu_\tau^{k+1}$ is the discrete minimizer starting from $\bar{\mu}_\tau(k\tau) = \mu_\tau^k$,

$$W_2^2(\bar{\mu}_\tau((k+1)\tau), \nu) - W_2^2(\bar{\mu}_\tau(k\tau), \nu) \leq 2\tau[\phi(\nu) - \phi(\bar{\mu}_\tau((k+1)\tau))] \tag{3.5.19}$$

for all $\nu \in D(\phi)$. Replacing τ with $\tau/2$, μ with ν, ν with θ and choosing $k = 0$ and $k = 1$ we obtain the inequalities

$$\begin{aligned} W_2^2(\bar{\nu}_{\tau/2}(\tau/2), \theta) - W_2^2(\nu^0, \theta) &\leq \tau\left[\phi(\theta) - \phi(\bar{\nu}_{\tau/2}(\tau/2))\right], \\ W_2^2(\bar{\nu}_{\tau/2}(\tau), \theta) - W_2^2(\bar{\nu}_{\tau/2}(\tau/2), \theta) &\leq \tau\left[\phi(\theta) - \phi(\bar{\nu}_{\tau/2}(\tau))\right], \end{aligned} \tag{3.5.20}$$

for all $\theta \in D(\phi)$. Summing up we have

$$W_2^2(\bar\nu_{\tau/2}(\tau),\theta) - W_2^2(\nu^0,\theta) \le \tau\left[2\phi(\theta) - \phi(\bar\nu_{\tau/2}(\tau/2)) - \phi(\bar\nu_{\tau/2}(\tau))\right] \tag{3.5.21}$$

for all $\theta \in D(\phi)$. Still from (3.5.19) we get

$$W_2^2(\bar\mu_\tau(\tau),\theta) - W_2^2(\mu^0,\theta) \le 2\tau\left[\phi(\theta) - \phi(\bar\mu_\tau(\tau))\right] \qquad \forall \theta \in D(\phi). \tag{3.5.22}$$

Setting $\theta = \bar\mu_\tau(\tau)$ in (3.5.21) and $\theta = \nu^0$ in (3.5.22), we can add the resulting inequalities to obtain

$$W_2^2(\bar\mu_\tau(\tau),\bar\nu_{\tau/2}(\tau)) - W_2^2(\bar\mu_\tau(0),\bar\nu_{\tau/2}(0)) \le 2\tau\left(\phi(\nu^0) - \phi(\bar\nu_{\tau/2}(\tau)\right), \tag{3.5.23}$$

yielding (3.5.18) with $t = \tau$; by adding the inequalities analogous to (3.5.23) between consecutive times $m\tau$, $(m+1)\tau$, for $m = 0,\ldots,N-1$, we obtain

$$W_2^2(\bar\mu_\tau(N\tau),\bar\nu_{\tau/2}(N\tau)) - W_2^2(\mu^0,\nu^0) \le 2\tau\left(\phi(\nu^0) - \phi(\bar\nu_{\tau/2}(N\tau))\right), \tag{3.5.24}$$

that yields (3.5.18).

Now, from (3.5.18) with $\nu^0 = \mu^0$ we get

$$W_2(\bar\mu_{\tau/2^m}(t),\bar\mu_{\tau/2^{m+1}}(t)) \le 2^{-m/2}\sqrt{2\tau(\phi(\mu^0) - \imath)} = 2^{-m/2}\sqrt{2C\tau}$$

for all t that are integer multiples of $\tau/2^m$. If t is an integer multiple of $\tau/2^j$, $j \in \mathbb{N}$, whenever $m \ge j$ we have that t is also integer multiple of $\tau/2^m$, and by triangle inequality we deduce

$$W_2(\bar\mu_{\tau/2^m}(t),\bar\mu_{\tau/2^n}(t)) \le \sum_{i=m}^{n-1} 2^{-i/2}\sqrt{2C\tau} \tag{3.5.25}$$

for all $n > m$. So, for any t which is integer multiple of $\tau/2^j$ (and therefore on a dense set of times, since j is arbitrary) the sequence $(\bar\mu_{\tau/2^n}(t))$ has the Cauchy property and converges in $\mathscr{P}_2(X)$ to some limit, that we shall denote by $\mathscr{S}(t)\mu^0$. Using the discrete $C^{0,1/2}$ estimate (3.5.17) we obtain convergence for all times, as well as the uniform Hölder continuity (3.5.14) of $t \mapsto \mathscr{S}(t)\mu^0$.

We prove now that the curve $t \mapsto \mathscr{S}(t)\mu^0$ is a gradient flow starting from μ^0, by showing that it satisfies the evolution variational inequalities introduced in Theorem 3.3.2. Indeed, we can read (3.5.19) as follows:

$$\frac{d}{dt}\frac{1}{2}W_2^2(\bar\mu_\tau(t),\nu) \le \tau\sum_{i=1}^{\infty}[\phi(\nu) - \phi(\bar\mu_\tau(i\tau))]\delta_{\{i\tau\}}$$

for all $\nu \in D(\phi)$, in the sense of distributions. Passing to the limit as $n \to \infty$, with τ replaced by $\tau/2^n$, the lower semicontinuity of ϕ gives

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mathscr{S}(t)\mu^0,\nu) \le [\phi(\nu) - \phi(\mathscr{S}(t)\mu^0)] \qquad \forall \nu \in D(\phi)$$

in the sense of distributions. This proves that $\mathscr{S}(t)\mu^0$ is a gradient flow starting from $\bar\mu$, and since we proved uniqueness for gradient flows, the semigroup property holds and from now on we let $S(t)\mu^0 = \mathscr{S}(t)\mu^0$. Contractivity of the semigroup for $\mu^0 \in D(\phi)$ follows taking the limit as $\tau \downarrow 0$ in (3.5.18).

Next we prove *i)*. Passing to the limit as $n \to \infty$ in (3.5.25), with $m = j = 0$, we obtain

$$W_2(S(t)\mu^0, \bar\mu_\tau(t)) \le 2(\sqrt{2}+1)\sqrt{C\tau}$$

when t/τ is an integer. Otherwise, let s/τ be the closest integer to t/τ, so that $|t-s| < \tau$. By (3.5.17), (3.5.14) and the latter inequality we have

$$\begin{aligned} W_2(S(t)\mu^0, \bar\mu_\tau(t)) &\le W_2(S(t)\mu^0, S(s)\mu^0) + W_2(S(s)\mu^0, \bar\mu_\tau(s)) + W_2(\bar\mu_\tau(s), \bar\mu_\tau(t)) \\ &\le \sqrt{2}\sqrt{C|t-s|} + 2(\sqrt{2}+1)\sqrt{C\tau} + \sqrt{2}\sqrt{C(|t-s|+\tau)} \\ &\le (3\sqrt{2}+4)\sqrt{C\tau}, \end{aligned}$$

that is *i)*.

For the proof of *ii)* when $\mu^0 \in D(\phi)$, consider the inequalities, following from (3.5.19) and (3.4.6),

$$\begin{aligned} W_2^2(\bar\mu_\tau((k+1)\tau), \nu) - W_2^2(\bar\mu_\tau(k\tau), \nu) &\le 2\tau[\phi(\nu) - \phi(\bar\mu_\tau((k+1)\tau))] \\ &\le 2\tau[\phi(\nu) - \phi(\bar\mu_\tau(N\tau))], \end{aligned}$$

holding for $k = 0, \ldots, N-1$, $\nu \in D(\phi)$. Summing up we get

$$W_2^2(\bar\mu_\tau(N\tau), \nu) - W_2^2(\mu^0, \nu) \le 2N\tau[\phi(\nu) - \phi(\bar\mu_\tau(N\tau))].$$

Replacing now τ by $\tau/2^m$ in this inequality we have

$$\frac{2N\tau}{2^m}\,\phi(\bar\mu_{\tau/2^m}(N\tau/2^m)) \le W_2^2(\mu^0, \nu) + \frac{2N\tau}{2^m}\,\phi(\nu) \quad \forall \nu \in D(\phi),$$

and defining N as the integer part of $2^m t/\tau$ (so that $N\tau/2^m \to t$), we can let $m \to \infty$ to obtain (3.5.15).

In order to prove contractivity and *ii)* when $\mu^0 \in \overline{D(\phi)}$ we use a density argument. Indeed, let $\mu^0_n \in D(\phi)$ be converging to $\mu^0 \in \overline{D(\phi)}$ in $\mathscr{P}_2(X)$: by contractivity we obtain that $S(t)\mu^0_n$ is a Cauchy sequence for all $t \ge 0$, and therefore converges to some limit, that we shall denote by $S(t)\mu^0$. It is not difficult to prove by approximation that $S(t)\mu^0$ is a gradient flow, and it remains to show that it starts from μ^0. We have indeed $W_2(S(t)\mu^0, S(t)\mu^0_n) \le W_2(\mu^0_n, \mu^0)$, so that

$$\limsup_{t\downarrow 0} W_2(S(t)\mu^0, \mu^0) \le 2W_2(\mu^0_n, \mu^0) + \limsup_{t\downarrow 0} W_2(S(t)\mu^0_n, \mu^0_n) = 2W_2(\mu^0_n, \mu^0).$$

Letting $n \to \infty$ we obtain that $S(t)\mu^0 \to \mu^0$ as $t \downarrow 0$.

Notice that, letting $\mu^0 \in \overline{D(\phi)}$, since $\mu_t = S(t)\mu_0$ is a gradient flow, it satisfies in particular (3.3.2). This shows that $\partial\phi(\mu_t)$, $t > 0$, is not empty. By Theorem 3.5.1 we conclude that $\mu_t \in D(|\partial\phi|)$. This way, *ii)* is proved.

Finally, the inequality of *iii)* follows from (3.3.5), with $\lambda = 0$, as well as the monotonicity property for the map $t \mapsto W_2(S(t)\mu^0, \mu_0)$, μ_0 being a minimum point for ϕ.

$\square$

Chapter 4

Nonlinear degenerate diffusion equations in Hilbert spaces

4.1 Description of the problem

We are going to present the first application of the theory outlined in Chapter 3, that is, the construction of solutions of PDEs as gradient flows in the space of probability measures, endowed with the quadratic optimal transportation distance W_2, of suitable energy functionals. Starting from the seminal papers [57, 43], many studies have been devoted to the description of classical and non-classical PDE's in such framework. Here we just mention [1, 26, 27, 58, 59, 60] and we refer to the monographs [4, 73, 72] for a detailed (but already not completely up to date) description of the literature. We have seen in Section 3.5 that this interpretation as a gradient flow, when associated to a convex structure, inherits its full power and is extremely useful to derive existence, stability results and trends to equilibrium. In [4] (see chapters 9, 10 and 11 therein) the approach has been shown to be particularly suitable for evolution equation of diffusion type in $\mathbb{R}^n$, that is

$$\partial_t \rho - \operatorname{div}\left(\rho \nabla (\frac{\delta \mathscr{F}}{\delta \rho})\right) = 0 \quad \text{in } R^n \times (0, +\infty), \tag{4.1.1}$$

characterized by the first variation $\frac{\delta \mathscr{F}}{\delta \rho} = F_z(x, \rho, \nabla \rho) - \operatorname{div} F_p(x, \rho, \nabla \rho)$ of an integral functional like

$$\mathscr{F}(\rho) = \int_{\mathbb{R}^n} F(x, \rho(x), \nabla \rho(x))\, dx.$$

Here $F = F(x, z, p) : \mathbb{R}^n \times [0, +\infty) \times \mathbb{R}^n \to \mathbb{R}$ is a smooth integrand. We see that (4.1.1) has the form of a continuity equation whose velocity vector field is a gradient.

We will focus the attention to the particular case in which F depends directly only on ρ, an instance which gives rise to different equations of interest in the applications. In [4], a full theory for this case has been developed. On the other hand, the analogous result in a infinite dimensional Hilbert space X has been proven therein only for the case of the linear Fokker planck equation (corresponding to $F(\rho) = \rho \log \rho$) with respect to a Gaussian measure on X. Later, in [9], the authors obtained general existence and stability results for infinite-dimensional Fokker-Planck equations spaces associated to log-concave probability measures γ: as we will see, when the PDE is regarded as the gradient flow of the relative entropy functional

$$\rho\gamma \mapsto \int_X \rho \log \rho \, d\gamma$$

with respect to W_2, the log-concavity of γ is (see [4]) precisely the property needed for convexity. More recently these results have also been extended to the Ornstein-Uhlenbeck operator in Wiener spaces (see [37, 49]).

In this chapter we investigate more in detail the nonlinear counterpart of these results, corresponding to general energies

$$\mu = \rho\gamma \mapsto \mathscr{F}(\mu) := \int_X F(\rho)\, d\gamma, \qquad \rho\gamma \in \mathscr{P}_2(X) \tag{4.1.2}$$

(set equal to $+\infty$ if μ is not absolutely continuous with respect to γ). In particular we obtain well-posedness and regularizing properties for nonlinear evolution equations of the form

$$\begin{cases} \partial_t \mu_t - \nabla \cdot (\nabla (L \circ \rho_t)\gamma) = 0 \quad \text{in } X \times (0, +\infty), \\ \lim_{t \downarrow 0} \mu_t = \bar{\mu}, \end{cases} \tag{4.1.3}$$

where ρ_t represents the density of μ_t with respect to γ and $L = L_F : \mathbb{R} \to \mathbb{R}$ is the Legendre transform of F. The reader may consult [30, 31] for a systematic study of evolution PDE's in infinite dimensions and the monograph [71] for the finite-dimensional theory of porous media equations.

As soon as a convex structure is identified, the results Section of 3.5 provide existence and uniqueness of the gradient flow, and several equivalent formulations of the evolution problem; but, the interpretation of this evolution in conventional PDE terms might not be immediate; in the case of Fokker-Planck equations, the connection with the point of view of Dirichlet forms and of Markov processes is completely analyzed in [9], and tools from the theory of optimal transportation are used to show closability of the Dirichlet form $\int \|\nabla u\|^2 \, d\gamma$.

In the nonlinear context provided by (4.1.2), our goal is relate the evolution semigroup in $\mathscr{P}_2(X)$ to the classical viewpoint based on Sobolev spaces and integration by parts. To this aim, we assume that an orthonormal system (that will be considered the reference basis denoted by $\mathbf{e}_j$) of X exists, such that $\partial_{\mathbf{e}_j}\gamma \ll \gamma$ for all $j \geq 1$; notice that this assumption is consistent with the model case of Gaussian measures γ. Notice however that it is not needed for the existence of the evolution semigroup in $\mathscr{P}_2(X)$. On the other hand, in order to have a convex structure we need some structural assumptions on F which cover all nonlinearities $F(z) = z^m$, $m > 1$ (see Assumption 4.4.1) and the log-concavity of γ. This last hypothesis covers all measures γ of the form $e^{-V}\gamma_G$ with γ_G Gaussian and V

convex and lower semicontinuous, but we won't need any absolute continuity assumption w.r.t. a Gaussian.

By Definition 3.3.1, we know that a gradient flow μ_t in $(\mathscr{P}_2(X), W_2)$ is characterized by the continuity equation (in the weak sense of duality with cylindrical functions)

$$\frac{d}{dt}\mu_t + \nabla \cdot (v_t \mu_t) = 0 \tag{4.1.4}$$

coupled with a constitutive equation relating $v_t \in L^2(\mu_t; X)$ to μ_t, namely $-v_t = \partial^0 \mathscr{F}(\mu_t)$. In this context, $\partial^0 \mathscr{F}(\rho\gamma)$ is the element with minimal $L^2(X, \rho\gamma; X)$ norm of $\partial \mathscr{F}(\rho\gamma)$ and the subdifferential relation is of course the Wasserstein one, described in Definition 3.1.7 (recall also the uniqueness of the minimal selection in the convex case, given by Theorem 3.5.1). The optimal transport map $\mathbf{t}$ between $\mu = \rho\gamma$ and ν appears therein, and it turns out that the absolute continuity of all measures $\partial_{\mathbf{e}_j}\gamma$ suffices to show in Theorem 4.3.1 (following with minor variants Theorem 2.1.7) existence and uniqueness of optimal maps. We remark that in comparison with the subdifferential analysis of [4, Chapter 10], our proofs are simplified by the choice of the quadratic exponent ($p = 2$) and by the existence of optimal maps, so that Kantorovich plans do not play an explicit role.

So, most of this chapter will be devoted to the identification of $\partial \mathscr{F}^0(\rho\gamma)$ and, in comparison to the linear Gaussian case considered in [4, 10.4.8], new difficulties are due to the nonlinearity and to the generality of γ. If $\rho \in L^\infty(X, \gamma)$, we shall prove that $\partial \mathscr{F}(\rho\gamma)$ is not empty if and only if $L_F \circ \rho \in W^{1,1}(X, \gamma)$ and $\nabla(L_F \circ \rho)/\rho \in L^2(X, \rho\gamma; X)$; if this is the case, then

$$\frac{\nabla(L_F \circ \rho)}{\rho} = \partial^0 \mathscr{F}(\rho\gamma). \tag{4.1.5}$$

In the case of unbounded densities ρ, membership to the Sobolev space can not be defined because we assume only $\partial_{\mathbf{e}_j}\gamma \ll \gamma$ (the assumption $|\partial_{\mathbf{e}_j}\gamma| \leq C\gamma$ would be incompatible even with the Gaussian case) and the integration by parts formula does not make sense. To overcome this difficulty, we define (in the same spirit of [12, 29]) generalized Sobolev spaces $GW^{1,1}(X, \gamma)$ in the "entropy" sense, by requiring that the truncated functions $T_\alpha(\rho) = -\alpha \vee \rho \wedge \alpha$ belong to $W^{1,1}(X, \gamma)$ for all $\alpha \geq 0$ (see also [25] for a definition of entropy solutions to some degenerate evolution equations). In this class a gradient can still be defined and (4.1.5) remains true. Replacing (4.1.5) into (4.1.4) we find equation (4.1.3).

The main result, namely a well-posedness result for (4.1.3) (see Theorem 4.7.4), will be obtained invoking Theorem 3.5.8. The solutions will inherit the additional properties listed therein, and in particular they will be described by a contraction semigroup on $\mathscr{P}_2(X)$. We conclude noticing that our strategy (based on the perturbation argument, as in [4, Remark 10.4.7]) identifies only the element with minimal norm and not the whole $\partial \mathscr{F}(\rho\gamma)$, in contrast with the known finite-dimensional result, recalled in Theorem 4.5.1. Since the differential inclusion $v_t \in -\partial \mathscr{F}(\mu_t)$ is equivalent to the equation $v_t = -\partial^0 \mathscr{F}(\mu_t)$, our result is sufficient to identify the PDE (4.1.3). A direct analysis of the subdifferential relation seems to require change of variables formulas relative to γ, a problem still open under our weak assumptions on γ.

4.2 Partial derivatives and gradient in Hilbert spaces

In this section we will introduce the weak directional derivatives, through the integration by parts formula, and the Sobolev spaces over X. Moreover we will prove a useful chain rule formula.

Let γ be a probability measure on X and $v \in X$, $v \neq 0$. The Fomin distributional derivative (see for instance [13]) $\partial_v \gamma$ is defined by the canonical duality

$$\langle \partial_v \gamma, \varphi \rangle = -\int_X \partial_v \varphi \, d\gamma, \quad \varphi \in \mathrm{Cyl}(X)$$

where $\partial_v \varphi$ is the partial derivative of φ in the direction v. We say that $\partial_v \gamma$ is an absolutely continuous measure with respect to γ if there exists $g \in L^1(X, \gamma)$ such that

$$\int_X \partial_v \varphi d\gamma = -\int_X g \varphi d\gamma, \quad \forall \varphi \in \mathrm{Cyl}(X). \tag{4.2.1}$$

Throughout this chapter we shall make the following assumption:

Assumption 4.2.1 $\partial_{\mathbf{e}_j} \gamma \ll \gamma$ *for all* $j \geq 1$. *The corresponding Radon-Nikodym derivatives will be denoted by* g_j*.*

Now we can define the distributional partial derivative of a bounded function (see for instance [13]).

Definition 4.2.2 (Partial derivative, gradient, Sobolev spaces) *Under* Assumption 4.2.1*, a function* $u \in L^\infty(X, \gamma)$ *has partial derivative* $\eta_j \in L^1(X, \gamma)$ *if*

$$\int_X \partial_{\mathbf{e}_j} \zeta(x) u(x) \, d\gamma(x) = -\int_X \eta_j(x) \zeta(x) \, d\gamma(x) - \int_X u(x) \zeta(x) g_j(x) \, d\gamma(x) \quad \forall \zeta \in \mathrm{Cyl}(X). \tag{4.2.2}$$

In this case, we write $\eta_j := \partial^\gamma_{\mathbf{e}_j} u$*, and simply* $\partial_{\mathbf{e}_j} u$ *when no ambiguity arises. In addition, if this happens for all* $j \geq 1$ *and* $\sqrt{\sum_j (\partial_{\mathbf{e}_j} u)^2} \in L^p(X, \gamma)$*, we write* $u \in W^{1,p}(X, \gamma)$ *and set*

$$\nabla u := \sum_{j=1}^\infty (\partial_{\mathbf{e}_j} u) \mathbf{e}_j \in L^p(X, \gamma; X).$$

We shall also use the fact that (recall the definition of cylindrical projection u^d given by (1.4.2))

$$\partial_{\mathbf{e}_j}(u^d) = (\partial_{\mathbf{e}_j} u)^d \tag{4.2.3}$$

whenever $\partial_{\mathbf{e}_j} u$ exists and $j \leq d$. In fact, if $j \leq d$, clearly γ_{x^d}, which disintegrates γ with respect to $\gamma^d = \Pi^d_{\#} \gamma$, can be seen as a family of measures on $(X^d)^\perp$, so that, for fixed $x^d \in X^d$, it does not depend on $\langle x, \mathbf{e}_j \rangle$ and then

$$\partial_{\mathbf{e}_j} \int_X u(x) \, d\gamma_{x^d}(x) = \int_X \partial_{\mathbf{e}_j} u(x) \, d\gamma_{x^d}(x).$$

We shall also need a chain rule formula and an existence result γ-a.e. of directional derivatives of Lipschitz functions; we recall briefly their proofs, that can be achieved by standard arguments.

Theorem 4.2.3 (Chain rule) *Let $u \in L^\infty(X,\gamma)$ with $\partial_{\mathbf{e}_j} u \in L^1(X,\gamma)$, and let $f \in Lip(\mathbb{R})$. Then $\partial_{\mathbf{e}_j}(f\circ u) \in L^1(X,\gamma)$ and*

$$\partial_{\mathbf{e}_j}(f\circ u) = f'(u)\partial_{\mathbf{e}_j} u \quad \gamma\text{-a.e. in } X. \tag{4.2.4}$$

More precisely, denoting by Σ the set where f is not differentiable, both $\partial_{\mathbf{e}_j} u = 0$ and $\partial_{\mathbf{e}_j}(f\circ u) = 0$ γ-a.e. on $u^{-1}(\Sigma)$, where (4.2.4) does not make sense.

Proof. We denote by Y the orthogonal subspace to $\mathbf{e}_j$, by $\boldsymbol{\pi} : X \to Y$ the orthogonal projection and by $\gamma_y \in \mathscr{P}(\mathbb{R})$ the disintegrating family of measures, namely

$$\int_X f(x)\,d\gamma(x) = \int_Y \left(\int_{(\boldsymbol{\pi})^{-1}(y)} f(x)\,d\gamma_y(x)\right) d(\boldsymbol{\pi}_\#\gamma)(y) = \int_X \int_{\mathbb{R}} f(y+t\mathbf{e}_j)\,d\gamma_y(t)\,d(\boldsymbol{\pi}_\#\gamma)(y)$$

for any Borel map $f : X \to \mathbb{R}$, or

$$\gamma(B) = \int_Y \gamma_y\big(\{t:\ y + t\mathbf{e}_j \in B\}\big)\,d\boldsymbol{\pi}_\#\gamma$$

for all Borel sets $B \subset X$, as seen with (1.3.3) and (1.3.4). We claim that $\gamma_y \ll \mathscr{L}^1$ for $\boldsymbol{\pi}_\#\gamma$-a.e. y. To prove this, we shall prove that γ_y has derivative equal to $f_y\gamma_y$, where $f_y(t) = g_j(y+t\mathbf{e}_j)$, and use the well known fact that this property, on the real line, implies absolute continuity.

To prove the claim, fix $\zeta \in \mathrm{Cyl}(Y)$ and $\psi \in C^1_c(\mathbb{R})$ and notice that, disintegrating,

$$\begin{aligned}
\int_Y \zeta(y)\int_{\mathbb{R}} \psi'(t)\,d\gamma_y(t)\,d\boldsymbol{\pi}_\#\gamma(y) &= \int_X \zeta(\boldsymbol{\pi}(x))\psi'(\langle x,\mathbf{e}_j\rangle)\,d\gamma(x)\\
&= -\int_X \zeta(\boldsymbol{\pi}(x))\psi(\langle x,\mathbf{e}_j\rangle)g_j(x)\,d\gamma(x)\\
&= -\int_Y \zeta(y)\int_{\mathbb{R}} \psi(t)f_y(t)\,d\gamma_y(t)\,d\boldsymbol{\pi}_\#\gamma(y),
\end{aligned}$$

where we have integrated by parts with (4.2.2) and used the fact that $\partial_{\mathbf{e}_j}\zeta(\boldsymbol{\pi}(x)) = 0$. Since ζ is arbitrary, $\int_{\mathbb{R}} \psi'(t)\,d\gamma_y(t) = -\int_{\mathbb{R}} \psi(t)f_y(t)\,d\gamma_y(t)$ for $\boldsymbol{\pi}_\#\gamma$-a.e. y. We can find a $\boldsymbol{\pi}_\#\gamma$-negligible set $Y' \subset Y$ such that the equality holds for all $y \in Y\setminus Y'$ and all ψ in a countable dense set in $C^1_c(\mathbb{R})$. By density, the claimed property holds for all $y \in Y\setminus Y'$.

With a very similar argument one can prove a second claim, that $u_y(t) := u(y + t\mathbf{e}_j)$ is differentiable according to (4.2.2) with $X = \mathbb{R}$, $\gamma = \gamma_y$, for $\boldsymbol{\pi}_\#\gamma$-a.e. y, with $\partial^{\gamma_y} u_y(t) = \partial^\gamma_{\mathbf{e}_j} u(y + t\mathbf{e}_j)$. In fact, choose again a cylindrical function $\zeta(x), x \in X$, of the form $\zeta(\boldsymbol{\pi}(x))\psi(t)$, where $\boldsymbol{\pi}(x) = y \in Y$ and $\psi \in C^1_c(\mathbb{R})$, and disintegrating all the three terms of (4.2.2) with respect to $\boldsymbol{\pi}_\#\gamma$ we get, since $\partial_{\mathbf{e}_j}\zeta(x) = \zeta(y)\psi'(t)$,

$$\begin{aligned}\int_Y \zeta(y)\int_{\mathbb{R}} \psi'(t)u_y(t)\,d\gamma_y(t)\,d\boldsymbol{\pi}_\#\gamma(y) = &-\int_Y \zeta(y)\int_{\mathbb{R}} \partial^\gamma_{\mathbf{e}_j}u(y+t\mathbf{e}_j)\psi(t)\,d\gamma_y(t)\,d\boldsymbol{\pi}_\#\gamma(y)\\ &-\int_Y \zeta(y)\int_{\mathbb{R}} u_y(t)\psi(t)f_y(t)\,d\gamma_y(t)\,d\boldsymbol{\pi}_\#\gamma(y),\end{aligned}$$

so that, for $\boldsymbol{\pi}_\#\gamma$-a.e. $y \in Y$, we have

$$\int_{\mathbb{R}} \psi'(t)u_y(t)\,d\gamma_y(t) = -\int_{\mathbb{R}} \partial^\gamma_{\mathbf{e}_j}u(y+t\mathbf{e}_j)\psi(t)\,d\gamma_y(t) - \int_{\mathbb{R}} u_y(t)\psi(t)f_y(t)\,d\gamma_y(t),$$

which proves the second claim invoking the same density argument. Having proved the claims, the conclusion of the proof is standard: first the statement is proved for u_y, γ_y, and then using the disintegration of γ, it is extended to u, γ.

So, it remains to prove the chain rule formula in the case when $X = \mathbb{R}$, $\gamma = h\mathcal{L}^1$, with $h' = hg \in L^1(\mathbb{R})$. In this case we shall use use the fact that this property holds for the classical distributional derivative (see for instance [34, Chapter 4]), or [3, Theorem 3.99] for a more general result); we can read the integration by parts formula

$$\int_{\mathbb{R}} uh\zeta'\,dt = -\int_{\mathbb{R}} ugh\zeta\,dt - \int_{\mathbb{R}} \partial^\gamma uh\zeta\,dt \tag{4.2.5}$$

by saying that $v := uh \in W^{1,1}(\mathbb{R})$ and $h\partial^\gamma u = v' - uh'$. Since h is continuous it follows that $u = v/h \in W^{1,1}_{\rm loc}(\{h > 0\})$ and the classical product rule in Sobolev spaces gives $\partial^\gamma u = u'$ in $\{h > 0\}$. Conversely, if a bounded function w belongs to $W^{1,1}_{\rm loc}(\{h > 0\})$ and $w' \in L^1(\gamma)$, then $w \in W^{1,1}(\mathbb{R}, \gamma)$ and $\partial^\gamma w = w'$: indeed, under these assumptions (4.2.5) with $u = w$ holds when ζ has support contained in $\{h > 0\}$, and by approximation it holds for all ζ of the form $\tilde\zeta h/\sqrt{h^2+\varepsilon^2}$ with $\tilde\zeta \in C^1_c(\mathbb{R})$. Letting $\epsilon \to 0$ easily gives

$$\int_{\mathbb{R}} wh\tilde\zeta'\,dt = -\int_{\mathbb{R}} wgh\tilde\zeta\,dt - \int_{\mathbb{R}} \partial^\gamma wh\tilde\zeta\,dt$$

because the extra term

$$\int_{\mathbb{R}} wh\zeta\Big(\frac{h}{\sqrt{h^2+\varepsilon^2}}\Big)'\,dt$$

coming from the differentiation of $h/\sqrt{h^2+\varepsilon^2}$, can be estimated, up to the multiplicative constant $\sup|w\zeta|$, by

$$\frac{\epsilon^2 h|h'|}{\sqrt{h^2+\epsilon^2}^3} \le \frac{\epsilon^2|h'|}{h^2+\epsilon^2} \le |h'|$$

and tends to 0 pointwise.

Obviously $w = f(u)$ is locally Sobolev on $\{h > 0\}$ and $w' = f'(u)u'$ on $\{h > 0\}\setminus u^{-1}(\Sigma)$, and equal to 0 on $u^{-1}(\Sigma)$. See Proposition 3.92 and Theorem 3.99 in [3]. □

Theorem 4.2.4 (Partial derivatives) *Let $f : X \to \mathbb{R}$ be Lipschitz and assume that $\partial_v \gamma \ll \gamma$. Then*

$$\exists \lim_{t \downarrow 0} \frac{f(x+tv)-f(x)}{t} \qquad \text{for } \gamma\text{-a.e. } x. \tag{4.2.6}$$

Proof. The disintegration arguments of the previous proof can be repeated, so one can see that the conditional measures γ_y induced by the map $x \mapsto x - \langle x, v\rangle v$, indexed by $y \in \{v\}^\perp$, are absolutely continuous with respect to $\mathcal{L}^1$ for $\boldsymbol{\pi}_\# \gamma$-a.e. $y \in \{v\}^\perp$, where $\boldsymbol{\pi}$ is the orthogonal projection on $\{v\}^\perp$ (clearly here v plays the role of $\mathbf{e}_j$, so that $\{v\}^\perp$ corresponds to Y, in the notation of the proof of Theorem 4.2.3). Then, the existence $\mathcal{L}^1$-a.e. of the derivative of $t \mapsto f(y+tv)$ yields existence of the derivative γ_y-a.e. in X. We conclude that the limit (4.2.6) exists γ-a.e. in X. □

Definition 4.2.2 makes sense for $L^\infty(X,\gamma)$ functions. In order to treat the unbounded case, we will need a generalized definition of Sobolev spaces, based on truncation. For $u : X \to \mathbb{R}$ and $\alpha \geq 0$, define the α-truncate of u by

$$T_\alpha(u) := -\alpha \vee u \wedge \alpha. \tag{4.2.7}$$

Suppose that $T_n(u) \in W^{1,1}(X,\gamma)$ for every integer n. Thanks to Theorem 4.2.3, there holds $\nabla T_n u = 0$ γ-a.e. on $\{|u| > n\}$. Moreover,

$$\nabla T_n u = \nabla T_m u \qquad \gamma\text{-a.e. on } \{|u| < n\} \tag{4.2.8}$$

for $n < m$, since the two functions are equal on $\{|u| < n\}$. Hence we can define

$$\partial^\gamma_{\mathbf{e}_j} u := \partial^\gamma_{\mathbf{e}_j} T_n(u) \qquad \gamma\text{-a.e. on } \{|u| < n\}, \tag{4.2.9}$$

$$\nabla u := \nabla T_n(u) \qquad \gamma\text{-a.e. on } \{|u| < n\} \tag{4.2.10}$$

and this is a good definition, up to γ-negligible sets, because of (4.2.8) (and because we used only a countable set of truncation levels).

Definition 4.2.5 (Generalized Sobolev spaces) *Moreover, we say that a Borel map $u : X \to \mathbb{R}$ belongs to $GW^{1,p}(X,\gamma)$ if $T_\alpha(u) \in W^{1,p}(X,\gamma)$ for all $\alpha \geq 0$. The partial derivatives and the gradient of u are defined as in* (4.2.9) *and* (4.2.10).

Remark 4.2.6 Notice that we might equivalently require only $T_n(u) \in W^{1,p}(X,\gamma)$ for all integers n: this follows by applying the chain rule with $f = T_\alpha$ to the identity $T_\alpha = T_\alpha \circ T_n$, for $n > \alpha$. Similarly one can prove that any unbounded sequence of truncation levels would provide an equivalent definition.

4.3 Existence of optimal transport maps

We keep working under Assumption 3.1.6. As a matter of fact, existence of optimal maps simplifies considerably some proofs and constructions, although almost all arguments can be reproduced working with transport plans. The assumption will be satisfied (even if we don't ask that measures in $D(|\partial\phi|)$ vanish on Gaussian null sets) if $\phi(\mu)$ finite implies $\mu \ll \gamma$ and Assumption 4.2.1 holds. In fact, we have the following

Theorem 4.3.1 (Existence of optimal maps) *Assume that $\partial_{\mathbf{e}_j}\gamma \ll \gamma$ for all $j \geq 1$, μ, $\nu \in \mathscr{P}_2(X)$ and $\mu \ll \gamma$. Then there exists a unique optimal transport plan from μ to ν, and this plan is induced by a map.*

Proof. The proof is very similar to the one of Theorem 2.1.7: one reduces to the case when ν has a bounded support and finds an optimal plan β and a maximizing pair (φ, ψ) of Kantorovich potentials, so that $\varphi(x) + \psi(y) \leq |x-y|^2$ and equality holds on suppβ; since

$$\varphi(x) = \inf_{y \in \operatorname{supp}\nu} |x-y|^2 - \psi(y)$$

we have that φ is Lipschitz on bounded sets. Then, by applying a local version of Theorem 4.2.4, we find a γ-negligible set $N \subset X$ such that $\partial_{\mathbf{e}_j}\varphi$ exists at all points of $X \setminus N$ for all $j \geq 1$. Since $|x'-y|^2 - \varphi(x')$ attains its minimum at $x' = x$ (equal to $-\psi(y)$) for points $(x,y) \in$ suppβ, if $x \notin N$ partial differentiation gives

$$2\langle x-y, \mathbf{e}_j\rangle = \partial_{\mathbf{e}_j}\varphi(x), \qquad \forall j \geq 1.$$

Since $\beta(N \times X) = \mu(N) = 0$, this proves that y is uniquely determined by x β-a.e., hence β is concentrated on a graph. This provides the optimal transport map. Since any optimal plan β is concentrated on the graph of a map, the optimal map is unique (otherwise a combination of two optimal maps would produce an optimal plan not concentrated on a graph) and, as a consequence, β is unique as well. □

Lemma 4.3.2 (Stability of optimal maps) *Let $\mu, \nu \in \mathscr{P}_2(X)$ be such that $\Gamma_0(\mu,\nu)$ contains a unique optimal plan induced by a map $\mathbf{r} \in L^2(X,\mu;X)$. Let $\nu_n \to \nu$ and let $\mathbf{r}_n \in L^2(X,\mu;X)$ be optimal transport maps from μ to ν_n. Then $\mathbf{r}_n \to \mathbf{r}$ in $L^2(X,\mu;X)$.*

Proof. Let $\varphi : X \times X \to \mathbb{R}$ be a continuous function with 2-growth. Then, by means of Proposition 1.1.4, we find

$$\lim_{n\to\infty}\int_X \varphi(x, \mathbf{r}_n(x))\,d\mu(x) = \lim_{n\to\infty}\int_{X\times X}\varphi(x,y)\,d((\mathbf{I},\mathbf{r}_n)_\#\mu)(x,y) = \int_X \varphi(x,\mathbf{r}(x))\,d\mu(x),$$

since $(\mathbf{I},\mathbf{r}_n)_\#\mu$ are optimal plans converging (thanks to Lemma 2.2.6) to the unique element of $\Gamma_0(\mu,\nu)$, namely $(\mathbf{I},\mathbf{r})_\#\mu$. If $\varphi(x,y) = |y|^2$ we see that $|\mathbf{r}_n|_{L^2(X,\mu;X)} \to |\mathbf{r}|_{L^2(X,\mu;X)}$, and then let $\widetilde{\mathbf{r}}$ be a weak $L^2(X,\mu;X)$ limit of $\mathbf{r}_n$. Choose now $\varphi(x_1,x_2) = \zeta(x_1)\langle z, x_2\rangle$, with ζ continuous and bounded and $z \in X$, getting

$$\lim_{n\to\infty}\int_X \zeta(x)\langle z, \mathbf{r}_n(x)\rangle\,d\mu = \int_X \zeta(x)\langle z, \mathbf{r}(x)\rangle\,d\mu(x).$$

On the other hand, by weak convergence of $\mathbf{r}_n$ to $\widetilde{\mathbf{r}}$ we have

$$\lim_{n\to\infty}\int_X \zeta(x)\langle z, \mathbf{r}_n(x)\rangle\,d\mu = \int_X \zeta(x)\langle z, \widetilde{\mathbf{r}}(x)\rangle\,d\mu.$$

Hence $\langle z, \widetilde{\mathbf{r}}(x)\rangle = \langle z, \mathbf{r}(x)\rangle$ for μ-a.e. $x \in X$, for all $z \in X$, yielding $\mathbf{r} = \widetilde{\mathbf{r}}$ μ-a.e. in X. Weak convergence and convergence of norms in L^2 give strong convergence. □

The gradient flow μ_t of functional ϕ is the limit of the implicit Euler discrete scheme: given $\mu^0 \in D(\phi)$, one constructs a sequence $(\mu_\tau^k) \subset \mathscr{P}_2(X)$, with $\mu_\tau^0 = \mu^0$, whose k-th element is found minimizing the functional (3.2.1). For $t > 0$ and $k > 0$, we can define a discrete gradient flow $\bar{\mu}_\tau(t)$ by (3.2.4). In fact, by the theory of Chapter 3, we know that, $\bar{\mu}_\tau(t) \to \mu_t$ for all $t \geq 0$, where μ_t is the gradient flow. Let us focus the attention on the discrete problem. The following approximation result of the minimal selection in terms of vectors taking the form of (3.4.1) will be useful in the sequel. It extends the result of Lemma 3.2.3.

Lemma 4.3.3 *Let $\phi : \mathscr{P}_2(X) \to (-\infty, +\infty]$ be a convex functional along geodesics and let $\mu^0 \in D(|\partial\phi|)$. If μ_τ is a minimizer of $\Phi_\tau(\cdot, \mu^0)$ and $\boldsymbol{\omega}_\tau$ is the vector introduced by* (3.4.1)*, then there exist $\tau_n \downarrow 0$ such that, as $n \to \infty$, $\mu_{\tau_n} \to \mu^0$, $\phi(\mu_{\tau_n}) \to \phi(\mu^0)$ and, more precisely,*

$$|\partial\phi(\mu^0)|^2 = \lim_{n\to\infty} \frac{W_2^2(\mu_\tau, \mu^0)}{\tau_n^2} = \lim_{n\to\infty} \frac{\phi(\mu^0) - \phi(\mu_{\tau_n})}{\tau_n} = \lim_{n\to\infty} \|\boldsymbol{\omega}_{\tau_n}\|^2_{L^2(X,\,\mu_{\tau_n};\,X)} . \qquad (4.3.1)$$

Moreover, $\boldsymbol{\omega}_{\tau_n} \in L^2(X, \mu_{\tau_n}; X)$ converge, strongly in the sense of Definition 1.5.3*, to the unique vector $\partial^0\phi(\mu^0)$ with minimal norm in $\partial\phi(\mu^0)$.*

Proof. The first three equalities of (4.3.1) follow from Lemma 3.2.3. Moreover, notice that $\boldsymbol{\omega}_{\tau_n}$, being defined as $(\mathbf{t}_{\mu_{\tau_n}}^{\mu^0} - \mathbf{I})/\tau$, is a subdifferential by Theorem 3.5.1, and, as remarked in the proof the same theorem, since $\mathbf{t}_{\mu_{\tau_n}}^{\mu^0}$ is the optimal transport map between μ_{τ_n} and μ^0, satisfies

$$\|\boldsymbol{\omega}_{\tau_n}\|^2_{L^2(X,\,\mu_{\tau_n};\,X)} = \int_X |\boldsymbol{\omega}_{\tau_n}(x)|^2 \, d\mu_{\tau_n}(x) = \frac{W_2^2(\mu_{\tau_n}, \mu^0)}{\tau_n^2},$$

so that (4.3.1) holds. The strong convergence in the sense of Definition 1.5.3 is itself a consequence of Theorem 3.5.1. □

4.4 The internal energy functional $\mathscr{F}$

Given a Borel probability measure γ on $\mathbb{R}^d$, we define the finite-dimensional internal energy functional relative to γ as follows:

$$\mathscr{F}_d(\mu|\gamma) = \begin{cases} \displaystyle\int_{\mathbb{R}^d} F\left(\frac{d\mu}{d\gamma}\right) d\gamma & \text{if } \mu \ll \gamma, \\ +\infty & \text{otherwise} . \end{cases}$$

The definition can be extended easily to the case of a Borel probability measure γ in an infinite-dimensional Hilbert space X:

$$\mathscr{F}(\mu|\gamma) = \begin{cases} \displaystyle\int_X F\left(\frac{d\mu}{d\gamma}\right) d\gamma & \text{if } \mu \ll \gamma, \\ +\infty & \text{otherwise} . \end{cases}$$

Assumption 4.4.1 *We consider the following assumptions on the integrand* $F : [0, +\infty) \to (-\infty, +\infty]$*:*

i) F *is strictly convex;*

ii) the map $s \mapsto e^s F(e^{-s})$ *is convex and nonincreasing in* $\mathbb{R}$*;*

iii) $F(0) = 0$*;*

iv) F *has a superlinear growth at infinity.*

Condition *ii)* is needed for the geodesical convexity of $\mathscr{F}$, and in fact it has been introduced in [4] as a dimension-free extension of the one introduced by McCann (see [53]) for the d-dimensional case, namely

$$x \mapsto x^d F(x^{-d}) \text{ is convex and nonincreasing in } (0, +\infty). \tag{4.4.1}$$

Indeed, it can be shown that *ii)* implies (4.4.1). It is convenient to introduce the continuous function

$$L_F(x) := xF'_+(x) - F(x), \tag{4.4.2}$$

where F'_+ denotes the right derivative. In fact, we will write the velocity vector field of the gradient flow of $\mathscr{F}_d$ and $\mathscr{F}$ in terms of L_F, which will indeed be the same function L of equation (4.1.3). Notice also that the monotonicity condition in *(ii)* is equivalent to $xL'_F(x) - L_F(x) \geq 0$, while the convexity condition yields

$$e^s F(e^{-s}) - F'(e^{-s}) + e^{-s} F''(e^{-s}) \geq 0,$$

which implies convexity of F.

Let us introduce (see [4, Lemma 9.4.4]) the following dual representation of $\mathscr{F}$:

$$\mathscr{F}(\mu|\gamma) = \sup\left\{ \int_X g(x) d\mu(x) - \int_X F^*(g(x)) d\gamma(x) : g \in C_b^0(X) \right\}, \tag{4.4.3}$$

where F^* denotes the Fenchel conjugate of F. We notice from (4.4.3) that $\mathscr{F}$ is sequentially l.s.c. with respect to the weak convergence. For μ, $\nu \in \mathscr{P}_2(X)$, we also introduce the notation

$$\Phi_\tau^{\mathscr{F}}(\nu, \mu) := \mathscr{F}(\nu|\gamma) + \frac{1}{2\tau} W_2^2(\nu, \mu). \tag{4.4.4}$$

The typical example of function F one can consider is the n-th power:

$$F(s) = \frac{x^n}{n-1}, \quad n > 1, \tag{4.4.5}$$

with $L_F(x) = x^n$. Another important example is $F(x) = x \log x$, corresponding to the relative entropy functional (see Remark 4.5.2 below), whose gradient flow is a linear Fokker-Planck equation (see [43] and the infinite-dimensional theory in [9]).

Geodesical convexity of $\mathscr{F}$

In this subsection we recall some results on the convexity properties of $\mathscr{F}$.

Definition 4.4.2 (Log-concavity) *A probability measure on X is said to be log-concave if, for any couple of open sets A, B in X, there holds*

$$\log \gamma((1-t)A + tB) \geq (1-t)\log\gamma(A) + t\log\gamma(B). \tag{4.4.6}$$

If $X = \mathbb{R}^d$ and γ is non-degenerate (i.e. it is not supported in a proper subspace of X), then Borell proved (see also [4, Theorem 9.4.10]) that γ is log-concave if and only if $\gamma = e^{-V}\mathscr{L}^d$ for some lower semicontinuous and convex function $V : \mathbb{R}^d \to (-\infty, +\infty]$ whose domain has nonempty interior.

For the internal energy functional relative to γ, convexity along geodesics is strictly related to the log-concavity of γ, as shown by the following result (see [4, Theorem 9.4.12]).

Theorem 4.4.3 *Let F be satisfying* Assumption 4.4.1 *ii)-iii)-iv), and suppose that γ is log-concave. Then $\mathscr{F}(\cdot|\gamma)$ is strongly convex in $\mathscr{P}_2(X)$.*

Remark 4.4.4 Let $\gamma \in \mathscr{P}(X)$ be log-concave, let $\mu \in \mathscr{P}_2(X)$ and consider the constrained minimization problem

$$\min_{\nu \leq M\gamma} \Phi_\tau^{\mathscr{F}}(\nu, \mu).$$

Then this problem admits a unique minimizer, as the unconstrained one. In fact, the functional

$$\mathscr{F}^M(\mu|\gamma) := \begin{cases} \displaystyle\int_X F^M\left(\frac{d\mu}{d\gamma}\right) d\gamma & \text{if } \mu \ll \gamma, \\ +\infty & \text{otherwise,} \end{cases} \tag{4.4.7}$$

where

$$F^M(z) := \begin{cases} +\infty & \text{if } z > M, \\ F(z) & \text{otherwise,} \end{cases}$$

trivially satisfies the hypotheses of Theorem 4.4.3, so it is strongly convex and we can apply Proposition 3.5.6 with $\phi = \mathscr{F}^M$.

Discrete minimizers of $\mathscr{F}$ in the bounded case

The following result extends the one of [1, §2.1] to the infinite dimensional case, basically with the same proof.

Lemma 4.4.5 *Let F satisfy* Assumption 4.4.1 *and suppose that γ is log-concave. Let $\mu = \rho\gamma \in \mathscr{P}_2(X)$, with $\rho \leq M$ γ-a.e. in X. Then there exists a unique minimizer μ_τ of $\Phi_\tau^{\mathscr{F}}(\cdot, \mu)$, and $\mu_\tau \leq M\gamma$.*

Proof. We assume without loss of generality that M is a point of differentiability for F. As a first step, we consider the problem of minimizing $\Phi_\tau^{\mathscr{F}}(\cdot,\mu)$ under the constraint $\nu \le M'\gamma$, where $M' \ge M$. In view of Remark 4.4.4, we know that in this case there exists a unique minimizer $\overline{\mu}_\tau = \overline{\rho}_\tau\gamma \le M'\gamma$.

Let β denote the optimal transport plan between μ and $\overline{\mu}_\tau$. Suppose by contradiction that $\overline{\rho}_\tau > M$ on some Borel set $\Omega \subset X$ with $\gamma(\Omega) > 0$ and let Ω^c be the complement of Ω in X.

Now let $\beta_\Omega = \chi_{\Omega^c\times\Omega}\beta$. It is clear that $\pi^1_\#\beta_\Omega \le \mu$ and $\pi^2_\#\beta_\Omega \le \overline{\mu}_\tau$. Then, letting $\tilde{\rho}$ and $\tilde{\rho}_\tau$ be the densities with respect to γ of the first and second marginal of β_Ω, we have

$$\tilde{\rho} \le \rho \quad \text{and} \quad \tilde{\rho}_\tau \le \overline{\rho}_\tau. \tag{4.4.8}$$

Moreover, the following properties are easily seen to hold γ-a.e.:

$$\tilde{\rho} \le M, \qquad \tilde{\rho} = 0 \ \text{ on } \Omega, \qquad \tilde{\rho}_\tau = 0 \ \text{ on } \Omega^c. \tag{4.4.9}$$

Let us introduce the competitor of μ_τ as

$$\rho^\varepsilon_\tau\gamma := (\overline{\rho}_\tau + \varepsilon(\tilde{\rho} - \tilde{\rho}_\tau))\,\gamma. \tag{4.4.10}$$

By the definition of $\tilde{\rho}$ and $\tilde{\rho}_\tau$ it is immediate to check that $\int_X \tilde{\rho}\,d\gamma = \int_X \tilde{\rho}_\tau\,d\gamma = \beta(\Omega^c\times\Omega)$. As a consequence $\rho^\varepsilon_\tau\gamma \in \mathscr{P}_2(X)$. Moreover, since $\overline{\rho}_\tau > M$ γ-a.e. in Ω, making use of (4.4.8) and (4.4.9) we obtain, for small enough ε,

$$\rho^\varepsilon_\tau = \overline{\rho}_\tau - \varepsilon\tilde{\rho}_\tau > 0 \qquad \gamma\text{-a.e. on } \Omega. \tag{4.4.11}$$

Then, denoting by F'_- and F'_+ respectively the left and right derivative of F, thanks to the convexity of F we have, for small enough ε,

$$\begin{aligned}
\int_X (F(\rho^\varepsilon_\tau) - F(\overline{\rho}_\tau))\,d\gamma &\le \int_{\Omega^c} (F(\overline{\rho}_\tau + \varepsilon\tilde{\rho}) - F(\overline{\rho}_\tau))\,d\gamma + \int_\Omega (F(\overline{\rho}_\tau - \varepsilon\tilde{\rho}_\tau) - F(\overline{\rho}_\tau))\,d\gamma \\
&\le \varepsilon\int_{\Omega^c} F'_+(\overline{\rho}_\tau + \varepsilon\tilde{\rho})\tilde{\rho}\,d\gamma - \varepsilon\int_\Omega F'_-(\overline{\rho}_\tau - \varepsilon\tilde{\rho}_\tau)\tilde{\rho}_\tau\,d\gamma \\
&\le \varepsilon\int_{\Omega^c} F'_+(M + \varepsilon\tilde{\rho})\tilde{\rho}\,d\gamma - \varepsilon\int_\Omega F'_-(M - \varepsilon\tilde{\rho}_\tau)\tilde{\rho}_\tau\,d\gamma \\
&= \varepsilon\int_{X\times X} \left[F'_+(M + \varepsilon\tilde{\rho}(x)) - F'_-(M - \varepsilon\tilde{\rho}_\tau(y))\right]\,d\beta_\Omega(x,y) \\
&= \varepsilon\int_{X\times X} o(1)\,d\beta_\Omega(x,y).
\end{aligned}$$

Since $\tilde{\rho}$ and $\tilde{\rho}_\tau$ are bounded above γ-a.e. by M', we conclude that

$$\int_X (F(\rho^\varepsilon_\tau) - F(\overline{\rho}_\tau))\,d\gamma \le o(\varepsilon). \tag{4.4.12}$$

On the other hand, let $\mathbf{t} : X \times X \to X \times X$ be defined by $\mathbf{t}(x, y) = (x, x)$, and let

$$\beta_\varepsilon = \beta - \varepsilon\beta_\Omega + \varepsilon\mathbf{t}_\#\beta_\Omega.$$

By the composition rule of the push forward we have $\pi^2_\#\mathbf{t}_\#\beta_\Omega = (\pi^2 \circ \mathbf{t})_\#\beta_\Omega = \pi^1_\#\beta_\Omega$, so that the second marginal of $\mathbf{t}_\#\beta_\Omega$ is equal to the first marginal of β_Ω, namely $\tilde{\rho}$; analogously the first marginal of $\mathbf{t}_\#\beta_\Omega$ coincides with the first marginal of β_Ω. Hence it is clear that $\beta_\varepsilon \in \Gamma(\mu, \rho^\varepsilon_\tau\gamma)$. So we can estimate

$$W_2^2(\rho^\varepsilon_\tau\gamma, \mu) - W_2^2(\rho_\tau\gamma, \mu) \leq \int_{X\times X} |x - y|^2 \, d(\beta_\varepsilon - \beta)(x, y) = -\varepsilon \int_{\Omega^c\times\Omega} |x - y|^2 \, d\beta(x, y). \tag{4.4.13}$$

Together with (4.4.12), this gives

$$\Phi^{\mathscr{F}}_\tau(\rho^\varepsilon_\tau\gamma, \mu) - \Phi^{\mathscr{F}}_\tau(\overline{\rho}_\tau\gamma, \mu) \leq -\frac{\varepsilon}{2\tau}\int_{\Omega^c\times\Omega} |x - y|^2 \, d\beta(x, y) + o(\varepsilon). \tag{4.4.14}$$

But consider that

$$\beta(\Omega \times \Omega) \leq \beta(\Omega \times X) = \int_X \chi_\Omega(x) \, d(\pi^1_\#\beta)(x) = \int_\Omega \rho(x) \, d\gamma(x) \leq M\gamma(\Omega). \tag{4.4.15}$$

This forces $\beta(\Omega^c \times \Omega)$ to be strictly positive, otherwise

$$\beta(\Omega \times \Omega) = \beta(X \times \Omega) = \overline{\mu}_\tau(\Omega) = \int_\Omega \overline{\rho}_\tau(x) \, d\gamma(x) > M\gamma(\Omega)$$

against (4.4.15). Back to (4.4.14), if ε is chosen small enough, we contradict the minimality of $\overline{\mu}_\tau = \overline{\rho}_\tau\gamma$. We have proved that $\overline{\rho}_\tau \leq M$, independently of the initial choice of M'.

Since these properties hold for all $M' > M$, it turns out that the minimizer is independent of M', hence $\overline{\mu}_\tau$ is a minimizer under the constraint $\nu = \rho\gamma$ with $\rho \in L^\infty(\gamma)$. Then, a simple truncation argument provides the minimality of $\overline{\mu}_\tau$ in the unconstrained problem. □

4.5 The finite-dimensional case

A key ingredient of our analysis will be the finite-dimensional framework, which has been studied in detail in [4, Chapter 10]. We now recall the main result therein (see [4, Theorem 10.4.9]). We will make use, in the rest of the chapter, of the notation $|\partial\mathscr{F}|(\mu|\gamma)$ to indicate, for fixed γ, the slope of $\mathscr{F}(\cdot|\gamma)$ at the point μ, and similarly for $\mathscr{F}_d$ and for the respective subdifferentials.

Theorem 4.5.1 *Let $\gamma = e^{-V}\mathcal{L}^d$ be a non-degenerate log-concave probability measure on $\mathbb{R}^d$, let Ω be the nonempty interior of $D(V)$ and consider the functional $\mathscr{F}_d(\cdot|\gamma)$ and $\mu = \rho\gamma \in D(\mathscr{F}_d)$. Then $\rho \in D(|\partial\mathscr{F}_d|)$ if and only if*

$$L_F \circ \rho \in W^{1,1}(\Omega) \qquad and \qquad \frac{\nabla(L_F \circ \rho)}{\rho} \in L^2(\mathbb{R}^d, \mu). \tag{4.5.1}$$

If these conditions hold, $\frac{\nabla(L_F\circ\rho)}{\rho}$ realizes the minimal selection in $|\partial\mathscr{F}_d|$ at the point μ, so that

$$\frac{\nabla(L_F \circ \rho)}{\rho} = \partial^0\mathscr{F}_d(\mu|\gamma) \qquad and \qquad \left\|\frac{\nabla(L_F \circ \rho)}{\rho}\right\|_{L^2(\mathbb{R}^d,\,\mu)} = |\partial\mathscr{F}_d|(\mu|\gamma). \tag{4.5.2}$$

Remark 4.5.2 In the case $X = \mathbb{R}^d$, let $\gamma = e^{-V}\mathcal{L}^d$, where V is a convex l.s.c. potential, and $\mu_t = u_t\mathcal{L}^d$. As a consequence we have $\rho_t = u_t e^V$ and (4.1.3) becomes

$$\partial_t u_t - \nabla\cdot(\nabla(L_F \circ u_t) + u_t\nabla V) = 0. \tag{4.5.3}$$

In (4.5.3) we recognize different PDEs. In particular, if $V = 0$ and $L_F(x) = x^m$, $m > 1$ (which corresponds to $F = (m-1)^{-1}x^m$) we obtain the porous media equations. If $F(x) = x\log x$, then $\mathscr{F}_d$ is the well known entropy functional

$$\mathscr{H}_d(\mu|\gamma) = \begin{cases} \displaystyle\int_X \left(\frac{d\mu}{d\gamma}\right)\log\left(\frac{d\mu}{d\gamma}\right)d\gamma & \text{if } \mu \ll \gamma, \\ +\infty & \text{otherwise}. \end{cases}$$

In this case $L_F(x) = x$ and (4.5.3) becomes the linear Fokker Planck equation with potential V:

$$\partial_t u_t - \Delta u_t - \nabla\cdot(u_t\nabla V) = 0. \tag{4.5.4}$$

See [9] for a detailed comparison between the different approaches to (4.5.4) in infinite dimensions.

4.6 Γ-convergence results

For the characterization of the subdifferential of $\mathscr{F}$, we will perform finite dimensional approximations, and we need a Γ-convergence result. First of all, if ϕ_n is a sequence of functionals, we introduce the notation

$$\Phi^n_\tau(\cdot,\mu) := \phi_n(\cdot) + \frac{1}{2\tau}W_2^2(\cdot,\mu). \tag{4.6.1}$$

Next we define the Γ-convergence.

Definition 4.6.1 (Γ-convergence) *We say that $\phi_n : \mathscr{P}_2(X) \to [-\infty,+\infty]$ $\Gamma(\mathscr{P}_2(X))$-converge to ϕ if* *i) for any sequence* $(\mu_n) \subset \mathscr{P}_2(X)$ *weakly convergent to* μ*, there holds*

$$\phi(\mu) \leq \liminf_{n\to\infty} \phi_n(\mu_n); \tag{4.6.2}$$

ii) for any $\mu \in \mathscr{P}_2(X)$ *there exists* $(\mu_n) \subset \mathscr{P}_2(X)$ *converging to* μ *in* $\mathscr{P}_2(X)$ *such that*

$$\lim_{n\to\infty} \phi_n(\mu_n) = \phi(\mu). \tag{4.6.3}$$

Γ-convergence guarantees the convergence of minimizers to minimizers, as in the next lemma.

Lemma 4.6.2 *Let* $\phi_h : \mathscr{P}_2(X) \to (-\infty, +\infty]$ *be geodesically convex functionals satisfying* Assumption 3.1.1 *and* $\Gamma(\mathscr{P}_2(X))$*-convergent to* ϕ*, still satisfying* Assumption 3.1.1*. Assume also that for all* $M > 0$ *the set*

$$\bigcup_{h=1}^{\infty} \{\mu \in \mathscr{P}_2(X) : \ \phi_h(\mu) \leq M\} \tag{4.6.4}$$

is relatively compact in the weak topology of $\mathscr{P}(X)$*. Let* $\mu^h \to \mu$*. Let* (μ^h_τ) *denote, for* τ *fixed, a family of minimizers of functionals* $\Phi^h_\tau(\cdot, \mu^h)$*, defined by* (4.6.1)*. Then* μ^h_τ *have limit points in* $\mathscr{P}_2(X)$ *and* $\boldsymbol{\omega}^h_\tau \in \partial\phi_h(\mu^h_\tau)$*, constructed in* Lemma 3.4.1*, have strong limit points in the sense of* Definition 1.5.3*. If* (h_n) *is any subsequence along which we have convergence, and* $\mu_\tau, \boldsymbol{\omega}_\tau$ *are the limits, then* μ_τ *is a minimizer of* $\Phi_\tau(\cdot, \mu)$ *and* $\boldsymbol{\omega}_\tau$ *belongs to* $\partial\phi(\mu_\tau)$*. Moreover*

$$\phi_{h_n}(\mu^{h_n}_\tau) \to \phi(\mu_\tau).$$

Proof. Let $\tau > 0$ be fixed during all the proof. Let $\nu \in \mathscr{P}_2(X)$, and let $\nu^h \to \nu$ be such that $\phi_h(\nu^h) \to \phi(\nu)$. We can find such a sequence thanks to (4.6.3). Since μ^h_τ minimizes $\Phi^h_\tau(\cdot, \mu^h)$, we have immediately

$$\phi_h(\mu^h_\tau) + \frac{1}{2\tau} W_2^2(\mu^h_\tau, \mu^h) \leq \phi_h(\nu^h) + \frac{1}{2\tau} W_2^2(\nu^h, \mu^h). \tag{4.6.5}$$

Consider the second member, as $h \to \infty$: the first term goes to $\phi(\nu)$. The second converges to $W_2^2(\nu, \mu)/(2\tau)$, by the continuity properties of the Wasserstein distance with respect to the convergence with moments (indeed, the liminf inequality follows from Lemma 2.2.6, while the limsup one follows from Lemma 2.2.7 and from the triangle inequality which entails $W_2(\nu^h, \mu^h) \leq W_2(\nu^h, \nu) + W_2(\nu, \mu) + W_2(\mu, \mu^h)$). Hence, $\phi_h(\mu^h_\tau)$ is uniformly bounded in h for τ fixed, so that μ^h_τ belongs to the set in (4.6.4) for some positive M, yielding the compactness of the family (μ^h_τ) in $\mathscr{P}(X)$. Now let (h_n) be a sequence along which we have convergence and let μ_τ be the corresponding limit, so that $\mu^{h_n}_\tau \rightharpoonup \mu_\tau$ as $n \to \infty$. We have from (4.6.5), and taking advantage again of the semicontinuity of the Wasserstein distance given by Lemma 2.2.6, and of (4.6.2), we get

$$\begin{aligned}\phi(\mu_\tau)+\frac{1}{2\tau}W_2^2(\mu_\tau,\mu)&\le\limsup_{n\to\infty}\left(\phi_{h_n}(\mu_\tau^{h_n})+\frac{1}{2\tau}W_2^2(\mu_\tau^{h_n},\mu^{h_n})\right)\\&\le\limsup_{n\to\infty}\left(\phi_{h_n}(\nu^{h_n})+\frac{1}{2\tau}W_2^2(\nu^{h_n},\mu^{h_n})\right)\\&=\phi(\nu)+\frac{1}{2\tau}W_2^2(\nu,\mu),\end{aligned}$$

which shows, by the arbitrariness of ν, that μ_τ is a minimizer for $\Phi_\tau(\cdot,\mu)$. These inequalities are of course equalities if we choose $\nu=\mu_\tau$, then

$$\lim_{n\to\infty}\left(\phi_{h_n}(\mu_\tau^{h_n})+\frac{W_2^2(\mu_\tau^{h_n},\mu^{h_n})}{\tau}\right)=\phi(\mu_\tau)+\frac{W_2^2(\mu_\tau,\mu)}{\tau}.$$

The two terms are separately l.s.c., hence, as $n\to\infty$,

$$\phi_{h_n}(\mu_\tau^{h_n})\to\phi(\mu_\tau)\quad\text{and}\quad W_2(\mu_\tau^{h_n},\mu^{h_n})\to W_2(\mu_\tau,\mu).\tag{4.6.6}$$

Notice that, by (4.6.6), and since $\mu^{h_n}\to\mu$, we have

$$\begin{aligned}\limsup_{n\to\infty}\int_X|x|^2\,d\mu_\tau^{h_n}(x)&\le2\limsup_{n\to\infty}\int_X|\mathbf{t}_{\mu_\tau^{h_n}}^{\mu^{h_n}}(x)-x|^2\,d\mu_\tau^{h_n}+2\limsup_{n\to\infty}\int_X|\mathbf{t}_{\mu_\tau^{h_n}}^{\mu^{h_n}}(x)|^2\,d\mu_\tau^{h_n}\\&=2\limsup_{n\to\infty}\int_X|\mathbf{t}_{\mu_\tau^{h_n}}^{\mu^{h_n}}(x)-x|^2\,d\mu_\tau^{h_n}+2\limsup_{n\to\infty}\int_X|x|^2\,d\mu^{h_n}\\&=2W_2^2(\mu_\tau,\mu)+2\int_X|x|^2\,d\mu(x).\end{aligned}$$

This shows that the sequence $(\mu_\tau^{h_n})$ has uniformly integrable 2-moments, and then, by Proposition 1.1.4, we have in fact $\mu_\tau^{h_n}\to\mu_\tau$ in $\mathscr{P}_2(X)$. Now let $\boldsymbol{\omega}_\tau^{h_n}\in\partial\phi_{h_n}(\mu_\tau^{h_n})$ be constructed as in Lemma 3.4.1, that is, $\boldsymbol{\omega}_\tau^{h_n}=(\mathbf{t}_{\mu_\tau^{h_n}}^{\mu^{h_n}}-\mathbf{I})/\tau$. Thanks to (4.6.6), we have

$$\lim_{n\to\infty}\int_X|\boldsymbol{\omega}_\tau^{h_n}|^2\,d\mu_\tau^{h_n}=\lim_{n\to\infty}\frac{W_2^2(\mu_\tau^{h_n},\mu^{h_n})}{\tau^2}=\frac{W_2^2(\mu_\tau,\mu)}{\tau^2}.\tag{4.6.7}$$

But μ_τ is a minimizer of $\Phi_\tau(\cdot,\mu)$, so that thanks to Lemma 3.4.1 it belongs to $D(|\partial\phi|)$. Then, since we are working under Assumption 3.1.6, there exists a unique optimal transport map $\mathbf{t}_{\mu_\tau}^{\mu}$ between μ_τ and μ. Invoking Lemma 3.4.1 again, we see that $\boldsymbol{\omega}_\tau=(\mathbf{t}_{\mu_\tau}^{\mu}-\mathbf{I})/\tau$ belongs to $\partial\phi(\mu_\tau)$. Since $W_2^2(\mu_\tau,\mu)/\tau^2=\int_X|\boldsymbol{\omega}_\tau|^2\,d\mu_\tau$, from (4.6.7) we conclude

$$\lim_{n\to\infty}\int_X|\boldsymbol{\omega}_\tau^{h_n}|^2\,d\mu_\tau^{h_n}=\int_X|\boldsymbol{\omega}_\tau|^2\,d\mu_\tau,$$

that is, the sequence of subdifferentials $(\boldsymbol{\omega}_\tau^{h_n})$ converges strongly, in the sense of Definition 1.5.3, to $\boldsymbol{\omega}_\tau\in\partial\phi(\mu_\tau)$.

□

It is clear from Lemma 4.3.3 that there exist μ_{τ_n}, minimizers of $\Phi_{\tau_n}(\cdot,\mu)$, such that the respective subdifferentials converge to $\partial^0\phi(\mu)$. With the next result we want to show that the approximating $\boldsymbol{\omega}_n$ can also be chosen to be subdifferentials of functionals ϕ_n, if ϕ_n is Γ-convergent to ϕ.

Theorem 4.6.3 *Let $\phi_n, \phi : \mathscr{P}_2(X) \to (-\infty,+\infty]$ be as in* Lemma 4.6.2. *Then, for every $\mu \in D(|\partial\phi|)$ there exist a subsequence $n(m)$, $\mu_{n(m)}$ converging to μ in $\mathscr{P}_2(X)$ and subdifferentials $\boldsymbol{\omega}_{n(m)} \in \partial\phi_{n(m)}(\mu_{n(m)})$ such that*

$$\boldsymbol{\omega}_{n(m)} \to \partial^0\phi(\mu) \in L^2(X,\mu;X) \quad \textit{strongly in } L^2 \textit{ as in } \text{Definition 1.5.3} \tag{4.6.8}$$

and

$$\lim_{m\to\infty} \phi_{n(m)}(\mu_{n(m)}) = \phi(\mu). \tag{4.6.9}$$

In particular, since $|\partial\phi|(\mu)$ is the $L^2(X,\mu;X)$ norm of the minimal selection in $\partial\phi(\mu)$, this means that

$$\limsup_{m\to\infty} \int_X \|\boldsymbol{\omega}_{n(m)}\|^2 d\mu_{n(m)} \le |\partial\phi|^2(\mu). \tag{4.6.10}$$

Proof. We construct the approximating sequence in the following way. Let $\mu^h \to \mu$ in $\mathscr{P}_2(X)$ with $\phi_h(\mu^h) \to \phi(\mu)$ (such a sequence exists by Γ-convergence). Let μ^h_τ be a minimizer of

$$\Phi^h_\tau(\cdot,\mu^h) := \phi_h(\cdot) + \frac{1}{2\tau}W_2^2(\cdot,\mu^h).$$

Let moreover $\boldsymbol{\omega}^h_\tau$ be constructed as in (3.4.1). We will show that there is a subsequence of the family $\{\boldsymbol{\omega}^h_\tau : h \in \mathbb{N}, \tau > 0\}$ such that (4.6.8) holds. First, for fixed τ, we know from Lemma 4.6.2 that there is a subsequence $\mu^{h_n}_\tau$ converging in $\mathscr{P}_2(X)$ to μ_τ, where μ_τ minimizes $\Phi_\tau(\cdot,\mu)$. Moreover, the corresponding sequence $\boldsymbol{\omega}^{h_n}_\tau$ converge to $\boldsymbol{\omega}_\tau \in \partial\phi(\mu_\tau)$ in the sense of Definition 1.5.3. Hence, given $\varepsilon > 0$, for n large enough we have

$$\left|\int_X |\boldsymbol{\omega}^{h_n}_\tau|^2\, d\mu^{h_n}_\tau - \int_X |\boldsymbol{\omega}_\tau|^2\, d\mu_\tau\right| < \frac{\varepsilon}{2}. \tag{4.6.11}$$

and (taking Lemma 4.6.2 into account)

$$\left|\phi_{h_n}(\mu^{h_n}_\tau) - \phi(\mu_\tau)\right| < \frac{\varepsilon}{2}. \tag{4.6.12}$$

On the other hand, we know from Lemma 4.3.3 that there exists an infinitesimal sequence (τ_m) such that

$$\lim_{m\to\infty}\left|\int_X |\boldsymbol{\omega}_{\tau_m}|^2\, d\mu_{\tau_m} - \int_X |\boldsymbol{\omega}|^2\, d\mu\right| = 0 \tag{4.6.13}$$

and

$$\lim_{m\to\infty} |\phi(\mu_{\tau_m}) - \phi(\mu)| = 0. \tag{4.6.14}$$

Now, with $\tau = \tau_m$ and $\varepsilon = 1/m$ we can suitably choose $h_n = h_n(m)$ in (4.6.11) and (4.6.12) to conclude with a diagonal argument. □

Now we state the particular Γ-convergence result for our functionals.

Theorem 4.6.4 *If γ_n converge weakly to γ, then $\mathscr{F}(\cdot|\gamma_n)$ $\Gamma(\mathscr{P}_2(X))$-converge to $\mathscr{F}(\cdot|\gamma)$ and satisfy the equi-tightness condition* (4.6.4)*. Moreover, if $\mu \in \mathscr{P}_2(X)$ and $\gamma_n = \pi^n_\# \gamma$, a sequence satisfying condition* (4.6.3) *is $\pi^n_\#\mu$, so that*

$$\lim_{n\to\infty} \mathscr{F}(\pi^n_\#\mu|\gamma_n) = \mathscr{F}(\mu|\gamma).$$

Proof. We first prove the equi-tightness condition (4.6.4). Fix $\varepsilon > 0$ and two constants M', M'' large enough such that $M/M' < \varepsilon/2$ and $F(x) > M'x$ for $x > M''$ (this is possible in view of the superlinear growth of F at infinity). Let moreover K_ε be a compact subset of X such that $\gamma_n(K_\varepsilon) > 1 - \frac{\varepsilon}{2M''}$ for every n (the sequence (γ_n) is tight, since it is weakly convergent). If $\mu \in \mathscr{P}_2(X)$ satisfies $\mathscr{F}(\mu|\gamma_n) \le M$ for some n, we have

$$\begin{aligned}\mu(X\setminus K_\varepsilon) = \frac{1}{M'}\int_{X\setminus K_\varepsilon} M'\,d\mu &< \frac{1}{M'}\int_{(X\setminus K_\varepsilon)\cap\{\rho>M''\}} \frac{F(\rho)}{\rho}\,d\mu + \int_{(X\setminus K_\varepsilon)\cap\{\rho\le M''\}} M''\,d\gamma_n \\ &\le \frac{M}{M'} + M''\gamma_n(X\setminus K_\varepsilon) < \varepsilon.\end{aligned}$$

This shows that the set introduced in (4.6.4) is tight, hence relatively compact.

In order to prove Γ−convergence, let $\mu_n \to \mu$. For any $g \in C^0_b(X)$ there holds

$$\begin{aligned}\int_X g(x)d\mu(x) - \int_X F^*(g(x))d\gamma(x) &= \lim_{n\to\infty}\left(\int_X g(x)\,d\mu_n - \int_X F^*(g(x))\,d\gamma_n\right) \\ &\le \liminf_{n\to\infty} \mathscr{F}(\mu_n|\gamma_n),\end{aligned} \tag{4.6.15}$$

so that

$$\sup_{\mu\in C^0_b(X)}\left(\int_X g(x)d\mu(x) - \int_X F^*(g(x))d\gamma(x)\right) \le \liminf_{n\to\infty}\mathscr{F}(\mu_n|\gamma_n).$$

Taking into account the duality formula (4.4.3), the liminf inequality *i)* of the definition of Γ-convergence follows. The limsup inequality *ii)* and the last statement are proven exactly as in [9, Lemma 6.2]. □

Now consider finite dimensional approximations of the measure γ: letting $\gamma_n = \pi^n_\#\gamma$, from Theorem 4.6.4 we know that $\mathscr{F}(\cdot|\gamma_n)$ Γ-converge to $\mathscr{F}(\cdot|\gamma)$. From the next result it will follow that, if the role of Γ-converging functionals of Theorem 4.6.3 is played by $\mathscr{F}(\cdot|\gamma_n)$ and we choose a limit point $\mu \in L^\infty(X,\gamma)$, then the approximating μ_n can be chosen so that their densities have uniformly bounded $L^\infty(X,\gamma_n)$ norms.

Corollary 4.6.5 *For all μ with $\mu \le M\gamma$ and $|\partial\mathscr{F}|(\mu|\gamma)$ finite, there exist μ_n with $\mu_n \le M\gamma_n$, $\mu_n \to \mu$ in $\mathscr{P}_2(X)$, $\mathscr{F}(\mu_n|\gamma_n) \to \mathscr{F}(\mu|\gamma)$. In addition, there exist $\boldsymbol{\omega}_n \in \partial\mathscr{F}(\mu_n|\gamma_n)$ such that*

$$\boldsymbol{\omega}_n \to \partial^0\mathscr{F}(\mu|\gamma) \in L^2(X,\mu;X) \quad \textit{strongly in the sense of } \text{Definition 1.5.3.} \tag{4.6.16}$$

Proof. It suffices to revisit in this particular case the proof of Theorem 4.6.3: first, let $\mu^h \to \mu$ be such that $\mathscr{F}(\mu^h|\gamma^h) \to \mathscr{F}(\mu|\gamma)$, and by Theorem 4.6.4 we can choose $\mu^h = \Pi^h_\# \mu$, but the density of $\Pi^h_\# \mu$ with respect to γ^h is the cylindrical projection, which does not increase the L^∞ norm, so $\mu^h \le M\gamma^h$. Second, the minimizers μ^h_τ of

$$\nu \mapsto \mathscr{F}(\nu|\gamma^h) + \frac{1}{2\tau} W_2^2(\nu, \mu^h)$$

satisfy $\mu^h_\tau \le M\gamma^h$ by Lemma 4.4.5. □

4.7 Wasserstein subdifferential of $\mathscr{F}$

We will now characterize the subdifferential of $\mathscr{F}$. In this section we make Assumption 4.2.1 on γ, besides the log-concavity.

In the sequel we are using the stability of generalized Sobolev spaces under composition with L_F, namely $\rho \in GW^{1,1}(X,\gamma)$ implies $L_F \circ \rho \in GW^{1,1}(X,g)$. Indeed, since $L_F(z) \to +\infty$ as $z \to +\infty$ and L_F is strictly increasing, we have

$$T_\alpha(L_F \circ \rho) = L_F \circ T_{L_F^{-1}(\alpha)}(\rho), \tag{4.7.1}$$

(here T_α is the truncation operator) and since $T_\beta(\rho) \in W^{1,1}(X,\gamma)$ for any $\beta > 0$ we conclude that $L_F \circ \rho \in GW^{1,1}(X,\gamma)$ thanks to the chain rule.

We begin giving the following:

Definition 4.7.1 (Generalized Fisher information) *Let $\rho \in L^\infty(X,\gamma)$ and let $\rho \in W^{1,1}(X,\gamma)$. Assume that*

$$\sum_{j=1}^\infty \int_X \left| \frac{\partial_{\mathbf{e}_j}(L_F \circ \rho)}{\rho} \right|^2 d\mu(x) < +\infty. \tag{4.7.2}$$

We define the generalized Fisher information functional as follows:

$$\mathscr{G}(\rho\gamma|\gamma) := \left\| \frac{\nabla(L_F \circ \rho)}{\rho} \right\|^2_{L^2(X,\mu;X)}.$$

In the general case $\rho \in L^1(X,\gamma)$, $\rho \in GW^{1,1}(X,\gamma)$, the generalized Fisher information is defined by the same formula, using the fact that $L_F \circ \rho \in GW^{1,1}(X,\gamma)$, so its gradient is still well defined.

Lemma 4.7.2 (Lower semicontinuity of $\mathscr{G}$) *Let $(\rho_n) \subset W^{1,1}(X,\gamma)$, with $\rho_n \to \rho$ γ-a.e. and with $\mathscr{G}(\rho_n\gamma|\gamma)$ uniformly bounded. Then $\rho \in GW^{1,1}(X,\gamma)$ and*

$$\mathscr{G}(\rho\gamma|\gamma) \le \liminf_{n\to\infty} \mathscr{G}(\rho_n\gamma|\gamma).$$

Proof. We set $\rho_{n,k} := T_k(\rho_n)$. By dominated convergence, it is clear that $\rho_{n,k} \to T_k(\rho)$ in $L^2(X,\gamma)$ and that

$$L_F \circ \rho_{n,k} \to L_F \circ T_k(\rho) \quad \text{in } L^2(X,\gamma). \tag{4.7.3}$$

By the chain rule proven in Theorem 4.2.3, $\nabla(L_F \circ \rho_{n,k})$ is equal to $L'_F(\rho_{n,k})\nabla\rho_{n,k}$, so it vanishes where $\rho_n > k$ and coincides with $\nabla(L_F \circ \rho_n)$ where $\rho_n \le k$. As a consequence, there holds

$$\int \frac{\|\nabla(L_F \circ \rho_{n,k})\|^2}{\rho_{n,k}}\, d\gamma \le \int \frac{\|\nabla(L_F \circ \rho_n)\|^2}{\rho_n}\, d\gamma, \tag{4.7.4}$$

where the second term is uniformly bounded by hypothesis. In particular, $L_F \circ \rho_{n,k}$ is bounded in $W^{1,2}(X,\gamma)$ and therefore $L_F \circ T_k(\rho) \in W^{1,2}(X,\gamma)$. Since k is arbitrary, we can use $L_F(k)$ as truncation levels to prove that $L_F \circ \rho \in GW^{1,2}(X,\gamma)$; in addition, $\nabla(L_F \circ \rho_{n,k})$ weakly converge in $L^2(X,\gamma;X)$ to $\nabla(L_F \circ T_k(\rho))$.

We can take advantage of Ioffe's lower semicontinuity Theorem under strong-weak convergence (see for instance [3, Theorem 5.8]) to obtain

$$\int \frac{\|\nabla(L_F \circ T_k(\rho))\|^2}{T_k(\rho)}\, d\gamma \le \liminf_{n\to\infty} \int \frac{\|\nabla(L_F \circ \rho_{n,k})\|^2}{\rho_{n,k}}\, d\gamma. \tag{4.7.5}$$

This, in combination with (4.7.4), gives

$$\int \frac{\|\nabla(L_F \circ T_k(\rho))\|^2}{T_k(\rho)}\, d\gamma \le \liminf_{n\to\infty} \int \frac{\|\nabla(L_F \circ \rho_n)\|^2}{\rho_n}\, d\gamma.$$

To conclude, it suffices to show that the left hand side converges to $\mathscr{G}(\rho\gamma|\gamma)$ as $k \to \infty$. To this aim, it suffices to remind that $\nabla(L_F \circ T_k(\rho))$ vanishes where $\rho > k$ and coincides with $\nabla(L_F \circ \rho)$ where $\rho \le k$. □

We are ready for the result which identifies the Wasserstein subdifferential of $\mathscr{F}$.

Theorem 4.7.3 *Let $\mu = \rho\gamma \in \mathscr{P}_2(X)$, and assume that F satisfies* Assumption 4.4.1. *Then the metric slope of $\mathscr{F}(\cdot|\gamma)$ at μ is finite if and only if*

$$L_F \circ \rho \in GW^{1,1}(X,\gamma) \quad \textit{and} \quad \frac{\|\nabla(L_F \circ \rho)\|^2}{\rho} \in L^1(X,\gamma). \tag{4.7.6}$$

Moreover, in this case

$$\frac{\nabla(L_F \circ \rho)}{\rho} = \partial^0\mathscr{F}(\mu|\gamma) \quad \textit{and} \quad \mathscr{G}(\mu|\gamma) = |\partial\mathscr{F}|^2(\mu|\gamma).$$

Proof. **Step 1.** We prove that finiteness of slope at $\mu = \rho\gamma$ implies the regularity properties (4.7.6). First, assume $\rho \le M$, set $\phi(\nu) = \mathscr{F}(\nu|\gamma)$, $\gamma^d = \Pi^d_\#\gamma$ and $\phi_d(\nu) = \mathscr{F}(\nu|\gamma^d)$ and recall that $\gamma^d \rightharpoonup \gamma$ (see Proposition 1.2.3). Thanks to Theorem 4.6.4, ϕ_d $\Gamma(\mathscr{P}_2(X))$-converge to ϕ as $d \to \infty$. By Theorem 4.6.3 we can find sequences

$$\mu_d \to \mu \quad \text{in } \mathscr{P}_2(X), \quad \phi_d(\mu_d) \to \phi(\mu)$$

$$\boldsymbol{\omega}_d \in \partial\phi_d(\mu_d) \text{ such that } \boldsymbol{\omega}_d \to \boldsymbol{\omega} = \partial^0\phi(\mu) \text{ strongly in } L^2 \text{ as in Definition 1.5.3,} \tag{4.7.7}$$

and thanks to (4.6.10) we have also that $|\partial\phi_d|(\mu_d)$ is finite and uniformly bounded in d. We can also choose μ_d so that the additional property $\mu_d \le M\gamma_d$ holds, by Corollary 4.6.5. We stress that here μ_d is not, in general, the projection $\Pi^d_{\#}\mu$. Let ρ_d be the density of μ_d with respect to γ^d, and since $\gamma^d \rightharpoonup \gamma$ and $\mu_d \to \mu$ in $\mathscr{P}_2(X)$, we have that $\rho_d \to \rho$ in the sense of Definition 1.5.3, in its scalar version. Together with (4.7.7), which guarantees convergence of the energies, this also implies, thanks to Lemma 2.2.9 and the strict convexity of F, that

$$\int_X \varphi(x) L_F \circ \rho_d(x) d\gamma^d(x) \to \int_X \varphi(x) L_F \circ \rho(x) d\gamma(x) \qquad \forall \varphi \in L^1(X,\gamma). \tag{4.7.8}$$

Indeed, (4.7.8) holds independently of the growth of L_F for all $\varphi \in C^0_b(X)$, as ρ and ρ_d are essentially bounded, uniformly with respect to d, and the same uniform bound allows to extend the validity of the formula to all $\varphi \in L^1(X,\gamma)$.

The theorem holds if X is finite-dimensional, and since γ^d is supported in $\Pi^d(X)$ we can use the implication in finite dimension (Theorem 4.5.1) to obtain, for $\zeta \in \mathrm{Cyl}(X)$, $j \le d$ and d large enough (depending on ζ only),

$$\begin{aligned}\int_X \partial_{\mathbf{e}_j}\zeta(x) L_F \circ \rho_d(x)\, d\gamma^d(x) = &-\int_X \partial_{\mathbf{e}_j}(L_F \circ \rho_d)(x)\zeta(x)\, d\gamma^d(x) \\ &- \int_X L_F \circ \rho_d(x)\zeta(x) g^d_j(x)\, d\gamma^d(x),\end{aligned} \tag{4.7.9}$$

where we used also the fact that $\partial_{\mathbf{e}_j}\gamma = g^j\gamma$ implies $\partial_{\mathbf{e}_j}\gamma^d = g^d_j\gamma^d$, g^d_j being the cylindrical projection of g^j (see Definition 1.4.2 and Lemma 1.4.1). The finite dimensional result also tells us that

$$\omega^d_j := \frac{\partial_{\mathbf{e}_j}(L_F \circ \rho_d)}{\rho_d} \in L^2(X,\mu_d), \qquad j = 1,\dots,d,$$

so we can rewrite (4.7.9) as

$$\begin{aligned}\int_X \partial_{\mathbf{e}_j}\zeta(x) L_F \circ \rho_d(x)\, d\gamma^d(x) = &-\int_X \omega^d_j(x)\zeta(x)\, d\mu_d(x) \\ &- \int_X L_F \circ \rho_d(x)\zeta(x) g^d_j(x)\, d\gamma^d(x).\end{aligned} \tag{4.7.10}$$

Now we pass to the limit in (4.7.10) as $d \to \infty$. The first term converges to the analogous term involving γ and ρ by (4.7.8), the second one converges too, thanks to (4.7.7). Adding and subtracting g_j in the last term and using (4.7.8) with $\varphi = g_j$ we have also convergence of that term. Hence, we find (letting $\omega_j = \langle \boldsymbol{\omega}, \mathbf{e}_j\rangle$)

$$\begin{aligned}\int_X \partial_{\mathbf{e}_j}\zeta(x) L_F \circ \rho(x)\, d\gamma(x) = &-\int_X \omega_j(x)\zeta(x)\, d\mu(x) \\ &- \int_X L_F \circ \rho(x)\zeta(x) g_j(x)\, d\gamma(x) \quad \forall j \in \mathbb{N},\end{aligned} \tag{4.7.11}$$

that is, $\partial_{\mathbf{e}_j}(L_F \circ \rho) = \rho\omega_j \in L^1(X,\gamma)$. Finally, since $\boldsymbol{\omega} \in L^2(X,\mu;X)$, we obtain $L_F \circ \rho \in W^{1,1}(X,\gamma)$ and

$$\boldsymbol{\omega} = \frac{\nabla(L_F \circ \rho)}{\rho}, \tag{4.7.12}$$

and since $\boldsymbol{\omega}$ is the minimal selection we have also

$$\mathscr{G}(\mu|\gamma) = |\partial\mathscr{F}|^2(\mu|\gamma).$$

We have proven the implication for the bounded case. Now we shall pass to the general one. Let $n \in \mathbb{N}$ and consider functionals $\mathscr{F}^n(\cdot|\gamma)$, defined in (4.4.7). These functionals are strongly convex, as noticed in Remark 4.4.4, and $\Gamma(\mathscr{P}_2(X))$-converge to $\mathscr{F}(\cdot|\gamma)$ as $n \to \infty$ (indeed, condition (4.6.2) is trivial, whereas (4.6.3) can be achieved by a truncation argument). Moreover, since $\mathscr{F}^n \geq \mathscr{F}$, it is easy to show tightness for the sets corresponding to the ones in (4.6.4). Then, by means of Theorem 4.6.3 again, we find subsequences (that we don't relabel) $\mu_n \to \mu$ in $\mathscr{P}_2(X)$ and $\boldsymbol{\omega}_n \in \partial\mathscr{F}^n(\mu_n|\gamma)$ such that $\mathscr{F}^n(\mu_n|\gamma) \to \mathscr{F}(\mu|\gamma)$ and

$$\boldsymbol{\omega}_n \to \boldsymbol{\omega} = \partial^0\mathscr{F}(\mu|\gamma) \text{ strongly in } L^2 \text{ as in Definition 1.5.3.} \tag{4.7.13}$$

We have $\rho_n \leq n$, since $\mathscr{F}^n(\mu_n|\gamma)$ is finite. So, the already obtained result for the bounded case entails $L_F \circ \rho_n \in W^{1,1}(X,\gamma)$ and ensures that the square of the metric slope at μ_n is characterized as

$$\mathscr{G}(\rho_n\gamma|\gamma) = \int_X \frac{\|\nabla L_F \circ \rho_n\|^2}{\rho_n}\,d\gamma.$$

Notice that the weak convergence of $\rho_n\gamma$ to $\rho\gamma$ and the convergence of $\mathscr{F}(\rho_n\gamma|\gamma) = \mathscr{F}^n(\rho_n\gamma|\gamma)$ to $\mathscr{F}(\rho\gamma|\gamma)$ imply, thanks to the strict convexity of F, that $\rho_n \to \rho$ in γ-measure (see [74, Theorem 3] or [17]); in particular a subsequence of (ρ_n) converges to ρ γ-a.e. Hence, we can apply Lemma 4.7.2 to that subsequence to conclude that $L_F \circ \rho \in GW^{1,1}(X,\gamma)$ and that

$$\int_X \frac{\|\nabla(L_F \circ \rho)\|^2}{\rho}\,d\gamma \leq |\partial\mathscr{F}|^2(\mu|\gamma). \tag{4.7.14}$$

Step 2. Now we prove that Sobolev regularity of $L_F \circ \rho$ and integrability of $\|\nabla(L_F \circ \rho)\|^2/\rho$ imply the opposite inequality in (4.7.14), hence finiteness of slope. First, assume that ρ is bounded and distant from zero. Since ρ^{-1} is bounded we have $\|\nabla(L_F \circ \rho)\| \in L^2(X,\gamma)$, and since L_F has a locally Lipschitz inverse by strict convexity of F, Theorem 4.2.3 yields $\rho \in W^{1,2}(X,\gamma)$. Let ρ^d be the d-dimensional cylindrical projection of ρ. By (4.2.3), $\rho^d \in W^{1,2}(X,\gamma)$ and again Theorem 4.2.3 gives

$$L_F \circ \rho^d \in W^{1,2}(X,\gamma). \tag{4.7.15}$$

Moreover, by the chain rule (4.2.4) we have

$$\nabla(L_F \circ \rho) = L_F'(\rho)\nabla\rho \quad \text{and} \quad \nabla(L_F \circ \rho^d) = L_F'(\rho^d)\nabla\rho^d, \tag{4.7.16}$$

and these gradients are respectively 0 γ-a.e. on the set of all x such that L_F is not differentiable at $\rho(x)$, $\rho^d(x)$. Since ρ^d and ρ are distant from zero, by (1.4.2) there holds

$$\frac{(L'_F(\rho^d))^2\|\nabla(\rho^d)\|^2}{\rho^d} \to \frac{(L'_F(\rho))^2\|\nabla\rho\|^2}{\rho} \quad \text{in} \quad L^1(X,\gamma).$$

In fact

$$\|\nabla(\rho^d) - \nabla\rho\|^2 \leq \|(\nabla\rho)^d - \nabla\rho\|^2 + \sum_{j=d+1}^{\infty} |\partial_{\mathbf{e}_j}\rho|^2$$

converges to 0 in $L^1(X,\gamma)$ (we use (4.2.3) and the fact that the convergence (1.4.2) of cylindrical projections holds for maps with values in X, like $\nabla\rho$ with its projection $(\nabla\rho)^d$).

On the other hand, $(L'_F(\rho_d))^2/\rho^d$ converge to $(L'_F(\rho))^2/\rho$ in $L^1(X,\gamma)$ and are essentially bounded uniformly in d. Then

$$\lim_{d\to\infty}\int_X \frac{\|\nabla(L_F\circ\rho^d)\|^2}{\rho^d}\,d\gamma^d = \int_X \frac{\|\nabla(L_F\circ\rho)\|^2}{\rho}\,d\gamma. \tag{4.7.17}$$

In view of (4.7.15), we can apply Theorem 4.5.1 and obtain the finiteness of $|\partial\mathscr{F}(\mu^d|\gamma^d)|$, where $\mu^d = \rho^d\gamma^d$ (so $\mu^d = \Pi^d_\#\mu$), and also $|\partial\mathscr{F}|^2(\mu^d|\gamma^d) = \mathscr{G}(\mu^d|\gamma^d)$. Now we make use of the lower semicontinuity of the metric slope and of (4.7.17) to infer the finiteness of the slope:

$$|\partial\mathscr{F}|^2(\mu|\gamma) \leq \liminf_{d\to\infty}|\partial\mathscr{F}|^2(\mu^d|\gamma^d) \leq \int_X \frac{\|\nabla(L_F\circ\rho)\|^2}{\rho}\,d\gamma.$$

Now consider the case in which ρ is bounded but not necessarily distant from 0. Let $\rho_n = \max\{\rho, \frac{1}{n}\}$, so that ρ_n is distant from zero, and $\mu_n = \rho_n\gamma$.

Notice that ρ_n are not probability measures, but the results we apply are obviously still valid if, instead of working in $\mathscr{P}_2(X)$, one works in the space $z\mathscr{P}_2(X)$ with $z > 0$ (this can also be seen considering the map $F_z(s) = F(zs)$, to come back to probability measures, as we do in Step 3). Since L_F is nondecreasing, $L_F\circ\rho_n = \max\{L_F\circ\rho, L_F(\frac{1}{n})\}$, and by Theorem 4.2.3 we can infer that $L_F\circ\rho_n \in W^{1,1}(X,\gamma)$. The chain rule also gives

$$\int_X \frac{\|\nabla(L_F\circ\rho_n)\|^2}{\rho_n}\,d\gamma \leq \int_X \frac{\|\nabla(L_F\circ\rho)\|^2}{\rho}\,d\gamma, \tag{4.7.18}$$

since $\rho_n \geq \rho$ and $\nabla(L_F\circ\rho_n) = 0$ γ-a.e. on $\{\rho < 1/n\}$. Since we have proven the theorem for the case of a density distant from zero, we have by (4.7.18) that

$$|\partial\mathscr{F}|^2(\mu_n|\gamma) \leq \int_X \frac{\|\nabla(L_F\circ\rho)\|^2}{\rho}\,d\gamma.$$

Using the lower semicontinuity of the slope we conclude.

Finally, in the general unbounded case, we take advantage of the just achieved characterization of the slope at $T_n(\rho)\gamma$. The slope is lower semicontinuous, and reasoning as we did to obtain (4.7.4), we get

$$|\partial\mathscr{F}|^2(\mu|\gamma) \le \liminf_{n\to\infty} |\partial\mathscr{F}|^2(T_n(\rho)\gamma|\gamma) = \liminf_{n\to\infty} \int_X \frac{\|\nabla(L_F\circ T_n(\rho))\|^2}{T_n(\rho)}\,d\gamma$$
$$\le \int_X \frac{\|\nabla(L_F\circ\rho)\|^2}{\rho}\,d\gamma. \tag{4.7.19}$$

Step 3. Suppose now that either the metric slope at μ is finite or that (4.7.6) hold. Joining together (4.7.14) and (4.7.19) we get the desired equality $|\partial\mathscr{F}|^2(\mu|\gamma) = \mathscr{G}(\mu|\gamma)$. Then, in order to characterize the minimal selection $\partial^0\mathscr{F}(\mu|\gamma)$, we have to show that $\nabla(L_F\circ\rho)/\rho$ belongs to $\partial\mathscr{F}(\mu|\gamma)$. We know from (4.7.12) that this is true if ρ is bounded. In the general case we check the subdifferential relation (3.1.7) with $\lambda = 0$, $\phi = \mathscr{F}$ and $\xi = \nabla(L_F\circ\rho)/\rho$ by approximation; thanks to Lemma 4.3.2, it suffices to check the property for all $\nu = f\gamma$ with f bounded. Now we approximate ρ by $\rho_n := z_n^{-1}(\rho\wedge n)$, where $z_n \uparrow 1$ is a normalizing constant, and we write the subdifferential relation for ρ_n, $F_n(s) = F(z_n s)$, to obtain:

$$\int_X F(z_n f(x))\,d\gamma(x) \ge \int_X F(\rho(x)\wedge n)\,d\gamma(x) + \int_X \left\langle \frac{\nabla(L_{F_n}\circ\rho_n)(x)}{\rho_n(x)}, \mathbf{t}_n(x) - x\right\rangle \rho_n(x)\,d\gamma(x),$$

where $\mathbf{t}_n$ are the optimal maps from $\rho_n\gamma$ to ν. Since $L_{F_n}(s) = L_F(z_n s)$, $L_{F_n}\circ\rho_n = L_F\circ(\rho\wedge n)$, and using the chain rule this immediately gives

$$\lim_{n\to\infty} \int_X \left\| \frac{\nabla(L_{F_n}\circ\rho_n)}{\rho_n} - \frac{\nabla(L_F\circ\rho)}{\rho} \right\|^2 \rho_n\,d\gamma = 0.$$

Hence, we need only to check that

$$\lim_{n\to\infty} \int_X \left\langle \frac{\nabla(L_F\circ\rho)}{\rho}, \mathbf{t}_n - \mathbf{I}\right\rangle \rho_n\,d\gamma = \int_X \left\langle \frac{\nabla(L_F\circ\rho)}{\rho}, \mathbf{t} - \mathbf{I}\right\rangle \rho\,d\gamma.$$

By a density argument, it suffices to check that

$$\lim_{n\to\infty} \int_X \langle g(x), \mathbf{t}_n(x) - x\rangle\,\rho_n(x)\,d\gamma(x) = \int_X \langle g(x), \mathbf{t}(x) - x\rangle \rho(x)\,d\gamma(x)$$

for all $g \in C_b(X;X)$. Writing the integrals above in terms of optimal plans, the formula reduces to

$$\lim_{n\to\infty} \int_X \langle g(x), y - x\rangle\,d\beta_n(x,y) = \int_X \langle g(x), y - x\rangle\,d\beta(x,y).$$

The latter is a direct consequence of the tightness of (β_n) (because the marginals are tight), of the fact that any limit point is an optimal plan from $\rho\gamma$ to γ (see Lemma 2.2.6) and of the uniqueness of β proved in Theorem 4.3.1. □

After Theorem 4.7.3, we can give a straightforward proof of the main result of this chapter.

Theorem 4.7.4 *Assume that $L = L_F$, with F satisfying* Assumption 4.4.1*, and that γ satisfies* Assumption 4.2.1*. Then, for all $\mu^0 \in \mathscr{P}_2(X)$ there exists a distributional solution $\mu_t = \rho_t \gamma$ to* (4.1.3)*, satisfying $L_F \circ \rho_t \in GW^{1,1}(X, \gamma)$ for a.e. $t > 0$ and:*

$$\left\| \frac{\nabla(L_F \circ \rho_t)}{\rho_t} \right\|_{L^2(X,\mu_t;X)} \in L^2_{loc}(0, +\infty). \tag{4.7.20}$$

In the class of solutions μ_t satisfying (4.7.20) *this solution is unique. Furthermore, if $\bar{\mu} \le C\gamma$, then $\rho_t \le C$ γ-a.e. for all $t > 0$ and therefore $L_F \circ \rho_t \in W^{1,1}(X, \gamma)$ for a.e. $t > 0$.*

Proof. Notice that the domain $D(\mathscr{F}(\cdot|\gamma))$ is dense in $\mathscr{P}_2(X)$ and, under Assumption 4.4.1, $\mathscr{F}(\cdot|\gamma)$ is strongly convex. Hence we can apply Theorem 3.5.8 to obtain, for any $\mu^0 \in \mathscr{P}_2(X)$, existence and uniqueness of the gradient flow μ_t of $\mathscr{F}(\cdot|\gamma)$ starting from μ^0. Notice that, by the regularizing effect of the semigroup, $\mu_t \ll \gamma$ for any $t > 0$ even if μ^0 does not have a density with respect to γ. The curve μ_t satisfy (3.3.3) and (3.3.2), and with Theorem 4.7.3 we have characterized, under Assumption 4.2.1, the minimal selection in the Wasserstein subdifferential of $\mathscr{F}(\cdot|\gamma)$ at $\mu_t = \rho_t\gamma$ as $\frac{\nabla(L_F \circ \rho_t)}{\rho_t}$. We deduce that $\mu_t = \rho_t\gamma$ is a solution to (4.1.3). This solution is unique and satisfies all the additional properties of Theorem 3.5.8.

Finally, if $\mu^0 \le M\gamma$, we know by Lemma 4.4.5 that such a bound is preserved by the discrete minimizer of functional $\Phi^{\mathscr{F}}_\tau(\cdot, \mu^0)$ defined in (4.4.4) (independently of the value of τ). Since μ_t is the limit of discrete minimizers, we conclude that $\rho_t \le M$ γ-a.e. for all $t \ge 0$. $\square$

Chapter 5

Mean-field evolution model in superconductivity

This chapter is devoted to the application of the gradient flow theory to another evolution problem. The corresponding energy functional arises in the Ginzburg-Landau theory for superconductivity. In this case we won't have the λ-geodesical convexity property, which characterized the analysis of Chapter 4, but we will se how, applying the abstract theory of Section 3.4, we can still obtain a satisfactory description. We start with a brief review of the physical framework (for the general theory see for instance [70]).

5.1 The physical context

The well known *Ginzburg-Landau* energy functional is

$$J(u, A) = \frac{1}{2}\int_\Omega |\nabla_A u|^2 + |h - h_{ex}|^2 + \frac{1}{2\varepsilon^2}(1 - |u|^2)^2, \tag{5.1.1}$$

where $\Omega \subset \mathbb{R}^2$ is the section of the superconductor, h_{ex} represents the intensity of an external magnetic field, constant and orthogonal to the section, A is the potential vector of the magnetic field h induced in the material ($h = \nabla \times A$ and $\nabla_A = \nabla - iA$), and ε is a parameter depending on the material and on the temperature. The function u takes complex values and its modulus ($|u| \leq 1$) accounts for the density of superconducting electron pairs, so that a value close to 1 indicates a significant presence of the superconducting phenomenon (the *superconducting phase*). The case $|u| \simeq 0$ is called the *normal phase.* This functional is associated to the *Ginzburg-Landau equations*

$$\begin{cases} -\nabla_A^2 u = \dfrac{1}{\varepsilon^2} u(1 - |u|^2), \\ \nabla \times h = \langle iu, \nabla_A u\rangle. \end{cases}$$

The boundary conditions are $h = h_{ex}$ and $\langle \nabla_A u, \mathbf{n}\rangle = 0$.

Different behaviors are observed for different values of the applied magnetic field intensity h_{ex} with respect to the parameter ε. At low temperatures (small ε), if the applied magnetic field intensity is itself sufficiently low, the material is superconductive. This means that the magnetic field has no relevant effect (it is said to be 'expelled') and in this case there holds approximately

$$\begin{cases} -\nabla h + h = 0 & \text{in } \Omega, \\ h = h_{ex} & \text{on } \partial\Omega. \end{cases}$$

Increasing h_{ex}, an opposition to the superconductive phenomenon appears, so that $|u|$ tends to decrease and to reach the value 0 of the normal phase. For intermediate values of h_{ex}, the so-called *mixed phase* is observed. That is, the normal phase gets concentrated in small regions, the *Ginzburg-Landau vortices*, at the center of which $|u| = 0$. If C is a small circle around one of these zeros, the *degree* of the vortex is defined as the topological degree of the map $u/|u| : C \to S^1$. Vortices with same degree tend to repel each other. In particular, the mixed phase starts when h_{ex} reaches the order of $|\log\varepsilon|$. To be more precise, let, as in [64],

$$\lambda = \lim_{\varepsilon\to 0} \frac{|\log\varepsilon|}{h_{ex}(\varepsilon)}. \tag{5.1.2}$$

When λ is finite and positive or zero (in the latter case with a not too large magnetic field, that is $h_{ex} \ll \varepsilon^{-2}$), we are in the mixed phase, with the vortex structure. It is shown in [64] (see also [65]) that the functional

$$\Phi_\lambda(\mu) = \frac{\lambda}{2}|\mu|(\Omega) + \frac{1}{2}\int_\Omega |\nabla h_\mu|^2 + |h_\mu - 1|^2 \quad \lambda \geq 0, \tag{5.1.3}$$

where λ is given by (5.1.2), is the limit as $\varepsilon \to 0$, in a suitable sense, of the Ginzburg-Landau functional defined by (5.1.1). The measure μ represents the density of vortices, whereas h_μ is the induced magnetic field.

In [28], the authors proposed an evolution model for the vortex density, which reads

$$\begin{gathered} \frac{d}{dt}\mu(t) - \operatorname{div}(\nabla h_{\mu(t)}|\mu(t)|) = 0, \\ \begin{cases} -\Delta h_{\mu(t)} + h_{\mu(t)} = \mu(t) & \text{in } \Omega, \\ h_{\mu(t)} = 1 & \text{on } \partial\Omega. \end{cases} \end{gathered} \tag{5.1.4}$$

Here $|\cdot|$ denotes the total variation. This model and similar ones have been investigated for instance in [8, 48, 50, 52]. In particular, in [48, 52] the authors are concerned also with different couplings between μ and h_μ, like $\Delta h_\mu = \mu$, giving rise to the equation

$$\frac{d}{dt}\mu(t) + \operatorname{div}(\nabla\Delta^{-1}\mu(t)|\mu(t)|). \tag{5.1.5}$$

For positive measures, this is reminiscent of incompressible Euler equations, where $\nabla^\perp$ appears instead of ∇. This rotation makes (5.1.5) dissipative.

The aim of this chapter is to address (5.1.4) as the gradient flow of functional (5.1.3). In the first part we will assume that μ is a positive measure. Let us begin with a rigorous formulation.

5.2 Formulation of the problem

Let Ω be a bounded open connected region in $\mathbb{R}^2$, with smooth boundary, representing the section of the superconductor. Denote with $\mathscr{P}(\overline{\Omega})$ the space of probability measures over $\overline{\Omega}$ (we omit the notation $\mathscr{P}_2(\overline{\Omega})$, since in this case the spaces and the topologies coincide). We are concerned with the following evolution problem:

$$\frac{d}{dt}\mu(t) - \operatorname{div}(\chi_\Omega \nabla h_{\mu(t)}\mu(t)) = 0 \quad \text{in } \mathcal{D}'((0,+\infty)\times\mathbb{R}^2) \tag{5.2.1}$$

with the initial datum $\mu(0) = \mu_0 \in \mathscr{P}(\overline{\Omega}) \cap H^{-1}(\Omega)$. We look for a solution $\mu(t)$ (the vortex density) which is a measure in $\mathscr{P}(\overline{\Omega}) \cap H^{-1}(\Omega)$. For every t the velocity field $-\chi_\Omega \nabla h_\mu$ and μ are coupled by

$$\begin{cases} -\Delta h_\mu + h_\mu = \mu & \text{in } \Omega \\ h_\mu = 1 & \text{on } \partial\Omega. \end{cases} \tag{5.2.2}$$

Clearly, $H^{-1}(\Omega)$ is the natural ambient space for the problem, so we are working with measures on $\overline{\Omega}$ in order to treat masses in Ω which vary during the evolution. In fact, concentration of μ on $\partial\Omega$ will model the 'expulsion' phenomenon of the vortices. Masses on $\overline{\Omega}$ are also normalized to 1 without loss of generality.

Let $\mathcal{M}_+(\Omega)$ be the space of nonnegative measures on Ω. The weak (or narrow) topology here is again defined by the convergence in duality with continuous and bounded functions. Until Section 5.6, we will work with positive measures only, so that we can write (5.1.3) as

$$\Phi_\lambda(\mu) = \frac{\lambda}{2}\mu(\Omega) + \frac{1}{2}\int_\Omega |\nabla h_\mu|^2 + |h_\mu - 1|^2, \quad \lambda \geq 0. \tag{5.2.3}$$

For measures μ on $\overline{\Omega}$ we will write $\mu = \widehat{\mu} + \widetilde{\mu}$, where $\widehat{\mu} = \chi_\Omega \mu$ and $\widetilde{\mu} = \chi_{\partial\Omega}\mu$. Functional (5.2.3), which is defined in $\mathcal{M}_+(\Omega)$, will be understood to be defined as $\Phi_\lambda(\widehat{\mu})$ for $\mu \in \mathscr{P}(\overline{\Omega})$. So, it depends only on the internal part of the measure.

It is shown in [8] that equation (5.2.1), with the coupling described by (5.2.2), can be viewed as a gradient flow of functionals (5.2.3) with respect to the structure induced on $\mathscr{P}(\overline{\Omega})$ by the 2-Wasserstein distance W_2. Therefore, in [8] the problem is studied exploiting the techniques of gradient flows in metric spaces developed in Chapter 3, and a global existence result is proved therein. The method is the classical one: a family of minimizers of the discrete, minimizing movements scheme (see Section 3.2) is found, then the family of measures $\mu(t)$ is built as the limit of subsequences of interpolations, given by (3.2.4). This limit satisfies a continuity equation with a suitable velocity field (see Section 2.3). In this case we no more analyze directly the Wasserstein subdifferential of the energy functional. Rather, the velocity field is shown to be the same of the evolutionary problem under investigation, thanks to suitable Euler-Lagrange equations associated to (3.2.2).

In [8], with the introduction of some ‘entropies’ which are shown to decrease along the flow, a regularity result is also obtained, that is, if the initial datum μ_0 is such that $\widehat{\mu}_0 \in L^p(\Omega)$, $p \geq 4/3$, then there exist a global solution $\mu(t)$ such that $\|\widehat{\mu}(t)\|_p$ is uniformly controlled by the L^p norm of $\widehat{\mu}_0$.

Finally, in the case $p = +\infty$, a short time uniqueness theorem is established in [8, Theorem 3.6]. Here we are going to discuss also a global time uniqueness result, following [50], based on the introduction of a one-sided condition on the velocity vector field at the boundary. The condition is the following: for $t \in (0, T]$,

$$\langle \nabla h_{\mu(t)}(x), y - x\rangle \geq 0 \text{ for all } (x, y) \in \operatorname{supp}(\widetilde{\mu}(t)) \times \overline{\Omega}. \tag{5.2.4}$$

We stress that we allow $\mu(t)$ to have a nonzero boundary part. Concerning this new condition (5.2.4), we will show later in Section 3 that it is a byproduct of our Wasserstein variational approach. In Theorem 5.4.4 we will indeed prove the analogous property for discrete minimizers of $\Phi_\lambda(\mu) + \frac{1}{2\tau}W_2^2(\mu, \mu^0)$, in the case $\lambda = 0$.

Notice that, since the domain is supposed to be convex, (5.2.4) can be interpreted as follows: the gradient of $h_{\mu(t)}$ on the boundary (whenever some mass is there present) points towards the interior of the domain. This is in fact reminiscent of the nondecreasing boundary mass condition appearing in [8, Definition 3.1], which is meaningful since a gradient flow of Φ_λ, at least for $\lambda > 0$, is expected to enjoy such a behavior (see the energy comparison argument in [8, Section 3]. See also Lemma 5.3.2 below).

Inequalities about the functional

Now we introduce some basic results that will often be useful in the sequel. For the first we refer to [8, Lemma 2.1].

Proposition 5.2.1 *Φ_λ is lower semicontinuous with respect to the weak topology of $\mathcal{M}_+(\Omega)$.*

Proposition 5.2.2 (Representation formula) *There holds*

$$\Phi_\lambda(\mu) = \frac{1}{2}(\lambda\mu(\Omega) + |\Omega|) + \sup_{h-1 \in H_0^1(\Omega)} \left\{ \int_\Omega (h-1)\, d\mu - \frac{1}{2}\int_\Omega |\nabla h|^2 + |h|^2 \right\}, \tag{5.2.5}$$

and the supremum is attained for $h = h_\mu$.

Proof. The direct method gives existence and uniqueness of a maximizer h for (5.2.5), while taking the first variation we have

$$\int_\Omega \varphi\, d\mu - \int_\Omega \nabla h \cdot \nabla \varphi + h\varphi = 0 \qquad \forall \varphi \in C_c^\infty(\Omega),$$

and therefore $h = h_\mu$. Finally, we can use the identity

$$\int_\Omega (h_\mu - 1)\, d\mu = \int_\Omega (h_\mu - 1)(-\Delta h_\mu + h_\mu) = \int_\Omega h_\mu^2 - h_\mu + |\nabla h_\mu|^2$$

to obtain (5.2.5). □

Lemma 5.2.3 *For all $\mu, \nu \in \mathcal{M}_+(\Omega)$ there hold*

$$\Phi_\lambda(\mu) - \frac{\lambda}{2}\mu(\Omega) \geq \Phi_\lambda(\nu) - \frac{\lambda}{2}\nu(\Omega) + \int_\Omega (h_\nu - 1)\, d(\mu - \nu) \tag{5.2.6}$$

and

$$\Phi_\lambda(\mu) - \Phi_\lambda(\nu) = \left(\frac{\lambda}{2} - 1\right)(\mu(\Omega) - \nu(\Omega)) + \frac{1}{2}\int_\Omega (h_\mu + h_\nu) d(\mu - \nu) \tag{5.2.7}$$

Proof. For the first relation, apply (5.2.5) to the difference $\Phi_\lambda(\mu) - \Phi_\lambda(\nu)$, taking into account that the supremum therein is attained at h_ν when the argument is ν. The second one is obtained computing

$$\begin{aligned}
\Phi_0(\mu) - \Phi_0(\nu) &= \frac{1}{2}\int_\Omega (h_\mu + h_\nu - 2)(\Delta h_\nu - \Delta h_\mu) + \frac{1}{2}\int_\Omega (h_\mu - 1)^2 - (h_\nu - 1)^2 \\
&= \frac{1}{2}\int_\Omega (h_\mu + h_\nu - 2)(\mu - \nu - h_\mu + h_\nu) + \frac{1}{2}\int_\Omega (h_\mu - 1)^2 - (h_\nu - 1)^2 \\
&= \nu(\Omega) - \mu(\Omega) + \frac{1}{2}\int_\Omega (h_\mu + h_\nu)(\mu - \nu) \\
&= \frac{1}{2}\int_{\overline{\Omega}} (h_\mu + h_\nu)(\mu - \nu),
\end{aligned}$$

where we started integrating by parts the quantity $|\nabla h_\mu|^2 - |\nabla h_\nu|^2$, which can be written as $\nabla(h_\mu + h_\nu - 2) \cdot \nabla(h_\mu - h_\nu)$. □

Moreover, we have

Lemma 5.2.4 *For all $\mu, \nu \in P(\overline{\Omega})$ there hold*

$$\Phi_\lambda(\nu) - \Phi_\lambda(\mu) \geq \frac{\lambda}{2}(\widehat{\nu}(\Omega) - \widehat{\mu}(\Omega)) + \int_{\overline{\Omega}} h_\mu d(\nu - \mu) \tag{5.2.8}$$

and

$$\Phi_\lambda(\mu) - \Phi_\lambda(\nu) = \left(\frac{\lambda}{2} - 1\right)(\widehat{\mu}(\Omega) - \widehat{\nu}(\Omega)) + \frac{1}{2}\int_\Omega (h_\mu + h_\nu) d(\widehat{\mu} - \widehat{\nu}). \tag{5.2.9}$$

Proof. These are straightforward consequences of Lemma 5.2.3, taking into account that $h_\mu|_{\partial\Omega} = 1$ and that, since the solution of problem (5.2.2) does not depend on the boundary part of μ, we have $h_{\widehat{\mu}} = h_\mu$. □

Formal gradient flow

Here we briefly see how functionals (5.2.3) are related to the Chapman Rubinstein Schatzman model. We can show that the latter is (at least formally for now) the gradient flow of Φ_0 with respect to the Wasserstein structure, that is, ∇h_μ is the gradient of Φ_0 at μ along transport maps. The Wasserstein (sub)gradient is a vector $\xi \in L^2(\mu;\mathbb{R}^2)$ defined by the subdifferential relation (3.1.6). Now consider functional (5.2.3). By (5.2.6) we are led to

$$\begin{aligned}\Phi_\lambda(\mathbf{s}_{\#}\mu) - \Phi_\lambda(\mu) &\geq \frac{\lambda}{2}\left(s_{\#}\mu(\Omega) - \mu(\Omega)\right) + \int_{\overline{\Omega}}(h_\mu - 1)d(\mathbf{s}_{\#}\mu - \mu) \\ &= \frac{\lambda}{2}\left(s_{\#}\mu(\Omega) - \mu(\Omega)\right) + \int_{\overline{\Omega}}(h_\mu(\mathbf{s}(x)) - h_\mu(x))\,d\mu.\end{aligned}$$

Since

$$\int_{\overline{\Omega}}(h_\mu(\mathbf{s}(x)) - h_\mu(x))\,d\mu \sim \int_{\overline{\Omega}}\nabla h_\mu(x)\cdot(\mathbf{s}(x) - x)\,d\mu$$

as $\|\mathbf{s} - I\|_{L^2(\mu)} \to 0$, if $\lambda = 0$, the formal Wasserstein gradient of Φ_λ at μ (if $\mu = \widehat{\mu}$) is $\chi_\Omega \nabla h_\mu$. The argument works also with $\lambda > 0$ if we consider transports which do not increase the mass on $\partial\Omega$.

5.3 Global existence and regularity

The results of this section have been obtained in [8]. Here we reproduce them, since they will be important for the subsequent analysis.

Given $\tau > 0$ and $\mu \in \mathscr{P}(\overline{\Omega})$, we denote as usual by μ_τ a minimizer of

$$\min_{\nu\in\mathscr{P}(\overline{\Omega})} \nu \mapsto \Phi_\lambda(\nu) + \frac{1}{2\tau}W_2^2(\nu,\mu). \tag{5.3.1}$$

Since $\mathscr{P}(\overline{\Omega})$ is compact for the weak convergence, existence is an easy consequence of the lower semicontinuity of Φ in $\mathscr{P}(\overline{\Omega})$, given by Proposition 5.2.1, and of the continuity of $W_2^2(\cdot,\mu)$. Uniqueness of minimizers, on the other hand, is not completely clear, however, the next result suffices to prove that $\widehat{\mu}_\tau$ is unique.

Variational arguments

We first prove some simple property of all minimizers μ_τ that can be achieved by non-differential variations: no mass is moved within $\partial\Omega$, and if some mass in Ω is moved to $\partial\Omega$, in passing from μ to μ_τ, it has to be moved along a shortest connection to $\partial\Omega$. Furthermore, if $\lambda > 0$, for τ sufficiently small no mass moved on $\partial\Omega$ returns to Ω, i.e. no new mass enters Ω. The corresponding result for $\lambda = 0$ will be object of Section 5.4.

Lemma 5.3.1 *For any minimizer μ_τ of (5.3.1) and any $\beta \in \Gamma_0(\mu,\mu_\tau)$ we have*

$$|y - x| = \operatorname{dist}(x,\partial\Omega) \qquad \textit{for } \beta\textit{-a.e. } (x,y) \in \overline{\Omega}\times\partial\Omega. \tag{5.3.2}$$

In addition, if $\lambda > 0$ and $\tau > 0$ is sufficiently small (depending only on Ω and λ), we have that $\beta(\partial\Omega\times\Omega) = 0$.

Proof. Let χ be the characteristic function of $\{(x,y)\in(\overline{\Omega},\partial\Omega):|y-x|>\mathrm{dist}(x,\partial\Omega)\}$. Let us choose (in a Borel way) for any $x\in\overline{\Omega}$ a point $s(x)\in\partial\Omega$ with shortest distance, set $\sigma:=(1-\chi)\beta$ and let ν and ϑ be respectively its first and second marginals; notice that $\nu\le\mu$ and that $\widehat{\vartheta}=\widehat{\mu}_\tau$, because $\chi(x,y)$ can be nonzero only when $y\in\partial\Omega$. Defining

$$\beta':=\sigma+(I,s)_\#(\mu-\nu)$$

and μ'_τ as the second marginal of β', we have that β' still has μ as first marginal, and therefore $\beta'\in\Gamma(\mu,\mu'_\tau)$. Taking into account that $\mu-\nu$ is the first marginal of $\chi\beta$, we also have

$$\begin{aligned}W_2^2(\mu,\mu'_\tau) &\le \int_{\overline{\Omega}\times\overline{\Omega}}|x-y|^2(1-\chi)\,d\beta+\int_{\overline{\Omega}}|s(x)-x|^2\,d(\mu-\nu)\\ &= W_2^2(\mu,\mu_\tau)+\int_{\overline{\Omega}\times\partial\Omega}(|s(x)-x|^2-|y-x|^2)\chi\,\beta\\ &\le W_2^2(\mu,\mu_\tau),\end{aligned}$$

with strict inequality whenever $\int\chi\,d\beta>0$. On the other hand, since the second marginal of $(I,s)_\#(\mu-\nu)$ is concentrated on $\partial\Omega$ we have $\widehat{\mu}'_\tau=\widehat{\vartheta}=\widehat{\mu}_\tau$, and therefore $\Phi_\lambda(\widehat{\mu}'_\tau)=\Phi_\lambda(\widehat{\mu}_\tau)$. The minimality of μ_τ then gives that (5.3.2) holds.

Let us prove that $\beta(\partial\Omega\times\Omega)=0$, for τ sufficiently small. Let $r>0$. By contradiction, suppose that $\pi^2_\#(\chi_{\partial\Omega\times\Omega}\beta)>0$. Suppose first that $\mathrm{supp}(\pi^2_\#(\chi_{\partial\Omega\times\Omega}\beta))$ is contained in the set $\{x\in\Omega:\mathrm{dist}\,(x\partial\Omega)\le r\}$. Hence the mass that comes from the boundary remains at a distance smaller than r from the boundary itself. Let $\nu=\mu_\tau$ minimize (5.3.1), starting from μ. Let ν_0 be a competitor, obtained starting from ν by $\widehat{\nu}_0=\widehat{\nu}-\pi^2_\#(\chi_{\partial\Omega\times\Omega}\beta)$ and $\widetilde{\nu}_0=\widetilde{\nu}+\pi^1_\#(\chi_{\partial\Omega\times\Omega}\beta)$. That is, we bring back the mass coming from μ through β. Choose r small enough such that $h_0-1+\lambda/2>\lambda/4$ for any x with $\mathrm{dist}\,(x,\partial\Omega)\le r$. This is possible since h_0 is continuous till the boundary (by elliptic regularity), where its value is 1. Next, notice that for any positive measure μ, $h_\mu=h_0+w_\mu$, where w_μ solves

$$\begin{cases}\Delta w_\mu+w_\mu=\mu & \text{in }\Omega,\\ w_\mu=0 & \text{on }\partial\Omega.\end{cases}$$

By the maximum principle, $w_\mu\ge0$, then $h_\mu\ge h_0$. We apply this inequality to ν_0, obtaining, as $\widehat{\nu}-\widehat{\nu}_0=\pi^2_\#(\chi_{\partial\Omega\times\Omega}\beta)$,

$$\begin{aligned}\frac{\lambda}{2}\nu(\Omega)-\frac{\lambda}{2}\nu_0(\Omega)+\int_\Omega(h_{\nu_0}-1)\,d(\nu-\nu_0) &\ge \frac{\lambda}{2}\nu(\Omega)-\frac{\lambda}{2}\nu_0(\Omega)-\frac{\lambda}{4}\int_\Omega d(\pi^2_\#(\chi_{\partial\Omega\times\Omega}\beta))\\ &= \frac{\lambda}{4}\pi^2_\#(\chi_{\partial\Omega\times\Omega}\beta)(\Omega).\end{aligned}$$

By (5.2.6) we immediately conclude that $\Phi_\lambda(\nu)\ge\Phi_\lambda(\nu_0)+\frac{\lambda}{4}\pi^2_\#(\chi_{\partial\Omega\times\Omega}\beta)(\Omega)$. Hence there is a gain of at least $\pi^2_\#(\chi_{\partial\Omega\times\Omega}\beta)(\Omega)\lambda/4$ in the Φ_λ term, when passing from ν to ν_0. On the other hand, if r is chosen small enough it is readily seen that the $W^2/(2\tau)$ term cannot compensate such loss. This is a contradiction, since ν is a minimizer. This shows that, for any $\tau > 0$, if an amount α of mass in $\partial\Omega$ is sent by β inside Ω, it has to be sent at a distance larger than r On the other hand, suppose that the competitor ν_0 has some mass transported from the boundary in the interior of the domain, but at greater distance. Then, let us compare the energy of ν with the energy of a new competitor ν_0, obtained by transporting back to μ the mass coming from the boundary, but being at distance grater than r from it. The term $W_2^2(\nu,\mu)$ decreases at least by αr^2, while (still by Lemma 5.2.3), the term $\Phi_\lambda(\nu)$ increases at most by $C\alpha$, with $C = \sup(h_0 - 1 + \lambda/2)^-$. By the inequality $\alpha r^2/(2\tau) \le C\alpha$ it follows that $\alpha = 0$ if $\tau < r^2/(2C)$. So that for τ small enough we have again a contradiction. We conclude that $\beta(\partial\Omega\times\Omega) = 0$. $\square$

In order to derive the Euler-Lagrange equation associated to (5.3.1), we will need a family of approximating variational problems in $\mathscr{P}(\overline{\Omega})$.

Lemma 5.3.2 *Let $\delta > 0$, let $\Phi_\lambda^\delta : \mathscr{P}(\overline{\Omega}) \to [0,+\infty]$ be defined by*

$$\begin{cases} \Phi_\lambda^\delta(\nu) = \Phi_\lambda(\nu) + \delta \displaystyle\int_\Omega \hat{\nu}^4 & \text{if } \widehat{\nu} \ll \mathcal{L}^2, \\ \qquad\qquad\qquad\qquad +\infty & \text{otherwise.} \end{cases} \tag{5.3.3}$$

Let us consider the minimization problem

$$\min_{\nu\in\mathscr{P}(\overline{\Omega})} \Phi_\lambda^\delta(\nu) + \frac{1}{2\tau}W_2^2(\mu,\nu). \tag{5.3.4}$$

Then this problem has a solution μ_τ^δ, the family μ_τ^δ has limit points both for the strong H^{-1} topology and the $\mathscr{P}(\overline{\Omega})$ narrow topology, $\delta\int_\Omega(\hat{\mu}_\tau^\delta)^4 \to 0$ as $\delta \to 0$, and any limit point μ_τ as $\delta \to 0$ solves (5.3.1).

Proof. The existence of μ_τ^δ is given by the direct method, as for the existence of μ_τ. Let M_δ be the minimum in (5.3.4) and let M be the minimum of the functional in (5.3.1). It is clear that $M_\delta \ge M$; on the other hand, $\Phi_\lambda^\delta \to \Phi_\lambda$ at any point ν such that $\widehat{\nu} \in L^4(\Omega)$. Then

$$\limsup_{\delta\downarrow 0} M_\delta \le \Phi_\lambda(\nu) + \frac{1}{2\tau}W_2^2(\nu,\mu)$$

for all ν in the subspace $\{\nu \in \mathscr{P}(\overline{\Omega}) : \widehat{\nu} \in L^4(\Omega)\}$. By density we obtain $\limsup_\delta M_\delta \le M$, therefore $M_\delta \to M$ as $\delta \to 0$.

If μ_τ is a weak limit point in $\mathscr{P}(\overline{\Omega})$ of μ_τ^δ along some sequence $\delta_i \to 0$, the lower semicontinuity of Φ_λ gives, since $\Phi_\lambda^{\delta_i} \ge \Phi_\lambda$ for any i,

$$\Phi_\lambda(\mu_\tau) + \frac{1}{2\tau}W_2^2(\mu_\tau,\mu) \le \liminf_{i\to\infty} \Phi_\lambda(\mu_\tau^{\delta_i}) + \frac{1}{2\tau}W_2^2(\mu_\tau^{\delta_i},\mu) \le \liminf_{i\to\infty} M_{\delta_i} = M,$$

therefore μ_τ is a solution of (5.3.1). As a consequence

$$\lim_{i\to\infty} \Phi_\lambda(\mu_\tau^{\delta_i}) + \delta_i \int_\Omega |\widehat{\mu}_\tau^{\delta_i}|^4 + \frac{1}{2\tau}W_2^2(\mu_\tau^{\delta_i},\mu) = \Phi_\lambda(\mu_\tau) + \frac{1}{2\tau}W_2^2(\mu_\tau,\mu).$$

This gives, taking into account the continuity of $\nu \mapsto W_2^2(\nu, \mu)$, proved in Lemma 2.2.7,

$$\limsup_{i\to\infty} \Phi_\lambda(\widehat{\mu}_\tau^{\delta_i}) + \delta_i \int_\Omega (\widehat{\mu}_\tau^{\delta_i})^4 \le \Phi_\lambda(\mu_\tau).$$

By the lower semicontinuity of Φ_λ it follows that $\Phi_\lambda(\mu_\tau^{\delta_i}) \to \Phi_\lambda(\mu_\tau)$ and $\delta_i \int_\Omega (\widehat{\mu}_\tau^{\delta_i})^4 \to 0$. Now, since $\Phi_\lambda(\nu)$ is the sum of two lower semicontinuous terms, namely $\Phi_0(\nu)$ and $\lambda\nu(\Omega)/2$, we obtain

$$\lim_{i\to\infty} \lambda\mu_\tau^{\delta_i}(\Omega) = \lambda\mu_\tau(\Omega) \quad \text{and} \quad \lim_{i\to\infty} \int_\Omega |\nabla h_{\mu_\tau^{\delta_i}}|^2 + (h_{\mu_\tau^{\delta_i}} - 1)^2 = \int_\Omega |\nabla h_{\mu_\tau}|^2 + (h_{\mu_\tau} - 1)^2. \tag{5.3.5}$$

In particular $\widehat{\mu}_\tau^{\delta_i} \to \widehat{\mu}_\tau$ strongly in $H^{-1}(\Omega)$. □

The next proposition provides a Euler-Lagrange equation for (5.3.4).

Proposition 5.3.3 *Any minimizer μ_τ^δ of (5.3.4) satisfies*

$$-3\delta\nabla((\widehat{\mu}_\tau^\delta)^4) - \nabla h_{\mu_\tau^\delta}\widehat{\mu}_\tau^\delta = \frac{1}{\tau}\pi^1_{\#}\left(\chi_\Omega(x)(x-y)\gamma\right) \quad \text{in } \mathcal{D}'(\mathbb{R}^2), \tag{5.3.6}$$

where $\gamma \in \Gamma_0(\mu_\tau^\delta, \mu)$.

We omit the proof, since it is a particular case of Lemma 5.7.2, that we shall prove later.

Entropy and regularity

The next results shows that $\int_\Omega \varphi(\widehat{\mu})$ decreases along the discrete flow, for a suitable family of *entropies* φ. Let φ be a C^1 function from $[0, +\infty)$ to $\mathbb{R}$. We say that φ is an *entropy* if:

(a) φ is convex nondecreasing, C^1, and $x\varphi'(x) = \varphi(x)$ for all $x \in [0, 1]$;

(b) $2x\varphi''(x) \ge x\varphi'(x) - \varphi(x)$ (displacement convexity).

Proposition 5.3.4 *Let φ be an entropy and let $\mu \in \mathscr{P}(\overline{\Omega})$ be such that $\int_\Omega \varphi(\widehat{\mu}) < \infty$. Then, for any minimizer μ_τ^δ of (5.3.4), we have*

$$\int_\Omega \varphi(\widehat{\mu}_\tau^\delta) \le \int_\Omega \varphi(\widehat{\mu}).$$

Proposition 5.3.5 *Let $p \in (1, \infty]$, $\mu \in \mathscr{P}(\overline{\Omega})$ with $\widehat{\mu} \in L^p(\Omega)$. Then there exists a minimizer μ_τ of (5.3.1) satisfying* Proposition 5.3.6 *and*

i) If $p < \infty$, there exists a p-growing entropy φ

such that $\int_\Omega \varphi(\widehat{\mu}_\tau) \le \int_\Omega \varphi(\widehat{\mu}) < \infty$, hence $\widehat{\mu}_\tau \in L^p(\Omega)$.

ii) If $p = \infty$ and $M = \max\{1, \|\widehat{\mu}\|_\infty\}$, we have $\|\widehat{\mu}_\tau\|_\infty \le M$, $0 \le h_{\mu_\tau} \le M$.

For the proof of the last two propositions we refer to the more general results of Section 5.8.

Now we pass to the limit as $\delta \to 0$ in (5.3.6). We are led to the actual Euler equation.

Proposition 5.3.6 *There exist a minimizer μ_τ of* (5.3.1) *and $\gamma \in \Gamma_0(\mu_\tau, \mu)$ satisfying*

$$-\nabla h_{\mu_\tau} \widehat{\mu}_\tau = \frac{1}{\tau} \pi^1_\# (\chi_\Omega(x)(x-y)\gamma) \quad \text{in } \mathcal{D}'(\mathbb{R}^2). \tag{5.3.7}$$

Proof. We consider a narrow (in $\mathscr{P}(\overline{\Omega})$) limit point $\mu_\tau = \lim_i \mu_\tau^{\delta_i}$ as in Lemma 5.3.2, and consider plans $\gamma_i \in \Gamma_0(\mu_\tau^{\delta_i}, \mu)$. Moreover, Lemma 5.3.2 gives $\delta_i \int_\Omega (\widehat{\mu}_\tau^{\delta_i})^4 \to 0$ and $\widehat{\mu}_\tau^{\delta_i}(\Omega) \to \widehat{\mu}_\tau(\Omega)$. Let η be a smooth vector field. By (5.3.6) we have

$$\frac{1}{\tau} \int_{\Omega \times \overline{\Omega}} (x-y) \cdot \eta(x) \, d\gamma_i(x,y) + \int_\Omega \nabla h_i(x) \cdot \eta(x) \, d\widehat{\mu}_\tau^{\delta_i} - 3\delta_i \int_\Omega (\widehat{\mu}_\tau^{\delta_i})^4 \mathrm{div}\, \xi = 0, \tag{5.3.8}$$

where $h_i := h_{\mu_\tau^{\delta_i}}$. Next, we pass to the limit in (5.3.8) as $i \to \infty$, possibly along a subsequence: we have that $\delta_i \int_\Omega (\widehat{\mu}_\tau^{\delta_i})^4 \mathrm{div}\, \xi \to 0$. By Lemma 2.2.6, γ_i weakly converges in $\mathscr{P}(\overline{\Omega} \times \overline{\Omega})$ to some $\gamma \in \Gamma_0(\mu_\tau, \mu)$, and the fact that $\mu_\tau^{\delta_i}(\Omega)$ converges to $\mu_\tau(\Omega)$ tells us that there is no concentration of mass near $\partial\Omega$; using this fact we easily obtain that $\chi_\Omega(x)\gamma_i$ still weakly converges to $\chi_\Omega(x)\gamma$, so that also the first term in the left-hand side goes to the limit. Finally, by Proposition 5.3.4 $\widehat{\mu}_\tau^{\delta_i}$ is bounded in $L^4(\Omega)$, hence we find a subsequence (that we don't relabel) weakly converging to μ_τ in $L^4(\Omega)$. By elliptic regularity, this gives strong $W^{2,4}(\Omega)$ convergence (and then $C^1(\overline{\Omega})$ convergence by Sobolev immersion, as the boundary is smooth) of h_i to h_{μ_τ}. This readily entails convergence for the second term in (5.3.8). □

Making use of Theorem 3.4.4, one obtains the following result about existence of solution and regularity of its internal part (see Theorem 1.1 in [8]). The only difference is that here the discrete velocity is written in terms of transport plans, as in the right hand side of (5.3.7). In fact, a limit curve exists by the construction of Theorem 3.4.4. It satisfies the continuity equation, coupled with the limiting velocity $\mathbf{v}_t$, which is the limit of the discrete velocities. But (5.3.7) allows to characterize such limit as $\chi_\Omega \nabla h_{\mu(t)}$. Such velocity satisfies also the relaxed gradient flow relation (3.4.5). In the end, we get the following

Theorem 5.3.7 *Assume that $\widehat{\mu}^0 \in L^p$ for some $p \geq 4/3$ and $\lambda \geq 0$. Then there exists an absolutely continuous curve $t \mapsto \mu_t \in P(\overline{\Omega})$ such that:*

i) $\|\widehat{\mu}_t\|_p \leq C$, *with C depending only on $\widehat{\mu}^0$;*

ii) $\mu_0 = \mu^0$ *and the PDE*

$$\partial_t \mu_t - \mathrm{div}\,(\chi_\Omega \nabla h_{\mu_t} \mu_t) = 0 \qquad \text{in } \mathcal{D}'((0,+\infty) \times \mathbb{R}^2) \tag{5.3.9}$$

holds.

5.4 New variation

Consider problem (5.3.1). In this section we are going to prove the analogous of (5.2.4) at the discrete level. For, we will consider a boundary variation for the minimization problem, that is, the competitor of $\nu \in \mathscr{P}(\overline{\Omega})$ will be of the form

$$\nu_\alpha := \widehat{\nu} + \alpha T_\# \sigma + (1-\alpha)\widetilde{\nu}.$$

So, we are leaving steady the interior part, while we are transporting a fraction α of $\widetilde{\nu}$ to some diffused measure σ over Ω through a map T. We will be able to derive information on the orientation of ∇h_ν at the boundary.

We need a measure theoretic lemma before proceeding with the main proof. We recall also that, given two measures μ and ν in $\mathcal{M}_+(\mathbb{R}^2)$ with same mass, if μ is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}^2$, then there exists a unique optimal transport plan between μ and ν (for which the infimum in (5.3.1) is achieved), and such a plan is induced by a transport map (see Theorem 2.1.7 in Chapter 2).

Lemma 5.4.1 *Let $\mu, \nu \in \mathscr{P}(\overline{\Omega})$, $\sigma \ll \mathcal{L}^2 \llcorner \Omega$, with $\sigma(\Omega) = \nu(\partial\Omega)$, and let T be the optimal transport map between σ and $\widetilde{\nu}$.*
Then there exist $\gamma \in \Gamma_0(\nu,\mu)$, $\gamma_T \in \Gamma(\sigma, \mu_1)$, where μ_1 is the second marginal of $\chi_{\partial\Omega\times\overline{\Omega}}\gamma$, such that

$$W_2^2(\nu_S, \mu) - W_2^2(\nu,\mu) \le \int_{\Omega\times\overline{\Omega}} \left[|y - S(x)|^2 - |y - T(x)|^2\right] d\gamma_T(x,y)$$

for all $S : \Omega \mapsto \Omega$, where $\nu_S = \widehat{\nu} + S_\# \sigma$.

Proof. Let us introduce a sequence of auxiliary measures $\widetilde{\nu}_n$, with equicompact supports contained in $\mathbb{R}^2 \backslash \Omega$, such that $\widetilde{\nu}_n(\mathbb{R}^2\backslash\Omega) = \sigma(\Omega)$, $\widetilde{\nu}_n \ll \mathcal{L}^2$ and $\widetilde{\nu}_n \rightharpoonup \widetilde{\nu}$ as $n \to \infty$. Let T_n be the optimal transport maps between σ and $\widetilde{\nu}_n$. Moreover, let γ_n be optimal transport plans between ν_n and μ, where $\nu_n = \widehat{\nu} + \widetilde{\nu}_n$. As an optimal transport map between absolutely continuous measures, T_n is essentially invertible for every n (i.e. its restriction to the complement of a σ-negligible set in Ω is injective, see Lemma 2.1.9). So we can define

$$\begin{aligned}
\overline{\gamma}_n &= (S \circ T_n^{-1}, I)_\# \chi_{(\mathbb{R}^2\backslash\Omega)\times\overline{\Omega}}\gamma_n + \chi_{\Omega\times\overline{\Omega}}\gamma_n, \\
\gamma_{T_n} &= (T_n^{-1}, I)_\# \chi_{(\mathbb{R}^2\backslash\Omega)\times\overline{\Omega}}\gamma_n.
\end{aligned}$$

Clearly, $\overline{\gamma}_n \in \Gamma(\nu_S, \mu)$ and $\gamma_{T_n} \in \Gamma(\sigma, \mu_n)$ for every n, where we introduced μ_n as the second marginal of $\chi_{(\mathbb{R}^2\backslash\Omega)\times\overline{\Omega}}\gamma_n$. So with the change of variables $z = T_n^{-1}(x)$, for every n we have

$$\begin{aligned}
W_2^2(\nu_S,\mu) &\le \int_{\overline{\Omega}\times\overline{\Omega}} |y-x|^2 d\overline{\gamma}_n \\
&= \int_{(\mathbb{R}^2\backslash\Omega)\times\overline{\Omega}} |y-x|^2 d((S\circ T_n^{-1}, I)_\#\gamma_n) + \int_{\Omega\times\overline{\Omega}} |y-x|^2 d\gamma_n \\
&= \int_{\Omega\times\overline{\Omega}} |y - S(z)|^2 d\gamma_{T_n}(z,y) + \int_{\Omega\times\overline{\Omega}} |y-x|^2 d\gamma_n,
\end{aligned}$$

and

$$W_2^2(\nu_n,\mu) = \int_{(\mathbb{R}^2\setminus\Omega)\times\overline{\Omega}} |y-x|^2 d\gamma_n + \int_{\Omega\times\overline{\Omega}} |y-x|^2 d\gamma_n$$
$$= \int_{\Omega\times\overline{\Omega}} |y-T_n(z)|^2 d\gamma_{T_n}(z,y) + \int_{\Omega\times\overline{\Omega}} |y-x|^2 d\gamma_n. \tag{5.4.1}$$

We get, for every n,

$$W_2^2(\nu_S,\mu) - W_2^2(\nu_n,\mu) \le \int_{\Omega\times\overline{\Omega}} \left[|y-S(x)|^2 - |y-T_n(x)|^2\right] d\gamma_{T_n}(x,y). \tag{5.4.2}$$

Now we have to pass to the limit as $n\to\infty$. As $\widetilde{\nu}_n \rightharpoonup \widetilde{\nu}$, for the stability property of optimal transport maps, we have that $T_n \to T$ strongly in $L^p(\Omega,\sigma)$, $1\le p<\infty$, where T is the optimal transport map between σ and $\widetilde{\nu}$. Moreover, γ_n has a weak limit point in $\mathscr{P}(\mathbb{R}^2\times\overline{\Omega})$ which is an optimal plan $\gamma\in\Gamma_0(\nu,\mu)$ (see Lemma 2.2.6). We will not relabel the sequence for simplicity.
We can also show that

$$\chi_{\Omega\times\overline{\Omega}}\gamma_n \rightharpoonup \chi_{\Omega\times\overline{\Omega}}\gamma. \tag{5.4.3}$$

In fact, let $\eta(x)$ be a smooth cutoff function approximating χ_Ω, with $\eta(x)\equiv 0$ on $\mathbb{R}^2\setminus\Omega$ and

$$\int_\Omega |\eta(x)-1|d\widehat{\nu} < \varepsilon.$$

Let $f\in C^0(\mathbb{R}^2\times\overline{\Omega})$, with $M=\|f\|_\infty$ finite. Then

$$\begin{aligned}
&\int_{\mathbb{R}^2\times\overline{\Omega}} f(x,y)\chi_{\Omega\times\overline{\Omega}}(x,y)d(\gamma_n-\gamma)(x,y)\\
&= \int_{\mathbb{R}^2\times\overline{\Omega}} f(x,y)\left[\chi_\Omega(x)-\eta(x)+\eta(x)\right]d(\gamma_n-\gamma)(x,y)\\
&\le \int_{\mathbb{R}^2\times\overline{\Omega}} f(x,y)\eta(x)d(\gamma_n-\gamma)(x,y) + M\int_{\mathbb{R}^2\times\overline{\Omega}} |\chi_\Omega(x)-\eta(x)|d(\gamma_n+\gamma)(x,y)\\
&= \int_{\mathbb{R}^2\times\overline{\Omega}} f(x,y)\eta(x)d(\gamma_n-\gamma)(x,y) + 2M\int_\Omega |1-\eta(x)|d\widehat{\nu}\\
&\le \int_{\mathbb{R}^2\times\overline{\Omega}} f(x,y)\eta(x)d(\gamma_n-\gamma)(x,y) + 2M\varepsilon.
\end{aligned}$$

Now the first integral tends to zero, since $f\eta$ is continuous, and by arbitrariness of ε we get the convergence. Here we used the fact that the measures $\chi_{\Omega\times\overline{\Omega}}\gamma_n$ and $\chi_{\Omega\times\overline{\Omega}}\gamma$ have $\widehat{\nu}$ as first marginal. In the same way one can prove that $\chi_{(\mathbb{R}^2\setminus\Omega)\times\overline{\Omega}}\gamma_n \rightharpoonup \chi_{(\mathbb{R}^2\setminus\Omega)\times\overline{\Omega}}\gamma$. This implies that, letting $\mu_n := \pi^2_\#(\chi_{(\mathbb{R}^2\setminus\Omega)\times\overline{\Omega}}\gamma_n)$ be the second marginal of $\chi_{(\mathbb{R}^2\setminus\Omega)\times\overline{\Omega}}\gamma_n$, there holds $\mu_n \rightharpoonup \mu_1$. μ_n is also the second marginal of γ_{T_n} which by tightness has a limit point γ_T (again we avoid relabeling the sequence). The first marginal of γ_{T_n} is σ for every n, and as a consequence $\gamma_T\in\Gamma(\sigma,\mu_1)$.

Now consider the first integral in the second member of (5.4.2). We have the weak convergence of γ_{T_n} to γ_T, and we can pass to the limit even though the integrand is not continuous. Indeed, reasoning exactly as in the proof of (5.4.3), we can approximate it with continuous functions (in the Lusin sense) and use the fact that both the first marginal of γ_{T_n} and of γ_T are equal to the absolutely continuous measure σ.
Finally, consider the last term in (5.4.2). We have

$$\begin{aligned}\int_{\Omega\times\overline{\Omega}} |y-T_n|^2 d\gamma_{T_n} &= \int_{\Omega\times\overline{\Omega}} \left[|y-T|^2 + |y-T_n|^2 - |y-T|^2\right] d\gamma_{T_n} \\ &\le \int_{\Omega\times\overline{\Omega}} |y-T|^2 d\gamma_{T_n} + K\int_{\Omega\times\overline{\Omega}} |T_n(x)-T(x)| d\gamma_{T_n} \\ &\le \int_{\Omega\times\overline{\Omega}} |y-T|^2 d\gamma_{T_n} + K\int_{\Omega} |T_n(x)-T(x)| d\sigma,\end{aligned}$$

with K being a suitable positive constant depending on Ω. Now the second term goes to zero for the strong convergence of T_n, and the first one can be treated as before and shown to converge to

$$\int_{\Omega\times\overline{\Omega}} |y-T(x)|^2 d\gamma_T(x,y).$$

We have all what is needed to pass to the limit in (5.4.1) and (5.4.2) and obtain

$$W_2^2(\nu_S,\mu) - W_2^2(\nu,\mu) \le \int_{\Omega\times\overline{\Omega}} \left[|y-S(x)|^2 - |y-T(x)|^2\right] d\gamma_T(x,y) \tag{5.4.4}$$

as desired. □

We also state a slight generalization of the previous lemma.

Lemma 5.4.2 *Let μ,ν,σ and T be as in* Lemma 5.4.1. *Let $S:\Omega\mapsto\Omega$, $\theta\in[0,1]$ and*

$$\nu_S = \widehat{\nu} + \theta S_{\#}\sigma + (1-\theta)\widetilde{\nu}.$$

Then there exist $\gamma\in\Gamma_0(\nu,\mu)$, $\gamma_T\in\Gamma(\sigma,\mu_1)$, μ_1 being the second marginal of $\chi_{\partial\Omega\times\overline{\Omega}}\gamma$, such that

$$W_2^2(\nu_S,\mu) - W_2^2(\nu,\mu) \le \theta\int_{\Omega\times\overline{\Omega}} \left[|y-S(x)|^2 - |y-T(x)|^2\right] d\gamma_T(x,y).$$

Proof. The case $\theta=0$ is trivial. Otherwise, define $\widetilde{\nu}_n,\nu_n,T_n,\gamma_n$ and γ_{T_n} as in the proof of Lemma 5.4.1. Moreover, let $\nu_S^n = \widehat{\nu} + S_{\#}(\theta\sigma) + (1-\theta)\widetilde{\nu}_n$ and introduce transport plans $\overline{\overline{\gamma}}_n\in\Gamma(\nu_S^n,\mu)$ as follows:

$$\overline{\overline{\gamma}}_n = \theta(S\circ T_n^{-1}, I)_{\#}\chi_{(\mathbb{R}^2\setminus\Omega)\times\overline{\Omega}}\gamma_n + \chi_{\Omega\times\overline{\Omega}}\gamma_n + (1-\theta)\chi_{(\mathbb{R}^2\setminus\Omega)\times\overline{\Omega}}\gamma_n.$$

Then, with the change of variables $z=T_n^{-1}(x)$, we have

$$\begin{aligned}W_2^2(\nu_S^n,\mu) \le \int_{\overline{\Omega}\times\overline{\Omega}} |y-x|^2 d\overline{\overline{\gamma}}_n &= \theta\int_{\Omega\times\overline{\Omega}} |y-S(z)|^2 d\gamma_{T_n}(z,y) + \int_{\Omega\times\overline{\Omega}} |y-x|^2 d\gamma_n \\ &+ (1-\theta)\int_{(\mathbb{R}^2\setminus\Omega)\times\overline{\Omega}} |y-x|^2 d\gamma_n.\end{aligned}$$

We can rewrite (5.4.1) as

$$\begin{aligned} W_2^2(\nu_n,\mu) &= \theta \int_{(\mathbb{R}^2\setminus\Omega)\times\overline{\Omega}} |y-x|^2 d\gamma_n + \int_{\Omega\times\overline{\Omega}} |y-x|^2 d\gamma_n + (1-\theta)\int_{(\mathbb{R}^2\setminus\Omega)\times\overline{\Omega}} |y-x|^2 d\gamma_n \\ &= \theta \int_{\Omega\times\overline{\Omega}} |y-T_n(z)|^2 d\gamma_{T_n}(z,y) + \int_{\Omega\times\overline{\Omega}} |y-x|^2 d\gamma_n \\ &\quad + (1-\theta)\int_{(\mathbb{R}^2\setminus\Omega)\times\overline{\Omega}} |y-x|^2 d\gamma_n. \end{aligned}$$

This way, it is clear that

$$W_2^2(\nu_S^n,\mu) - W_2^2(\nu_n,\mu) \le \theta \int_{\Omega\times\overline{\Omega}} \left[|y-S(x)|^2 - |y-T_n(x)|^2\right] d\gamma_{T_n}(x,y).$$

Here we can pass to the limit in n exactly as done for (5.4.2), so we refer to the proof of Lemma 5.4.1 for the conclusion. The only element to add is the lower semicontinuity of W_2 for treating the first term, so that

$$W_2(\nu_S,\mu) \le \liminf_{n\to\infty} W_2(\nu_S^n,\mu)$$

as $\nu_S^n \rightharpoonup \nu_S$. □

Remark 5.4.3 With minor modifications one can also obtain the same result for the case

$$\nu_S = \widehat{\nu} + \theta S_\# \sigma + (1-\theta)\chi_A \widetilde{\nu} + \chi_{\partial\Omega\setminus A}\widetilde{\nu},$$

where A is an arc contained in $\partial\Omega$. In this case we have $\gamma \in \Gamma_0(\nu,\mu)$, $\sigma \ll \mathcal{L}^2 \llcorner \Omega$, $\sigma(\Omega) = \widetilde{\nu}(A)$, $(I,T)_\#\sigma \in \Gamma_0(\sigma, \chi_A\widetilde{\nu})$. μ_1 will be a suitable measure such that $\mu_1 \le \pi^2_\#(\chi_{\partial\Omega\times\overline{\Omega}}\gamma)$.

Taking into account also Proposition 5.3.5, we are ready to state and prove the main result about discrete minimizers.

Theorem 5.4.4 *Let $\nu = \mu_\tau$ be a minimizer of* (5.3.1)*, with $\lambda = 0$, such that $\widehat{\nu} \in L^4(\Omega)$. Let Ω be convex. Then*

$$\langle \nabla h_\nu(x), y-x\rangle \ge 0 \quad \forall (x,y) \in \operatorname{supp}(\widetilde{\nu}) \times \overline{\Omega}. \tag{5.4.5}$$

Proof.
Let $\sigma \ll \mathcal{L}^2 \llcorner \Omega$ have a bounded density, and let $\sigma(\Omega) = \widetilde{\nu}(\overline{\Omega})$. Let moreover T be the optimal transport map between σ and $\widetilde{\nu}$, and

$$T_\varepsilon = (1-\varepsilon)I + \varepsilon T, \quad \varepsilon \in [0,1].$$

We introduce the following perturbed measure

$$\nu_\varepsilon := \widehat{\nu} + \alpha^2 T_{\varepsilon\#}\sigma + (1-\alpha^2)\widetilde{\nu},$$

where $\alpha = (1-\varepsilon)^2$.
Now we apply Lemma 5.4.1, with T_ε in the role of S: there exist a transport plan $\gamma \in \Gamma_0(\nu,\mu)$ and a transport plan $\gamma_T \in \Gamma(\sigma,\mu_1)$, where μ_1 is the second marginal of $\chi_{\partial\Omega\times\overline{\Omega}}\gamma$, such that

$$W_2^2(\nu_\varepsilon,\mu) - W_2^2(\nu,\mu) \le \alpha^2 \int_{\Omega\times\overline{\Omega}} \left[|y-T_\varepsilon(x)|^2 - |y-T(x)|^2\right] d\gamma_T(x,y). \tag{5.4.6}$$

Next, we apply (5.2.9) to ν_ε and ν, and we find

$$\Phi_0(\nu_\varepsilon) - \Phi_0(\nu) = -(\widehat{\nu}_\varepsilon(\Omega) - \widehat{\nu}(\Omega)) + \frac{1}{2}\int_\Omega (h_{\nu_\varepsilon} + h_\nu) d(\widehat{\nu}_\varepsilon - \widehat{\nu}),$$

so that

$$\Phi_0(\nu_\varepsilon) - \Phi_0(\nu) = -\alpha^2 \widetilde{\nu}(\partial\Omega) + \frac{1}{2}\alpha^2 \int_\Omega (h_{\nu_\varepsilon} + h_\nu) d(T_{\varepsilon\#}\sigma). \tag{5.4.7}$$

Since ν is a minimizer, there holds

$$\Phi_0(\nu_\varepsilon) - \Phi_0(\nu) + \frac{1}{2\tau}\left(W_2^2(\nu_\varepsilon,\mu) - W_2^2(\nu,\mu)\right) \ge 0,$$

for all $\mu \in \mathscr{P}(\overline{\Omega})$. Substituting (5.4.6) and (5.4.7) in this inequality, we obtain:

$$\frac{\alpha^2}{2\tau}\int_{\Omega\times\overline{\Omega}} \left[|y-T_\varepsilon(x)|^2 - |y-T(x)|^2\right] d\gamma_T - \alpha^2\widetilde{\nu}(\partial\Omega) + \frac{1}{2}\alpha^2\int_\Omega (h_{\nu_\varepsilon} + h_\nu) d(T_{\varepsilon\#}\sigma) \ge 0. \tag{5.4.8}$$

Since $T_\varepsilon = T + (1-\varepsilon)(I-T)$, we obtain the following expansion (of the first order centered in $\varepsilon = 1$)

$$|y-T_\varepsilon(x)|^2 = |y-T(x)|^2 + 2(\varepsilon-1)\langle y-T(x), x-T(x)\rangle + o(\varepsilon-1).$$

Of course the remainder is uniformly bounded with respect to $x \in \overline{\Omega}$. For treating the second integral in (5.4.8), notice that, as $\widehat{\nu} \in L^4(\Omega)$, $h_\nu \in W^{2,4}(\Omega)$, and by Sobolev embedding $h_\nu \in C^1(\overline{\Omega})$ (since Ω has smooth boundary). So we can perform the expansion

$$h_\nu \circ T_\varepsilon = h_\nu \circ T + (\varepsilon-1)\langle \nabla h_\nu \circ T, T-I\rangle + (\varepsilon-1)\langle(\nabla h_\nu \circ T_\theta - \nabla h_\nu \circ T), T-I\rangle, \tag{5.4.9}$$

for a suitable $\theta \in (\varepsilon, 1)$. If $K = \sup_{x\in\overline{\Omega}} |T(x) - x|$, the last term is bounded by $K(\varepsilon - 1)\,\omega(|T_\theta(x) - T(x)|)$, $\omega(\delta)$ being the modulus of continuity of ∇h_ν, which, as $\delta \to 0$, goes to zero uniformly with respect to $x \in \overline{\Omega}$, since $\nabla h_\nu \in C^0(\overline{\Omega})$. So there holds

$$h_\nu \circ T_\varepsilon = h_\nu \circ T + (\varepsilon-1)\langle \nabla h_\nu \circ T, T-I\rangle + o(\varepsilon-1), \tag{5.4.10}$$

and the remainder is uniform in x.
Finally, since

$$h_{\nu_\varepsilon} \circ T_\varepsilon = h_\nu \circ T_\varepsilon + (h_{\nu_\varepsilon} - h_\nu) \circ T_\varepsilon, \tag{5.4.11}$$

we have to estimate $h_{\nu_\varepsilon} - h_\nu$. This quantity is solution of the problem

$$\begin{cases} -\Delta u + u = \alpha^2 T_{\varepsilon\#}\sigma & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

Hence we can write

$$\sup_{x\in\overline{\Omega}} |h_{\nu_\varepsilon}(x) - h_\nu(x)| = \alpha^2 \sup_{x\in\overline{\Omega}} |\varphi_\varepsilon|, \tag{5.4.12}$$

where φ_ε satisfies

$$\begin{cases} -\Delta \varphi_\varepsilon + \varphi_\varepsilon = T_{\varepsilon\#}\sigma & \text{in } \Omega \\ \varphi_\varepsilon = 0 & \text{on } \partial\Omega. \end{cases}$$

But since for $\varepsilon \in (0,1)$ there holds $|\det(JT_\varepsilon)| \geq (1-\varepsilon)^2$ (see [4, §5.5], also for the push-forward change of variables formula), we have

$$\int_\Omega |\alpha T_{\varepsilon\#}\sigma|^4 = \alpha^4 \int_\Omega \left(\frac{\sigma}{|\det(JT_\varepsilon)|}\right)^4 |\det(JT_\varepsilon)| \leq \frac{\alpha^4}{(1-\varepsilon)^6} \int_\Omega |\sigma|^4 = (1-\varepsilon)^2 \int_\Omega |\sigma|^4.$$

Therefore $\alpha T_{\varepsilon\#}(\sigma)$ converges to 0 in $L^4(\Omega)$. This implies the $W^{2,4}(\Omega)$ convergence and the $C^1(\overline{\Omega})$ convergence of $\alpha\varphi_\varepsilon$ as $\varepsilon \to 1$. So there exists a constant C which bounds $\alpha\varphi_\varepsilon$ uniformly in x and ε, and from (5.4.12) we get

$$\sup_{x\in\overline{\Omega}} |h_{\nu_\varepsilon}(x) - h_\nu(x)| \leq C\alpha = C(1-\varepsilon)^2. \tag{5.4.13}$$

Making use of (5.4.10) and (5.4.13), from (5.4.11) we find

$$h_{\nu_\varepsilon} \circ T_\varepsilon = h_\nu \circ T + (\varepsilon - 1)\langle \nabla h_\nu \circ T, T - I\rangle + o(\varepsilon - 1), \tag{5.4.14}$$

where the remainder is again uniformly bounded in x.

Now, dividing by α^2, we expand to the first order in (5.4.8) with respect to $\varepsilon \to 1$, and with τ fixed, to find

$$\begin{aligned} &\frac{1-\varepsilon}{\tau} \int_{\Omega\times\overline{\Omega}} \langle y - T(x), T(x) - x\rangle d\gamma_T - \widetilde{\nu}(\partial\Omega) + \int_\Omega h_\nu(T(x))d\sigma \\ &+ (1-\varepsilon)\int_\Omega \langle \nabla h_\nu(T(x)), x - T(x)\rangle d\sigma + o(1-\varepsilon) \geq 0. \end{aligned}$$

As a consequence, since $\sigma(\Omega) = \widetilde{\nu}(\partial\Omega)$ and $h_\nu = 1$ on $\partial\Omega$, upon dividing by $(1-\varepsilon)$ we get

$$\frac{1}{\tau}\int_{\Omega\times\overline{\Omega}} \langle y - T(x), T(x) - x\rangle d\gamma_T + \int_\Omega \langle \nabla h_\nu(T(x)), x - T(x)\rangle d\sigma \geq 0.$$

As $T(x) \in \operatorname{supp}(\widetilde{\nu})$, in the first integral the scalar product is nonpositive for geometric reasons (we are working with a convex domain). It follows that

$$\int_\Omega \langle \nabla h_\nu(T(x)), x - T(x)\rangle d\sigma \geq 0. \tag{5.4.15}$$

Let $A \subset \partial\Omega$ be an arc such that $\widetilde{\nu}(A) > 0$. We point out that, redefining ν_ε as $\widehat{\nu} + \alpha^2 T_{\varepsilon\#}\sigma + (1-\alpha^2)\chi_A\widetilde{\nu} + \chi_{\partial\Omega\setminus A}\widetilde{\nu}$, with $T_\varepsilon = (1-\varepsilon)I + \varepsilon T$ and T now being the optimal transport map between an absolutely continuous σ and $\chi_A\widetilde{\nu}$, this proof works in the same way. Indeed, in view of Remark 5.4.3, inequality (5.4.6) still holds for some $\gamma_T \in \Gamma(\sigma, \mu_1)$, where $\mu_1 \leq \pi^2_\#(\chi_{\partial\Omega\times\overline{\Omega}}\gamma)$. So we obtain (5.4.15) with $T(x)$ taking values in $\mathrm{supp}(\widetilde{\nu}) \cap A$. Now, suppose by contradiction that

$$\langle \nabla h_\nu(\bar{z}), \bar{y} - \bar{z} \rangle < 0$$

for some $(\bar{z}, \bar{y}) \in \mathrm{supp}(\widetilde{\nu}) \times \overline{\Omega}$. Then, recalling that $\nabla h_\nu \in C^0(\overline{\Omega})$, there exist an arc $I \subset \partial\Omega$ containing $\bar{z}$ and a neighborhood Q of $\bar{y}$ such that the same inequality holds whenever $(z, y) \in I \times Q$. Because of the arbitrariness of A and σ, we can choose σ supported in $\Omega \cap Q$ and $A \subset I$. Since T transports σ to $\chi_A\widetilde{\nu}$, this implies $\langle \nabla h_\nu(T(x)), x - T(x) \rangle < 0$ for all $x \in \mathrm{supp}(\sigma)$, against (5.4.15). □

5.5 Uniqueness of the regular gradient flow

We now consider the problem of uniqueness of solutions for (5.2.1)-(5.2.2) in the case of measures with L^∞ internal part. Taking into account the result of Theorem 5.4.4, we focus on the following class of solutions.

Definition 5.5.1 (Regular gradient flow) *Let $T > 0$. A solution of problem* (5.2.1)-(5.2.2) *is a regular gradient flow if*

i) $\|\widehat{\mu}(t)\|_\infty \in L^\infty(0, T)$,

ii) $\langle \nabla h_{\mu(t)}(x), y - x \rangle \geq 0$ *for all* $(x, y) \in \mathrm{supp}(\widetilde{\mu}(t)) \times \overline{\Omega}$ *and* $t \in (0, T]$.

Remark 5.5.2 Condition *ii)* is related, as previously noticed, to the one appearing in [8, Definition 3.1], that is, $t \mapsto \widetilde{\mu}(t)$ is nondecreasing as a measure valued map. In fact, if the negative gradient at the boundary (that is the limit of velocities in Ω) is directed towards the exterior of the domain, we expect that no mass can move from $\partial\Omega$ to Ω during the evolution. Such a behavior was argued in [8] in the case $\lambda > 0$ by means of direct energy arguments (see Lemma 5.3.1), which do not extend for $\lambda = 0$. Actually, condition *ii)*, obtained in Theorem 5.4.4 only for $\lambda = 0$, will allow us to obtain a stronger uniqueness result.

Theorem 5.5.3 (Construction of a regular gradient flow) *Let Ω be convex. Let $\mu_0 \in \mathscr{P}(\overline{\Omega})$, with $\widehat{\mu}_0 \in L^\infty(\Omega)$. Then there exists a solution to problem* (5.2.1)-(5.2.2) *which is a regular gradient flow.*

Proof. The construction follows the one of Theorem 3.4.4. Let $\mu^0_\tau := \mu_0$. We find μ^{k+1}_τ solving (5.7.4) with $\lambda = 0$ recursively. We then define $\bar{\mu}_\tau(t)$ in the usual way, as in (3.2.4).

For $\tau \downarrow 0$, by Theorem 3.4.4 we can find limit points, that is, we can find sequences $\tau_n \downarrow 0$ such that weakly in the sense of measures

$$\lim_{n\to\infty} \bar{\mu}_{\tau_n}(t) = \mu(t) \qquad \forall t \geq 0. \tag{5.5.1}$$

Thanks to Lemma 5.3.5, the interior parts of all the discrete minimizers will belong to L^∞. Letting $T > 0$, and passing to the limit in τ, we will have $\widehat{\mu}(t) \in L^\infty((0,T); L^\infty(\Omega))$. Moreover, after Theorem 5.4.4, the discrete minimizers can also be chosen to satisfy (5.4.5), which, passing again to the limit in τ, becomes condition *ii)* of Definition 5.5.1. In fact, as a consequence of (5.5.1), with the regularity of the interior parts of $\bar{\mu}_{\tau_n}$, by elliptic regularity $h_{\bar{\mu}_{\tau_n}(t)} \to h_{\mu(t)}$ in $C^1(\overline{\Omega})$ for every $t \in [0,T]$. In conclusion, there exists a regular gradient flow as in such definition. □

The next inequality prepares the proof of the uniqueness theorem.

Lemma 5.5.4 *Let $\mu, \nu \in \mathscr{P}(\overline{\Omega})$, with $\widehat{\mu}, \widehat{\nu} \in L^\infty(\Omega)$ and $W_2^2(\mu,\nu) \leq e^{-3}$. Then there holds*

$$\Phi_\lambda(\nu) - \Phi_\lambda(\mu) \geq \frac{\lambda}{2}(\widehat{\nu}(\Omega) - \widehat{\mu}(\Omega)) + \int_{(\overline{\Omega}\times\overline{\Omega})\setminus(\partial\Omega\times\partial\Omega)} \langle \nabla h_\mu(x), y-x\rangle d\gamma(x,y) - \omega(W_2^2(\mu,\nu)), \tag{5.5.2}$$

where $\omega(t) = \widetilde{K} t |\log t|$, $\widetilde{K}$ being a suitable nonnegative constant depending only on Ω, $\|\widehat{\nu}\|_\infty$ and $\|\widehat{\mu}\|_\infty$.

Proof. We shall estimate the last term of inequality (5.2.8). For all $\gamma \in \Gamma_0(\mu,\nu)$ we have

$$\int_{\overline{\Omega}} h_\mu d(\nu - \mu) = \int_{(\overline{\Omega}\times\overline{\Omega})\setminus(\partial\Omega\times\partial\Omega)} (h_\mu(y) - h_\mu(x)) d\gamma(x,y)$$

and a Taylor expansion (with remainder in integral form) yields

$$\int_{\overline{\Omega}} h_\mu d(\nu - \mu) = \int_{(\overline{\Omega}\times\overline{\Omega})\setminus(\partial\Omega\times\partial\Omega)} \langle \nabla h_\mu(x), y-x\rangle d\gamma(x,y) + \frac{1}{2}\int_0^1 \int_{(\overline{\Omega}\times\overline{\Omega})\setminus(\partial\Omega\times\partial\Omega)} \langle \nabla^2 h_\mu((1-\theta)x + \theta y)(y-x), y-x\rangle d\gamma(x,y) d\theta. \tag{5.5.3}$$

In order to treat the remainder, we split it in two terms:

$$\begin{aligned} &\frac{1}{2}\int_0^1 \int_{(\overline{\Omega}\times\overline{\Omega})\setminus(\partial\Omega\times\partial\Omega)} |\langle \nabla^2 h_\mu((1-\theta)x + \theta y)(y-x), y-x\rangle| d\gamma(x,y) d\theta \\ \leq &\frac{1}{2}\int_0^1 \int_{\Omega\times\overline{\Omega}} |\langle \nabla^2 h_\mu((1-\theta)x + \theta y)(y-x), y-x\rangle| d\gamma(x,y) d\theta \\ &+ \frac{1}{2}\int_0^1 \int_{\overline{\Omega}\times\Omega} |\langle \nabla^2 h_\mu((1-\theta)x + \theta y)(y-x), y-x\rangle| d\gamma(x,y) d\theta. \end{aligned}$$

First term:
the measure $\chi_{\Omega\times\overline{\Omega}}\gamma$ is a transport plan between $\widehat{\mu}$ and σ_1 for a suitable $\sigma_1 \leq \nu$, then it is induced by a transport map T. Let

$$T_\theta = (1-\theta)I + \theta T, \quad \mu_\theta = T_{\theta\#}\widehat{\mu}.$$

It follows that

$$\begin{aligned}
&\int_0^1 \int_{\Omega\times\overline{\Omega}} |\langle \nabla^2 h_\mu((1-\theta)x+\theta y)(y-x), y-x\rangle| d\gamma(x,y) d\theta \\
&= \int_0^1 \frac{1}{\theta^2} \int_\Omega |\langle \nabla^2 h_\mu(T_\theta(x))(T_\theta(x)-x), T_\theta(x)-x\rangle| d\widehat{\mu}(x) d\theta \\
&= \int_0^1 \frac{1}{\theta^2} \int_\Omega |\langle \nabla^2 h_\mu(x)(x-T_\theta^{-1}(x)), x-T_\theta^{-1}(x)\rangle| d\mu_\theta(x) d\theta \\
&\leq \int_0^1 \frac{1}{\theta^2} \int_\Omega |\nabla^2 h_\mu(x)||x-T_\theta^{-1}(x)|^2 d\mu_\theta(x) d\theta \\
&\leq \int_0^1 \frac{1}{\theta^2} \left(\int_\Omega |\nabla^2 h_\mu(x)|^p d\mu_\theta(x)\right)^{1/p} \left(\int_\Omega |x-T_\theta^{-1}(x)|^{2p'} d\mu_\theta(x)\right)^{1/p'} d\theta, \qquad (5.5.4)
\end{aligned}$$

where $p>1$ and p' are conjugate exponents. Let ϱ and ϱ_θ be the densities of $\widehat{\mu}$ and μ_θ respectively. The change of variables formula gives

$$\begin{aligned}
\int_\Omega |\nabla^2 h_\mu|^p d\mu_\theta &= \int_\Omega |\nabla^2 h_\mu|^p \varrho_\theta \, d\mathcal{L}^2 \\
&\leq \left(\int_\Omega |\nabla^2 h_\mu|^{2p} d\mathcal{L}^2\right)^{1/2} \left(\int_\Omega |\varrho_\theta|^2 d\mathcal{L}^2\right)^{1/2} \\
&\leq \left(\int_\Omega |\nabla^2 h_\mu|^{2p} d\mathcal{L}^2\right)^{1/2} \left(\int_\Omega \left(\frac{\varrho}{|\det(JT_\theta)|}\right)^2 |\det(JT_\theta)| d\mathcal{L}^2\right)^{1/2} \\
&\leq M \left(\int_\Omega |\nabla^2 h_\mu|^{2p} d\mathcal{L}^2\right)^{1/2} \left(\int_\Omega \frac{d\mathcal{L}^2}{|\det(JT_\theta)|}\right)^{1/2}.
\end{aligned}$$

But for $\theta \in (0,1)$, there holds $|\det(J((1-\theta)I+\theta T))| \geq (1-\theta)^2$, yielding

$$\begin{aligned}
\int_\Omega |\nabla^2 h_\mu|^p d\mu_\theta &\leq M \left(\int_\Omega |\nabla^2 h_\mu|^{2p} d\mathcal{L}^2\right)^{1/2} \left(\int_\Omega \frac{d\mathcal{L}^2}{(1-\theta)^2}\right)^{1/2} \\
&\leq M \|\nabla^2 h_\mu\|^p_{L^{2p}(\Omega)} |\Omega|^{1/2} (1-\theta)^{-1}.
\end{aligned}$$

On the other hand

$$\begin{aligned}
\int_\Omega |I-T_\theta^{-1}|^{2p'} d\mu_\theta &= \int_\Omega |I-T_\theta^{-1}|^2 |I-T_\theta^{-1}|^{2p'-2} d\mu_\theta \leq (\mathrm{diam}\Omega)^{2(p'-1)} \int_\Omega |I-T_\theta^{-1}|^2 d\mu_\theta \\
&= (\mathrm{diam}\Omega)^{2(p'-1)} \theta^2 \int_\Omega |T-I|^2 d\widehat{\mu} = \theta^2 (\mathrm{diam}\Omega)^{2(p'-1)} W_2^2(\mu,\nu).
\end{aligned}$$

Substituting in (5.5.4), we get

$$\begin{aligned}
&\int_0^1 \int_{\Omega\times\overline{\Omega}} |\langle \nabla^2 h_\mu((1-\theta)x+\theta y)(y-x), y-x\rangle| d\gamma(x,y)d\theta \\
\leq & M^{1/p}|\Omega|^{1/(2p)}(\mathrm{diam}\Omega)^{2(p'-1)/p'}\|\nabla^2 h_\mu\|_{L^{2p}(\Omega)} W_2^{2/p'}(\mu,\nu)\int_0^1 \frac{1}{\theta^2}(1-\theta)^{-1/p}\theta^{2/p'}d\theta.
\end{aligned}$$

Now, for p sufficiently large (for example $p \geq 3$), the integral in the last term is finite and uniformly bounded in p. Moreover by elliptic regularity we have $\|\nabla^2 h_\mu\|_{L^{2p}(\Omega)} \leq c\,p\|\mu\|_\infty$, so that

$$\int_0^1 \int_{\Omega\times\overline{\Omega}} |\langle \nabla^2 h_\mu((1-\theta)x+\theta y)(y-x), y-x\rangle| d\gamma(x,y)d\theta \leq CpW_2^{2/p'}(\mu,\nu).$$

As done by Yudovich in the study of Euler equations in two dimensions (see [75, 76]), we minimize in p and, since $W_2^2(\mu,\nu) \leq e^{-3}$, we find

$$\min_{p\geq 3} pW_2^{2/p'}(\mu,\nu) = eW_2^2(\mu,\nu)|\log(W_2^2(\mu,\nu))|.$$

This is the desired logarithmic bound.

Second term:
it can be treated in the same way: for example we can consider $\chi_{\overline{\Omega}\times\Omega}\gamma \in \Gamma(\sigma_2, \widehat{\nu})$, where σ_2 is a suitable measure with $\sigma_2 \leq \mu$. Now there exists a transport map s such that $s_\#\widehat{\nu} = \sigma_2$. Letting $s_\theta = (1-\theta)s + \theta I$, we get

$$\begin{aligned}
&\int_0^1 \int_{\overline{\Omega}\times\Omega} |\langle \nabla^2 h_\mu((1-\theta)x+\theta y)(y-x), y-x\rangle| d\gamma(x,y)d\theta \\
\leq &\int_0^1 \frac{1}{(1-\theta)^2}\int_\Omega |\nabla^2 h_\mu||s_\theta^{-1} - I|\, d(s_{\theta\#}\widehat{\nu})d\theta.
\end{aligned}$$

The calculation is now analogous, taking into account that $|\det(Js_\theta)| \geq \theta^2$ and that $\int_\Omega |I-s|^2 d\widehat{\nu} \leq W_2^2(\mu,\nu)$.
Thanks to the logarithmic bound on the remainder of (5.5.3), from (5.2.8) we obtain (5.5.2). □

Eventually, we are going to state and prove our main result. The procedure is analogous to the one of [8, Theorem 3.2], but here we can show that uniqueness holds also if some mass is present on the boundary of Ω during the evolution. Even if the initial datum is not supported in Ω, this guarantees a global uniqueness result.

Theorem 5.5.5 (Uniqueness of the regular gradient flow) *Let Ω be convex. Let μ^1, μ^2 be solutions of* (5.2.1)-(5.2.2) *satisfying the conditions of* Definition 5.5.1. *Then $\mu^1(0) = \mu^2(0)$ implies $\mu^1(t) = \mu^2(t)$ for all $t \in [0,T]$.*

Proof. Let $\mu(t)$ be a regular gradient flow as in Definition 5.5.1 (it is coupled with the velocity field $-\nabla h_{\mu(t)}\chi_\Omega$), $\gamma_t \in \Gamma_0(\mu(t),\nu)$ and $\nu \in \mathscr{P}(\overline{\Omega})$. Applying (5.5.2) we find

$$\begin{aligned}\Phi_\lambda(\nu) - \Phi_\lambda(\mu(t)) \geq &\frac{\lambda}{2}(\widehat{\nu}(\Omega) - \widehat{\mu}(t)(\Omega)) \\ &+ \int_{(\overline{\Omega}\times\overline{\Omega})\setminus(\partial\Omega\times\partial\Omega)} \langle \nabla h_{\mu(t)}(x), y-x\rangle d\gamma_t(x,y) - \omega(W_2^2(\mu(t),\nu))\end{aligned}$$

whenever $W_2^2(\mu(t),\nu) \leq e^{-3}$. Since $\mu(t)$ satisfies the continuity equation, for almost every t there holds (see Lemma 2.3.5).

$$\frac{1}{2}\frac{d}{dt}W_2^2(\mu(t),\nu) = \int_{\overline{\Omega}\times\overline{\Omega}} \langle \chi_\Omega(x)\nabla h_{\mu(t)}(x), y-x\rangle d\gamma_t(x,y).$$

Substituting in the previous relation we get (for $W_2^2(\mu(t),\nu) \leq e^{-3}$)

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}W_2^2(\mu(t),\nu) \leq &\Phi_\lambda(\nu) - \Phi_\lambda(\mu(t)) - \frac{\lambda}{2}(\widehat{\nu}(\Omega) - \widehat{\mu}(t)(\Omega)) \\ &+ \omega(W_2^2(\mu(t),\nu) - \int_{\partial\Omega\times\Omega} \langle \nabla h_{\mu(t)}(x), y-x\rangle d\gamma_t(x,y).\end{aligned}$$

On $\mathrm{supp}(\widetilde{\mu}(t))$, $\nabla h_{\mu(t)}$ points towards the interior of the convex domain, then the last term is non positive, and so, for $W_2^2(\mu(t),\nu) \leq e^{-3}$,

$$\frac{1}{2}\frac{d}{dt}W_2^2(\mu(t),\nu) \leq \Phi_\lambda(\nu) - \Phi_\lambda(\mu(t)) - \frac{\lambda}{2}(\widehat{\nu}(\Omega) - \widehat{\mu}(t)(\Omega)) + \omega(W_2^2(\mu(t),\nu)). \tag{5.5.5}$$

Applying (5.5.5) first to $\mu = \mu^1(t)$, with $\nu = \mu^2(s)$, and then reversing the roles of μ^1 and μ^2, we get, by Lemma 3.3.3,

$$\frac{d}{dt}W_2^2(\mu^1(t),\mu^2(t)) \leq 4\omega(W_2^2(\mu^1(t),\mu^2(t)))$$

for almost every t such that $W_2^2(\mu^1(t),\mu^2(t)) \leq e^{-3}$. Now we make use of the logarithmic bound, that yields $\int_0^1 1/\omega(s)\,ds = \infty$. So Gronwall's lemma entails $\mu^1(t) = \mu^2(t)$ for all $t \in [0,T]$. □

5.6 The signed case

The actual model of Chapman Rubinstein and Schatzman (see [28]) involves signed measures. In fact, in Section 5.1 we have mentioned that Ginzburg-Landau vortices posses a degree and are subject to Coulombian interactions. In this last part of the chapter we would like to study the flow of the full functional (5.1.3), and possibly to relate it with the evolution model (5.1.4).

The generalized W_2 functional

A first difficulty arises at the theoretical point of view. The techniques of Chapter 3 are concerned with probability measures only, it is not clear how they could be generalized to signed measures. Moreover, we do not know how to rephrase the characterization of absolutely continuous curves by means of continuity equations given in Section 2.3.

On the other hand, the most flexible part of the theory is the minimizing movements approach. Indeed, the minimization problem

$$\min_{\nu\in\mathfrak{X}} \phi(\nu) + \frac{1}{2\tau}\, d^2(\nu,\mu), \quad \mu \in \mathfrak{X} \tag{5.6.1}$$

makes sense in any metric space $\mathfrak{X}$, d being the corresponding distance, where $\phi : \mathfrak{X} \to \mathbb{R}$. On top of that, it is not strictly needed for functional d appearing in (5.6.1) to be a distance. In fact, often the important thing is its behavior on small scale. We are going to make use of a functional which is bounded below by a distance.

Let $\mathcal{M}(\overline{\Omega})$ denote the set of real measures over $\overline{\Omega}$. Let us define the following measure subset of $\mathcal{M}(\overline{\Omega})$.

$$\mathcal{M}_{\kappa,M}(\overline{\Omega}) := \{\mu \in \mathcal{M}(\overline{\Omega}) : \mu(\overline{\Omega}) = \kappa, |\mu|(\overline{\Omega}) \le M\}, \tag{5.6.2}$$

where $\kappa \in \mathbb{R}$, $M > 0$. Let $\mu, \nu \in \mathcal{M}_{\kappa,M}(\overline{\Omega})$.

First cost

We define the following generalization of the 2-Wasserstein distance:

$$\mathbb{W}_2(\mu,\nu) := W_2(\mu^+ + \nu^-, \nu^+ + \mu^-), \tag{5.6.3}$$

It is immediate to check that, if μ and ν are nonnegative, then $\mathbb{W}_2$ reduces to the Wasserstein distance between positive measures of given mass κ on $\overline{\Omega}$ (here W_2 is naturally extended from probability measures to measures with a fixed total mass different from 1). Functional $\mathbb{W}_2$ accounts for the cost of transporting signed measures, and some heuristics on its behavior are worthy. We notice that, when transporting a signed measure μ, its positive and negative mass may change (only $\int \mu$ is fixed, as in (5.6.2)). So, in order to connect μ to ν, it can be convenient that some part of μ^+ is transported on μ^-, and this correspond to auto-annihilation of mass. On the other hand, if the total variation of ν is larger than the one of μ, one expects that, in the transport given by $\mathbb{W}_2$, a consistent part will come from moving some part of ν^- to ν^+. From the dynamic point of view, this is explained thinking of some fake mass which is created and then transported with a certain cost.

Remark 5.6.1 The framework given by $\mathbb{W}_2$ seems to fit the physical problem we are investigating, since we expect that vortices with opposite degrees can interact like dipoles and cancel themselves.

Although it is immediate to verify that $\mathbb{W}_2$ is symmetric and vanishes if and only if $\mu = \nu$, $\mathbb{W}_2$ is not a distance. Indeed, the following example shows that the triangle inequality fails. On the real line, let $\mu = \delta_0$, $\nu = \delta_4$ and $\eta = \delta_1 - \delta_2 + \delta_3$. Clearly $\mathbb{W}_2(\mu,\nu) = W_2(\mu,\nu) = 4$. But a transport plan between $\mu^+ + \eta^-$ and $\eta^- + \mu^+$ is given by $\delta_0 \times \delta_1 + \delta_2 \times \delta_3$, so that

$$\mathbb{W}_2(\mu,\eta) \le \sqrt{\int_{\mathbb{R}} |x-y|^2\, d(\delta_0 \times \delta_1) + \int_{\mathbb{R}} |x-y|^2\, d(\delta_2 \times \delta_3)} = \sqrt{2}.$$

Symmetrically, $\mathbb{W}_2(\nu,\eta) \le \sqrt{2}$, so that

$$\mathbb{W}_2(\mu,\nu) > \mathbb{W}_2(\mu,\eta) + \mathbb{W}_2(\nu,\eta).$$

On the other hand, we notice that, if $\gamma \in \Gamma_0(\mu^+ + \nu^-, \nu^+ + \mu^-)$, by Holder inequality

$$\left(\int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2\, d\gamma\right)^{1/2} \ge \sqrt{\frac{1}{2M}} \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|\, d\gamma, \tag{5.6.4}$$

but

$$\mathbb{W}_1(\mu,\nu) := \inf_{\gamma\in\Gamma(\mu^+ + \nu^-, \nu^+ + \mu^-)} \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|\, d\gamma \tag{5.6.5}$$

is indeed a distance between signed measures. The standard W_1 distance between probability measures is in fact characterized by the following duality (see for example [72])

$$W_1(\mu,\nu) = \sup_{\varphi\in Lip(\Omega)} \int_\Omega \varphi\, d(\mu-\nu), \tag{5.6.6}$$

which of course can be extended to signed measures, still remaining a distance. Notice in fact that it is not sensible to the addition of equal masses in the source and in the target of the transport. Such feature is typical of the 1-distance only, since (5.6.6) readily gives

$$\mathbb{W}_1(\mu,\nu) = \mathbb{W}_1(\mu+\sigma, \nu+\sigma), \quad \forall \sigma \in \mathcal{M}(\mathbb{R}^2). \tag{5.6.7}$$

It is worth to analyze some other features of $\mathbb{W}_2$. In the sequel we will often make use of the Hahn decomposition for a real measure μ, identifying its positive and negative parts, so that $\mu = \mu^+ - \mu^-$, where μ^+ and μ^- are two positive measures. This decomposition is minimal in the sense that, for any other couple of positive measures σ^1, σ^2 such that $\sigma^1 - \sigma^2 = \mu$, there holds $\mu^+ \le \sigma^1$ and $\mu^- \le \sigma^2$. In the next proposition we see that $\mathbb{W}_2$ 'metrizes' the weak topology of $\mathcal{M}_{\kappa, M}(\overline{\Omega})$.

Proposition 5.6.2 *Let μ_n, μ belong to $\mathcal{M}_{\kappa, M}(\overline{\Omega})$. Then $\mu_n \rightharpoonup \mu$ if and only if $\mathbb{W}_2(\mu_n, \mu) \to 0$.*

Proof. Assume that $\mu_n \rightharpoonup \mu$. Since $\mu_n^+(\overline{\Omega}) \le M$ and $\mu_n^-(\overline{\Omega}) \le M$, by tightness there exists a subsequence (n_k) such that $\mu_{n_k}^+ \rightharpoonup \sigma^+$ and $\mu_{n_k}^- \rightharpoonup \sigma^-$, with $\sigma^+ - \sigma^- = \mu$. By continuity of the Wasserstein distance, for each limit point we have $W_2(\mu_n^+ + \mu^-, \mu^+ + \mu_n^-) \to W_2(\sigma^+ + \mu^-, \mu^+ + \sigma^-) = 0$.

Assume that $\mathbb{W}_2(\mu_n, \mu) \to 0$, that is, $W_2(\mu_n^+ + \mu^-, \mu^+ + \mu_n^-) \to 0$. But W_2 metrizes the weak convergence, so there exists a positive measure ϑ such that $\mu_n^+ + \mu^- \rightharpoonup \vartheta$ and $\mu_n^- + \mu^+ \rightharpoonup \vartheta$, hence $\mu_n^+ - \mu_n^- \rightharpoonup \mu^+ - \mu^- = \mu$. □

We have seen that $\mathbb{W}_2$ is not a distance. With a similar simple construction, it is possible to see that the map $\nu \mapsto \mathbb{W}_2(\nu, \mu)$ is not weakly l.s.c. in $\mathcal{M}_{\kappa, M}(\Omega)$. The point is that if $\mu_n \rightharpoonup \mu$ in $\mathcal{M}_{\kappa, M}$, then μ_n^+ and μ_n^- are tight, but the limits are not in general μ^+ and μ^-. In order to overcome this problem, we consider the sequentially lower semicontinuous envelope of $\mathbb{W}_2$, that is

$$\mathbb{W}_2^-(\nu, \mu) := \inf_{\substack{\nu_n^+(\overline{\Omega}) \le \mu^+(\overline{\Omega}) \\ \nu_n^-(\overline{\Omega}) \le \mu^-(\overline{\Omega})}} \left\{ \liminf_{n\to\infty} \mathbb{W}_2(\nu_n, \mu) : (\nu_n) \subset \mathcal{M}_{\kappa, M}(\overline{\Omega}), \nu_n \rightharpoonup \nu \right\}. \tag{5.6.8}$$

As $\nu_n \rightharpoonup \nu$ in $\mathcal{M}_{\kappa, M}(\overline{\Omega})$, we have that $\nu_{n_k}^+ \rightharpoonup \sigma^1$ and $\nu_{n_k}^- \rightharpoonup \sigma^2$, with $\sigma^1(\overline{\Omega}) \le M$ and $\sigma^2(\overline{\Omega}) \le M$ and $\sigma^1 - \sigma^2 = \nu$. Here σ^1 and σ^2 are not the positive and negative parts of ν, but simply two measures such that $\sigma^1 - \sigma^2 = \nu$ (a non minimal decomposition). Hence we can write the l.s.c. envelope as

$$\inf \left\{ W_2(\sigma^1 + \mu^-, \mu^+ + \sigma^2) : \sigma^1(\overline{\Omega}) \le M,\, \sigma^2(\overline{\Omega}) \le M,\, \sigma^1 - \sigma^2 = \nu \right\}. \tag{5.6.9}$$

Notice that the boundedness of total variations prevents the envelope from being identically zero, and by continuity of W_2 the infimum above is attained. We denote by θ^1, θ^2 the corresponding optimal couple. By construction, the map

$$\nu \mapsto W_2(\theta^1 + \mu^-, \mu^+ + \theta^2)$$

is l.s.c. on $\mathcal{M}_{\kappa, M}(\overline{\Omega})$, and of course $W_2(\theta^1 + \mu^-, \mu^+ + \theta^2) \le \mathbb{W}_2(\nu, \mu)$.

In order to deal with optimal transport plans between signed measures, consider partitions of the positive and negative parts of ν and μ of the form

$$\begin{aligned} \mu_0^+ + \mu_1^+ = \mu^+, &\quad \mu_0^- + \mu_1^- = \mu^-, \\ \nu_0^+ + \nu_1^+ = \nu^+, &\quad \nu_0^- + \nu_1^- = \nu^-, \end{aligned} \tag{5.6.10}$$

where all the terms are positive measures. Some compatibility conditions have to be taken into account, and precisely

$$\nu_0^+(\overline{\Omega}) = \mu_0^+(\overline{\Omega}), \quad \mu_0^-(\overline{\Omega}) = \nu_0^-(\overline{\Omega}), \quad \mu_1^-(\overline{\Omega}) = \mu_1^+(\overline{\Omega}), \quad \nu_1^+(\overline{\Omega}) = \nu_1^-(\overline{\Omega}). \tag{5.6.11}$$

Of course there are many partitions of this kind. Moreover, we have the following

Lemma 5.6.3 (Splitting of the optimal plan) *Let $\gamma \in \Gamma_0(\nu^+ + \mu^-, \mu^+ + \nu^-)$. Then there exists a partition of the form* (5.6.10)-(5.6.11) *such that γ can be written as the sum of four plans γ_+^+, γ_-^-, γ_-^+, γ_+^- satisfying*

$$\begin{aligned} \gamma_+^+ \in \Gamma_0(\nu_0^+, \mu_0^+), &\quad \gamma_-^- \in \Gamma_0(\mu_0^-, \nu_0^-), \\ \gamma_-^+ \in \Gamma_0(\mu_1^-, \mu_1^+), &\quad \gamma_+^- \in \Gamma_0(\nu_1^+, \nu_1^-). \end{aligned} \tag{5.6.12}$$

Proof. Let $\vartheta_1 = \nu^+ + \mu^-$ and $\vartheta_2 = \mu^+ + \nu^-$. It is clear that ν^+ and μ^- are both absolutely continuous with respect to ϑ_1. Let $f_1, g_1 \in L^1(\mathbb{R}^2, \vartheta_1)$ denote the respective densities. Similarly, let f_2, g_2 be the densities of ν^- and μ^+ with respect to ϑ_2, so that

$$\nu^+ = f_1\vartheta_1, \quad \mu^- = g_1\vartheta_1, \quad \mu^+ = g_2\vartheta_2, \quad \nu^- = f_2\vartheta_2.$$

Clearly $f_1 + g_1 = f_2 + g_2 = 1$, so that we can write

$$\gamma = (f_1\circ\pi^1)(g_2\circ\pi^2)\gamma + (f_1\circ\pi^1)(f_2\circ\pi^2)\gamma + (g_1\circ\pi^1)(g_2\circ\pi^2)\gamma + (g_1\circ\pi^1)(f_2\circ\pi^2)\gamma. \quad (5.6.13)$$

Notice that

$$\pi^1_\#\left((f_1\circ\pi^1)(g_2\circ\pi^2)\gamma\right) = f_1\,\pi^1_\#\left((g_2\circ\pi^2)\gamma\right) \le f_1\,\pi^1_\#\gamma = f_1\vartheta_1 = \nu^+,$$
$$\pi^2_\#\left((f_1\circ\pi^1)(g_2\circ\pi^2)\gamma\right) = g_2\,\pi^2_\#\left((f_1\circ\pi^1)\gamma\right) \le g_2\,\pi^2_\#\gamma = g_2\vartheta_2 = \mu^+.$$

Moreover,

$$\pi^1_\#((f_1\circ\pi^1)(g_2\circ\pi^2)\gamma) + \pi^1_\#((f_1\circ\pi^1)(f_2\circ\pi^2)\gamma) = f_1\,\pi^1_\#((g_2\circ\pi^2 + f_2\circ\pi^2)\gamma) = f_1\,\pi^1_\#\gamma = \nu^+.$$

With the analogous computations for the other terms in the right hand side of (5.6.13), we see that the marginals of the four plans therein are submeasures of $\nu^+, \mu^-, \mu^+, \nu^-$ satisfying (5.6.10)-(5.6.11). Hence, in (5.6.13) γ is written as the sum of four plans on a partition of the desired form. Moreover, each of these plans is optimal, since their sum is. ☐

Second cost

For dealing with a sequence of measures with decreasing total mass, we introduce the following simplified version. Let $\mu, \nu \in \mathcal{M}_{\kappa, M}(\overline{\Omega})$. Let $|\nu|(\overline{\Omega}) \le |\mu|(\overline{\Omega})$. Define

$$\mathcal{W}_2^2(\nu,\mu) = \inf_{\substack{\nu_n \rightharpoonup \nu \\ \nu_n^+(\overline{\Omega}) = \mu^+(\overline{\Omega}) \\ \nu_n^-(\overline{\Omega}) = \mu^-(\overline{\Omega})}} \left\{ \liminf_{n\to\infty} \left(W_2^2(\nu_n^+, \mu^+) + W_2^2(\nu_n^-, \mu^-) \right) \right\}, \quad (5.6.14)$$

By its very definition, we see that the map $\nu \mapsto \mathcal{W}_2^2(\nu,\mu)$ is lower semicontinuous. Moreover, since any weak limit point of ν_n^+, ν_n^- is a couple σ^+, σ^- satisfying $\sigma^+ - \sigma^- = \nu$, $\mathcal{W}_2^2(\nu,\mu)$ can be also written as

$$\inf_{\substack{\sigma^+ - \sigma^- = \nu \\ \sigma^+(\overline{\Omega}) = \mu^+(\overline{\Omega}) \\ \sigma^-(\overline{\Omega}) = \mu^-(\overline{\Omega})}} \left\{ W_2^2(\sigma^+, \mu^+) + W_2^2(\sigma^-, \mu^-) \right\}. \quad (5.6.15)$$

Tightness and semicontinuity of the standard Wasserstein distance show that there exists an optimal couple ϑ^+, ϑ^- such that

$$\mathcal{W}_2^2(\nu,\mu) = W_2^2(\vartheta^+, \mu^+) + W_2^2(\vartheta^-, \mu^-), \quad (5.6.16)$$

where $\vartheta^+ - \vartheta^- = \nu$.

Notice that $\mathcal{W}_2$ is non symmetric. But symmetry is not a key point, since we are going to compute the costs corresponding to subsequent time steps: an evolution problem has a natural time direction. About the relation with the other objects we have defined, it is not difficult to show the following

Proposition 5.6.4 *Let $\mu, \nu \in \mathcal{M}_{\kappa, M}(\overline{\Omega})$ and $|\nu|(\overline{\Omega}) \le |\mu|(\overline{\Omega})$. Then*

$$\mathcal{W}_2(\nu, \mu) \ge \mathbb{W}_2^-(\nu, \mu) \ge \sqrt{\frac{1}{2M}}\mathbb{W}_1(\mu, \nu). \tag{5.6.17}$$

Proof. Let ϑ^1, ϑ^2 be the optimal couple for $\mathcal{W}_2$, so that the infimum in (5.6.15) is attained. Let $\gamma^1 \in \Gamma_0(\mu^+, \vartheta^+)$, $\gamma^2 \in \Gamma_0(\mu^-, \vartheta^-)$. Then $(\gamma^2)^{-1} \in \Gamma_0(\vartheta^-, \mu^-)$ and $\gamma^1 + (\gamma^2)^{-1} \in \Gamma(\mu^+ + \vartheta^-, \vartheta^+ + \mu^-)$. Hence

$$\begin{aligned}
\mathcal{W}_2^2(\mu, \nu) = W_2^2(\mu^+, \vartheta^+) + W_2^2(\mu^-, \vartheta^-) &= \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2 \, d(\gamma^1 + \gamma^2)(x,y) \\
&= \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2 \, d(\gamma^1 + (\gamma^2)^{-1})(x,y) \\
&\ge W_2^2(\mu^+ + \vartheta^-, \vartheta^+ + \mu^-) \ge \mathbb{W}_2^-(\nu, \mu).
\end{aligned}$$

Exploiting (5.6.4) we get the thesis. □

Notation 5.6.5 We let ϑ_1 denote the common part of ϑ^+ and ϑ^-, so that $\vartheta^+ = \nu^+ + \vartheta_1$ and $\vartheta^- = \nu^- + \vartheta_1$. Moreover, we let $\gamma^+ \in \Gamma_0(\vartheta^+, \mu^+)$ and $\gamma^- \in \Gamma_0(\vartheta^-, \mu^-)$ be the two optimal transport plans corresponding to $\mathcal{W}_2$. Thanks to (a simplified version of) Lemma 5.6.3, we can write these plans as

$$\gamma^+ = \gamma_0^+ + \gamma_1^+ \quad \text{and} \quad \gamma^- = \gamma_0^- + \gamma_1^-.$$

Each plan in the splitting is optimal:

$$\gamma_0^+ \in \Gamma_0(\nu^+, \mu_0^+), \quad \gamma_1^+ \in \Gamma_0(\vartheta_1, \mu_1^+), \quad \gamma_0^- \in \Gamma_0(\nu^-, \mu_0^-), \quad \gamma_1^- \in \Gamma_0(\vartheta_1, \mu_1^-), \tag{5.6.18}$$

where $\mu_0^+ + \mu_1^+ = \mu^+$ and $\mu_0^- + \mu_1^- = \mu^-$.

5.7 Fine characterization of discrete minimizers

The functional we are going to analyze is (5.1.3), defined on real measures. Notice that the Φ_0 part is the same as the positive measures case. As a consequence, the full Φ_λ is still lower semicontinuous, since $\mu \mapsto |\mu|(\Omega)$ is. Moreover, Proposition 5.2.2 works in the same way, giving the representation formula

$$\Phi_\lambda(\mu) = \frac{1}{2}(\lambda|\mu|(\Omega) + |\Omega|) + \sup_{h-1\in H_0^1(\Omega)} \left\{ \int_\Omega (h-1)\, d\mu - \frac{1}{2}\int_\Omega |\nabla h|^2 + |h|^2 \right\}, \tag{5.7.1}$$

the supremum being attained at $h = h_\mu$. By means of (5.7.1), the analogous of (5.2.6) is easily obtained: for any couple of real measures μ, ν there holds

$$\Phi_\lambda(\mu) - \Phi_\lambda(\nu) \geq \frac{\lambda}{2}|\mu|(\Omega) - \frac{\lambda}{2}|\nu|(\Omega) + \int_\Omega (h_\nu - 1)\, d(\mu - \nu). \tag{5.7.2}$$

On the other hand, (5.2.7) has its counterpart too, and in particular

$$\Phi_0(\mu) - \Phi_0(\nu) = \nu(\Omega) - \mu(\Omega) + \frac{1}{2}\int_\Omega (h_\mu + h_\nu)(\mu - \nu). \tag{5.7.3}$$

We are concerned with the discrete minimization problem: given $\mu \in \mathcal{M}_{\kappa,\,M}(\overline{\Omega})$, solve

$$\min_{\nu \in \mathcal{M}_{\kappa,\,M}(\overline{\Omega}),\, |\nu|(\overline{\Omega}) \leq |\mu|(\overline{\Omega})} \Phi_\lambda(\nu) + \frac{1}{2\tau}\mathcal{W}_2^2(\nu, \mu). \tag{5.7.4}$$

Again, we need a perturbed, regularized functional. Let

$$\Phi_\lambda^\delta(\nu) = \ \Phi_\lambda(\widehat{\nu}) + \delta \int_\Omega |\widehat{\nu}|^4 \tag{5.7.5}$$

if $\widehat{\mu} \ll \mathcal{L}^2$ and $+\infty$ otherwise. We have the following result (analogous of Lemma 5.3.2).

Lemma 5.7.1 *The perturbed minimization problem*

$$\min_{\nu \in \mathcal{M}_{\kappa,\,M}(\overline{\Omega}),\, |\nu|(\overline{\Omega}) \leq |\mu|(\overline{\Omega})} \Phi_\lambda^\delta(\nu) + \frac{1}{2\tau}\mathcal{W}_2^2(\nu, \mu) \tag{5.7.6}$$

has a solution μ_τ^δ, the family μ_τ^δ has limit points both for the strong H^{-1} topology and the $\mathcal{M}(\overline{\Omega})$ topology, $\delta \int_\Omega (\widehat{\mu}_\tau^\delta)^4 \to 0$ as $\delta \to 0$, and any limit point μ_τ as $\delta \to 0$ solves (5.7.4).

Proof. The existence of μ_τ^δ is given by the direct method, as for the existence of μ_τ. Let M_δ be the minimum in (5.7.6) and let M be the minimum of the functional in (5.7.4). It is clear that $M_\delta \geq M$; on the other hand, $\Phi_\lambda^\delta \to \Phi_\lambda$ at any admissible point ν such that $\widehat{\nu} \in L^4(\Omega)$. Then

$$\limsup_{\delta \downarrow 0} M_\delta \leq \Phi_\lambda(\nu) + \frac{1}{2\tau}\mathcal{W}_2^2(\nu, \mu)$$

for all ν in $\mathcal{M}_{\kappa,M}(\overline{\Omega})$ such that $|\nu|(\overline{\Omega}) \leq |\mu|(\overline{\Omega})$ and $\widehat{\nu} \in L^4(\Omega)$. By density we obtain $\limsup_\delta M_\delta \leq M$, therefore $M_\delta \to M$ as $\delta \to 0$.

If μ_τ is a weak limit point of μ_τ^δ along some sequence $\delta_i \to 0$, the lower semicontinuity of Φ_λ gives, since $\Phi_\lambda^{\delta_i} \geq \Phi_\lambda$ for any i,

$$\Phi_\lambda(\mu_\tau) + \frac{1}{2\tau}\mathcal{W}_2^2(\mu_\tau, \mu) \leq \liminf_{i\to\infty} \Phi_\lambda(\mu_\tau^{\delta_i}) + \frac{1}{2\tau}\mathcal{W}_2^2(\mu_\tau^{\delta_i}, \mu) \leq \liminf_{i\to\infty} M_{\delta_i} = M,$$

therefore μ_τ is a solution of (5.7.4). As a consequence

$$\lim_{i\to\infty} \Phi_\lambda(\mu_\tau^{\delta_i}) + \delta_i \int_\Omega |\widehat{\mu}_\tau^{\delta_i}|^4 + \frac{1}{2\tau}\mathcal{W}_2^2(\mu_\tau^{\delta_i}, \mu) = \Phi_\lambda(\mu_\tau) + \frac{1}{2\tau}\mathcal{W}_2^2(\mu_\tau, \mu).$$

By the lower semicontinuity of Φ_λ and $\nu \mapsto \mathcal{W}_2^2(\nu,\mu)$ it follows that $\Phi_\lambda(\mu_\tau^{\delta_i}) \to \Phi_\lambda(\mu_\tau)$, $\mathcal{W}_2^2(\mu_\tau^{\delta_i},\mu) \to \mathcal{W}_2^2(\mu_\tau,\mu)$ and $\delta_i \int_\Omega (\widehat{\mu}_\tau^{\delta_i})^4 \to 0$. Now, since $\Phi_\lambda(\nu)$ is itself the sum of two lower semicontinuous terms, namely $\Phi_0(\nu)$ and $\lambda\nu(\Omega)/2$, we obtain

$$\lim_{i\to\infty} \lambda\mu_\tau^{\delta_i}(\Omega) = \lambda\mu_\tau(\Omega) \quad \text{and} \quad \lim_{i\to\infty} \int_\Omega |\nabla h_{\mu_\tau^{\delta_i}}|^2 + (h_{\mu_\tau^{\delta_i}} - 1)^2 = \int_\Omega |\nabla h_{\mu_\tau}|^2 + (h_{\mu_\tau} - 1)^2.$$

In particular $\widehat{\mu}_\tau^{\delta_i} \to \widehat{\mu}_\tau$ strongly in $H^{-1}(\Omega)$. □

Next we derive an Euler equation for problem (5.7.6), which will give a characterization of the discrete velocity of the scheme. It is useful to begin with the analysis of the corresponding minimization problem on the whole plane. This way, we can deal with competitors of the form $\mathbf{t}_\#\nu$, which can have some mass outside $\overline{\Omega}$.

Lemma 5.7.2 *Any minimizer ν of*

$$\min\left\{\Phi_\lambda^\delta(\nu) + \frac{1}{2\tau}\mathcal{W}_2^2(\nu,\mu) : \nu \in \mathcal{M}_{\kappa,M}(\mathbb{R}^2),\ |\nu|(\mathbb{R}^2) \le |\mu|(\overline{\Omega}),\ \int_{\mathbb{R}^2} |x|^2\, d\nu < +\infty\right\} \tag{5.7.7}$$

satisfies

$$-3\delta\nabla((\widehat{\nu})^4) - \frac{1}{2}\nabla h_\nu \widehat{\nu} = \frac{1}{\tau}\pi^1_\#(\chi_\Omega(x)(x-y)\gamma_0^+) + \frac{1}{\tau}\pi^1_\#(\chi_\Omega(x)(x-y)\gamma_0^-) \quad \text{in } \mathcal{D}'(\mathbb{R}^2), \tag{5.7.8}$$

where γ_0^+ and γ_0^- are the optimal transport plans given by splitting, with the notation of (5.6.18)*: $\gamma_0^+ \in \Gamma_0(\nu^+,\mu_0^+)$ and $\gamma_0^- \in \Gamma_0(\nu^-,\mu_0^-)$, where μ_0^+ and μ_0^- are suitable submeasures of μ^+ and μ^- respectively.*

Proof.

We perform a variation of the internal part of the optimal measure ν along a smooth vector field $\boldsymbol{\xi} : \mathbb{R}^2 \to \mathbb{R}^2$.

Let ϑ^+, ϑ^- be the optimal couple for $\mathcal{W}_2$, such that (5.6.16) holds. If $\gamma^+ \in \Gamma_0(\vartheta^+,\mu^+)$ and $\gamma^- \in \Gamma_0(\vartheta^-,\mu^-)$, we can consider a splitting as (5.6.18). Accordingly, ϑ_1 denotes the common part of ϑ^+ and ϑ^- and

$$\mathcal{W}_2^2(\nu,\mu) = \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2\, d(\gamma_0^+ + \gamma_0^- + \gamma_1^+ + \gamma_1^-)(x,y). \tag{5.7.9}$$

Let

$$\nu_\varepsilon = \widetilde{\nu} + (\mathbf{I} + \varepsilon\boldsymbol{\xi})_\#\widehat{\nu} \tag{5.7.10}$$

and

$$\Omega_\varepsilon = \{x \in \Omega : x + \varepsilon\boldsymbol{\xi}(x) \in \Omega\}. \tag{5.7.11}$$

For small ε, $\mathbf{I} + \varepsilon\boldsymbol{\xi}$ is injective, and it is clear that $\nu_\varepsilon(\mathbb{R}^2) = \nu(\mathbb{R}^2)$ and $|\nu_\varepsilon|(\mathbb{R}^2) = |\nu|(\mathbb{R}^2)$. Let moreover

$$\begin{aligned} \gamma_\varepsilon^+ &= (\mathbf{I} + \varepsilon\boldsymbol{\xi}, \mathbf{I})_\#(\chi_{\Omega\times\overline{\Omega}}\gamma_0^+) + \chi_{\partial\Omega\times\overline{\Omega}}\gamma_0^+ + \gamma_1^+ \\ \gamma_\varepsilon^- &= (\mathbf{I} + \varepsilon\boldsymbol{\xi}, \mathbf{I})_\#(\chi_{\Omega\times\overline{\Omega}}\gamma_0^-) + \chi_{\partial\Omega\times\overline{\Omega}}\gamma_0^- + \gamma_1^-. \end{aligned} \tag{5.7.12}$$

We have

$$\mathcal{W}_2^2(\nu_\varepsilon,\mu) \le W_2^2(\widetilde{\vartheta}^+ + \widehat{\vartheta}_1 + (\mathbf{I}+\varepsilon\boldsymbol{\xi})_\#\widehat{\nu}^+,\mu^+) + W_2^2(\widetilde{\vartheta}^- + \widehat{\vartheta}_1 + (\mathbf{I}+\varepsilon\boldsymbol{\xi})_\#\widehat{\nu}^-,\mu^-),$$

but it is clear from (5.7.12) that

$$\gamma_\varepsilon^+ \in \Gamma(\widetilde{\vartheta}^+ + \widehat{\vartheta}_1 + (\mathbf{I}+\varepsilon\boldsymbol{\xi})_\#\widehat{\nu}^+,\mu^+) \quad \text{and} \quad \gamma_\varepsilon^- \in \Gamma(\widetilde{\vartheta}^- + \widehat{\vartheta}_1 + (\mathbf{I}+\varepsilon\boldsymbol{\xi})_\#\widehat{\nu}^-,\mu^-),$$

hence

$$\mathcal{W}_2^2(\nu_\varepsilon,\mu) \le \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2\, d(\gamma_\varepsilon^+ + \gamma_\varepsilon^-).$$

We write the last integral as

$$\begin{aligned}
\int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2\, d(\gamma_\varepsilon^+ + \gamma_\varepsilon^-) &= \int_{\Omega\times\overline{\Omega}} |x+\varepsilon\boldsymbol{\xi}(x)-y|^2\, d\gamma_0^+ + \int_{\partial\Omega\times\overline{\Omega}} |x-y|^2\, d\gamma_0^+ \\
&\quad + \int_{\Omega\times\overline{\Omega}} |x+\varepsilon\boldsymbol{\xi}(x)-y|^2\, d\gamma_0^- + \int_{\partial\Omega\times\overline{\Omega}} |x-y|^2\, d\gamma_0^- \\
&\quad + \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2\, d(\gamma_1^+ + \gamma_1^-) \\
&= \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2\, d\gamma_0^+ + 2\varepsilon \int_{\Omega\times\overline{\Omega}} \boldsymbol{\xi}(x)\cdot(x-y)\, d\gamma_0^+ \\
&\quad + \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2\, d\gamma_0^- + 2\varepsilon \int_{\Omega\times\overline{\Omega}} \boldsymbol{\xi}(x)\cdot(x-y)\, d\gamma_0^- + o(\varepsilon) \\
&\quad + \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2\, d(\gamma_1^+ + \gamma_1^-).
\end{aligned}$$

Then, recalling also (5.7.9),

$$\mathcal{W}_2^2(\nu_\varepsilon,\mu) - \mathcal{W}_2^2(\nu,\mu) \le 2\varepsilon \int_{\Omega\times\overline{\Omega}} \boldsymbol{\xi}(x)\cdot(x-y)\, d\gamma_0^+ + 2\varepsilon \int_{\Omega\times\overline{\Omega}} \boldsymbol{\xi}(x)\cdot(x-y)\, d\gamma_0^- + o(\varepsilon). \quad (5.7.13)$$

So we obtain

$$\limsup_{\varepsilon\to 0} \frac{(\mathcal{W}_2^2(\nu_\varepsilon,\mu) - \mathcal{W}_2^2(\nu,\mu))}{2\varepsilon} \le \int_{\overline{\Omega}} \boldsymbol{\xi}(z)\cdot d\big[\pi^1_\# \left(\chi_\Omega(x)(x-y)(\gamma_0^+(x,y) + \gamma_0^-(x,y))\right)\big](z). \quad (5.7.14)$$

Now for the derivative of $\Phi_\lambda^\delta(\nu_\varepsilon)$, we take advantage of the $L^4(\Omega)$ convergence of $\widehat{\nu}_\varepsilon$ to $\widehat{\nu}$ as $\varepsilon \to 0$, which gives the $W^{2,4}(\Omega)$ convergence of h_{ν_ε} to h_ν and, by smoothness of $\partial\Omega$, the $C^1(\overline{\Omega})$ convergence as well. We begin making use of the standard equality (5.7.3) about the functional Φ_0:

$$\begin{aligned}
&\Phi_0(\nu_\varepsilon) - \Phi_0(\nu) \\
=&\nu(\Omega) - \nu_\varepsilon(\Omega) + \frac{1}{2}\int_\Omega (h_{\nu_\varepsilon} + h_\nu)(\nu_\varepsilon - \nu) \\
=&\nu(\Omega) - \nu_\varepsilon(\Omega) + \frac{1}{2}\int_{\Omega_\varepsilon} (h_{\nu_\varepsilon} \circ (\mathbf{I} + \varepsilon\boldsymbol{\xi}) + h_\nu \circ (\mathbf{I} + \varepsilon\boldsymbol{\xi}))\, d\nu - \int_\Omega (h_{\nu_\varepsilon} + h_\nu)\, d\nu \\
=&\nu(\Omega) - \nu_\varepsilon(\Omega) + \frac{1}{2}\int_\Omega (h_{\nu_\varepsilon} \circ (\mathbf{I} + \varepsilon\boldsymbol{\xi}) - h_{\nu_\varepsilon} + h_\nu \circ (\mathbf{I} + \varepsilon\boldsymbol{\xi}) - h_\nu)\, d\widehat{\nu} \\
&- \frac{1}{2}\int_{\Omega\setminus\Omega_\varepsilon} (h_{\nu_\varepsilon} \circ (\mathbf{I} + \varepsilon\boldsymbol{\xi}) + h_\nu \circ (\mathbf{I} + \varepsilon\boldsymbol{\xi}))\, d\nu.
\end{aligned}$$

But the $C^1(\overline{\Omega})$ regularity, using the fact that $h_\nu = 1$ on $\partial\Omega$, yields

$$\int_{\Omega\setminus\Omega_\varepsilon} (h_{\nu_\varepsilon} \circ (\mathbf{I} + \varepsilon\boldsymbol{\xi}) + h_\nu \circ (\mathbf{I} + \varepsilon\boldsymbol{\xi}))\, d\nu = 2\nu(\Omega \setminus \Omega_\varepsilon) + O(\varepsilon)\widehat{\nu}(\Omega \setminus \Omega_\varepsilon) = 2\nu(\Omega \setminus \Omega_\varepsilon) + o(\varepsilon).$$

As a consequence

$$\Phi_0(\nu_\varepsilon) - \Phi_0(\nu) = \frac{\varepsilon}{2}\int_\Omega \nabla h_\nu \cdot \boldsymbol{\xi}\, d\nu + o(\varepsilon).$$

Since $|\nu_\varepsilon|(\Omega) \le |\nu|(\Omega)$ we have also

$$\Phi_\lambda(\nu_\varepsilon) - \Phi_\lambda(\nu) \le \frac{\varepsilon}{2}\int_\Omega \nabla h_\nu \cdot \boldsymbol{\xi}\, d\nu + o(\varepsilon). \tag{5.7.15}$$

For the regularizing term, we make use of the change of variables formula for the push forward (see for instance [4, Section 5.5]). Since $\det(J(\mathbf{I} + \varepsilon\boldsymbol{\xi})) = 1 + \varepsilon\nabla \cdot \boldsymbol{\xi} + o(\varepsilon)$, we get

$$\begin{aligned}
\frac{\delta}{\varepsilon}\left[\int_\Omega |\hat{\nu}_\varepsilon|^4 - \int_\Omega |\hat{\nu}|^4\right] &= \frac{\delta}{\varepsilon}\left[\int_{\Omega_\varepsilon} \frac{\hat{\nu}^4}{\det^3(J(\mathbf{I} + \varepsilon\boldsymbol{\xi}))} - \int_\Omega \hat{\nu}^4\right] \\
&\le -3\delta\int_\Omega \hat{\nu}^4 \nabla \cdot \boldsymbol{\xi} + o(1).
\end{aligned} \tag{5.7.16}$$

As in the proof of [8, Proposition 5.1], we join together (5.7.14), (5.7.15) and (5.7.16). By the minimality of ν, and considering that we can change sign to the arbitrary vector $\boldsymbol{\xi}$, we find the equality

$$-3\delta\int_\Omega \hat{\nu}^4 \nabla \cdot \boldsymbol{\xi} + \frac{1}{2}\int_\Omega \nabla h_\nu \cdot \boldsymbol{\xi}\, d\nu + \frac{1}{\tau}\int_{\overline{\Omega}} \boldsymbol{\xi} \cdot d\left[\pi^1_\#\left(\chi_\Omega(x)(x-y)(\gamma_0^+(x,y) + \gamma_0^-(x,y))\right)\right] = 0,$$

for any $\boldsymbol{\xi} \in C_c^\infty(\mathbb{R}^2; \mathbb{R}^2)$. □

Corollary 5.7.3 *Let $\nu \in \mathcal{M}_{\kappa,M}(\overline{\Omega})$ be a minimizer of* (5.7.6)*. Then* (5.7.8) *holds.*

Proof. Let ν_P be the minimizer of (5.7.7). Since any element of $\mathcal{M}_{\kappa,M}(\overline{\Omega})$ is admissible for such problem, there holds

$$\Phi_\lambda^\delta(\nu_P) + \frac{1}{2\tau}\mathcal{W}_2^2(\nu_P,\mu) \le \Phi_\lambda^\delta(\sigma) + \frac{1}{2\tau}(\sigma,\mu) \quad \forall\sigma \in \mathcal{M}_{\kappa,M}(\overline{\Omega}). \tag{5.7.17}$$

Let ϑ_P^+ and ϑ_P^- be the optimal couple corresponding to ν_P, such that the infimum in the definition of $\mathcal{W}_2$ is attained and $\mathcal{W}_2^2(\nu_P,\mu) = W_2^2(\vartheta_P^+,\mu^+) + W_2^2(\vartheta_P^-,\mu^-)$. Denote by γ_P^+ and γ_P^- the corresponding optimal transport plans. Consider the map $\Psi(x,y) = (x,y')$, where y' is equal to y if $y \in \Omega$, and is equal to the first point on the segment from x to y hitting $\partial\Omega$ otherwise; let θ^+ and θ^- be the first marginals of $\Psi_\#\gamma_P^+$ and $\Psi_\#\gamma_P^-$ respectively, and let $\nu = \theta^+ - \theta^-$. It is clear that $\nu \in \mathcal{M}_{\kappa,M}(\overline{\Omega})$. We claim that ν is the minimum for (5.7.6). For the proof, notice that $\widehat{\nu} = \widehat{\nu_P}$ (so that $\Phi_\lambda^\delta(\nu_P) = \Phi_\lambda^\delta(\nu)$). Moreover

$$W_2^2(\theta^+,\mu^+) \le W_2^2(\theta_P^+,\mu^+) \quad \text{and} \quad W_2^2(\theta^-,\mu^-) \le W_2^2(\theta_P^-,\mu^-),$$

since μ is supported in $\overline{\Omega}$ and the projection decreases distances. As $\theta^+ - \theta^- = \nu$, we get

$$\mathcal{W}_2^2(\nu,\mu) \le \mathcal{W}_2^2(\nu_P,\mu).$$

Combining these facts with (5.7.17), the claim is readily seen to follow. In order to conclude, it is sufficient to notice that (5.7.8) does depend only on the interior part of the minimizer. □

5.8 The entropy argument

In this section we are going to prove a key fact: the regularity of the initial datum is kept by the discrete minimizers. That is, the analogous result for positive measures (in [8]) actually extends to the general real measure framework. For, we need regularity for the reference measure μ in (5.7.6).

From now on, we will say that φ is an entropy function if it is nondecreasing, odd and C^1 and there hold

$$\begin{aligned} x\varphi'(x) &= \varphi(x) \quad \text{in } [0,1], \\ 2x^2\varphi''(x) &\ge x\varphi'(x) - \varphi(x) \quad \text{(displacement convexity).} \end{aligned} \tag{5.8.1}$$

Given an entropy φ, we will also consider an even convex function ψ on $\mathbb{R}$ such that $\psi'(x) = x\varphi'(x) - \varphi(x)$.

Lemma 5.8.1 *Let φ be an entropy and let $\mu \in \mathcal{M}_{\kappa,M}(\overline{\Omega})$ be such that $\mu = \widehat{\mu} \in L^4(\Omega)$ and $\int_\Omega \varphi(\widehat{\mu}) < \infty$. Then, for any minimizer μ_τ^δ of (5.7.6), we have*

$$\int_\Omega \varphi(\widehat{\mu}_\tau^\delta) \le \int_\Omega \varphi(\widehat{\mu}).$$

Proof. We know that $\widehat{\nu} := \widehat{\mu}_\tau^\delta$ has $L^4(\Omega)$ regularity. But in view of the Euler equation (5.7.8) we can find even more regularity. In fact, since $\widehat{\nu} \ll \mathcal{L}^2$, we know that $\chi_{\Omega\times\overline{\Omega}}\gamma_0^+$ and $\chi_{\Omega\times\overline{\Omega}}\gamma_0^-$ are plans induced by optimal transport maps r_1 and r_2. These maps correspond to the gradients of two convex Lipschitz functions (defined on $\mathbb{R}^2$). Therefore we have $r_1, r_2 \in BV_{loc}(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$ and

$$\pi^1_\# \left(\chi_\Omega(x)(x-y)\gamma_0^+\right) = (I - r_1)\nu^+$$
$$\pi^1_\# \left(\chi_\Omega(x)(x-y)\gamma_0^-\right) = (I - r_2)\nu^-.$$

This way (5.7.8) becomes

$$-3\delta\nabla((\widehat{\nu})^4) - \frac{1}{2}\nabla h_{\mu_\tau^\delta}\widehat{\nu} = \frac{1}{\tau}(I - r_1)\nu^+ + \frac{1}{\tau}(I - r_2)\nu^- \quad \text{in } \mathcal{D}'(\mathbb{R}^2), \tag{5.8.2}$$

As $r_1, r_2 \in L^\infty(\Omega)$, the right hand side is in $L^4(\Omega)$. But since $\nabla h_\nu \in C^0(\overline{\Omega})$, we have $\nabla h_\nu \widehat{\nu} \in L^4(\Omega)$, so that by comparison in (5.8.2) we find $\widehat{\nu}^4 \in W^{1,4}(\Omega)$, and by Sobolev embedding $\widehat{\nu} \in C^0(\overline{\Omega})$. Let us now define

$$r = r_1\chi_{\{\widehat{\nu}>0\}} + r_2\chi_{\{\widehat{\nu}<0\}}.$$

Let us divide (5.8.2) by $|\widehat{\nu}|$, obtaining $\widehat{\nu}$-a.e. in Ω

$$3\delta \mathrm{sgn}(\widehat{\nu})\frac{\nabla((\widehat{\nu})^4)}{\widehat{\nu}} + \frac{1}{2}\nabla h_\nu \mathrm{sgn}(\widehat{\nu})\widehat{\nu} = \frac{1}{\tau}(r_1 - I)\chi_{\{\widehat{\nu}>0\}} + \frac{1}{\tau}(r_2 - I)\chi_{\{\widehat{\nu}<0\}},$$

which, by definition of r, corresponds to

$$3\delta \mathrm{sgn}(\widehat{\nu})\frac{\nabla((\widehat{\nu})^4)}{\widehat{\nu}} + \frac{1}{2}\nabla h_\nu \mathrm{sgn}(\widehat{\nu})\widehat{\nu} = \frac{1}{\tau}(r - I) \tag{5.8.3}$$

Mind that r_1 transports $\widehat{\nu}^+$ to a submeasure of $\mu^+ = \widehat{\mu}^+ \in L^4(\Omega)$ (and similarly for r_2), so $r_{1\#}\widehat{\nu}^+ \le \widehat{\mu}^+$ and $r_{2\#}\widehat{\nu}^- \le \widehat{\mu}^-$. Since φ is not decreasing on $(0, +\infty)$, and since the relations (5.8.1) hold, we have (see [4, Lemma 10.4.4])

$$\int_{\mathbb{R}^2} \varphi(\widehat{\mu}^+) - \varphi(\widehat{\nu}^+) \ge \int_{\mathbb{R}^2} \varphi(r_{1\#}\widehat{\nu}^+) - \varphi(\widehat{\nu}^+) \ge -\int_{\mathbb{R}^2} \psi'(\widehat{\nu}^+)\,\mathrm{tr}(\nabla(r_1 - I))$$

and

$$\int_{\mathbb{R}^2} \varphi(\widehat{\mu}^-) - \varphi(\widehat{\nu}^-) \ge \int_{\mathbb{R}^2} \varphi(r_{2\#}\widehat{\nu}^-) - \varphi(\widehat{\nu}^-) \ge -\int_{\mathbb{R}^2} \psi'(\widehat{\nu}^-)\,\mathrm{tr}(\nabla(r_2 - I)).$$

We sum up the last two inequalities using the fact that $\varphi(0) = 0$, so that

$$\int_{\mathbb{R}^2} \varphi(|\widehat{\mu}|) - \varphi(|\widehat{\nu}|) \ge -\int_{\mathbb{R}^2} \psi'(\widehat{\nu}^+)\,\mathrm{tr}(\nabla(r_1 - I)) - \int_{\mathbb{R}^2} \psi'(\widehat{\nu}^-)\,\mathrm{tr}(\nabla(r_2 - I)). \tag{5.8.4}$$

But r_1 and r_2 are gradients of convex functions, so that we have $\mathrm{tr}(\nabla(r_1 - I)) \le \mathrm{div}\,(r_1 - I)$ and $\mathrm{tr}(\nabla(r_2 - I)) \le \mathrm{div}\,(r_2 - I)$, being the divergences in the distributional sense (they are measures, as r_1, r_2 are BV). Now consider the quantity $\operatorname{div}((r_1-I)\chi_{\{\widehat{\nu}>0\}})$. Formally by the Volpert formula for BV functions (see [3]) we have

$$\operatorname{div}((r_1-I)\chi_{\{\widehat{\nu}>0\}}) = \operatorname{div}(r_1-I)\chi_{\{\widehat{\nu}>0\}} + \langle r_1-I, n_{\{\widehat{\nu}=0\}}\rangle \, d\mathcal{H}^1 \llcorner(\{\widehat{\nu}=0\}), \tag{5.8.5}$$

where n denotes the normal. The computation is formal because the level set $\{\widehat{\nu}=0\}$ need not be $\mathcal{H}^1$-rectifiable. But for almost any $\varepsilon>0$ the boundaries of the sublevels $\{|\widehat{\nu}|<\varepsilon\}$ are, by the BV regularity of $\widehat{\nu}$. Since we are dealing with integrals of the form

$$\int_{\partial\{|\widehat{\nu}|<\varepsilon\}} \psi'(\widehat{\nu})\, d\mathcal{H}^1 = 0,$$

where ψ' vanishes in a whole interval containing 0, we can take ε small enough and use the formula above. As a consequence,

$$\int_{\mathbb{R}^2} \psi'(\widehat{\nu}^+)\operatorname{div}((r_1-I)\chi_{\{\widehat{\nu}>0\}}) = \int_{\mathbb{R}^2} \psi'(\widehat{\nu}^+)\operatorname{div}(r_1-I)\chi_{\{\widehat{\nu}>0\}}.$$

The same for $\widehat{\nu}^-$ on $\{\widehat{\nu}<0\}$. This way, from (5.8.4) we deduce

$$\begin{aligned}
\int_{\mathbb{R}^2} \varphi(|\widehat{\mu}|)-\varphi(|\widehat{\nu}|) &\geq -\int_{\mathbb{R}^2} \psi'(\widehat{\nu}^+)\operatorname{tr}(\nabla(r_1-I)) - \int_{\mathbb{R}^2} \psi'(\widehat{\nu}^-)\operatorname{tr}(\nabla(r_2-I)) \\
&\geq -\int_{\{\widehat{\nu}>0\}} \psi'(|\widehat{\nu}|)\operatorname{div}(r_1-I) - \int_{\{\widehat{\nu}<0\}} \psi'(|\widehat{\nu}|)\operatorname{div}(r_2-I) \\
&= -\int_{\mathbb{R}^2} \psi'(|\widehat{\nu}|)\operatorname{div}((r_1-I)\chi_{\{\widehat{\nu}>0\}}) - \int_{\mathbb{R}^2} \psi'(|\widehat{\nu}|)\operatorname{div}((r_2-I)\chi_{\{\widehat{\nu}<0\}}) \\
&= -\int_{\mathbb{R}^2} \psi'(|\widehat{\nu}|)\operatorname{div}(r-I) \\
&= -\int_{\Omega} \psi'(|\widehat{\nu}|)\operatorname{div}(r-I)
\end{aligned}$$

We make use of (5.8.3) to estimate the last integral, that is, by means of (5.8.3), from the latter inequality we have

$$\int_{\mathbb{R}^2} \varphi(|\widehat{\mu}|)-\varphi(|\widehat{\nu}|) \geq -\tau \int_{\Omega} \psi'(|\widehat{\nu}|)\operatorname{div}\left[3\delta \operatorname{sgn}(\widehat{\nu})\frac{\nabla((\widehat{\nu})^4)}{\widehat{\nu}} + \frac{1}{2}\operatorname{sgn}(\widehat{\nu})\nabla h_{\mu_\tau^\delta}\right]. \tag{5.8.6}$$

As ψ' is odd and $\widehat{\nu}$ vanishes on $\mathbb{R}^2\backslash\Omega$ we find

$$\int_{\Omega} \psi'(|\widehat{\nu}|)\operatorname{div}(\operatorname{sgn}(\widehat{\nu})\nabla h_\nu) = \int_{\Omega} \psi'(\widehat{\nu})\Delta h_\nu.$$

Moreover, by convexity of ψ we obtain

$$\int_{\Omega} \psi'(|\widehat{\nu}|)\operatorname{div}(\operatorname{sgn}(\widehat{\nu})\nabla h_\nu) = \int_{\Omega} \psi'(\widehat{\nu})(h_\nu-\widehat{\nu}) \leq \int_{\Omega} \psi(h_\nu)-\psi(\widehat{\nu}). \tag{5.8.7}$$

Now consider the equation $-\Delta h_\nu + h_\nu = \nu$ in Ω. Multiplying by $\psi'(h_\nu)$ and integrating by parts yields

$$\int_\Omega \psi''(h_\nu)|\nabla h_\nu|^2 + \psi'(h_\nu)(h_\nu - \widehat{\nu}) = 0,$$

so, with the convexity of ψ on $\mathbb{R}$, we obtain

$$\int_\Omega \psi''(h_\nu)|\nabla h_\nu|^2 \leq \int_\Omega \psi(\widehat{\nu}) - \psi(h_\nu).$$

Inserting this inequality in (5.8.7) we get

$$\int_\Omega \psi'(|\widehat{\nu}|)\mathrm{div}\,(\mathrm{sgn}(\widehat{\nu})\nabla h_\nu) \leq -\int_\Omega \psi''(h_\nu)|\nabla h_\nu|^2. \tag{5.8.8}$$

On the other hand, the oddness of ψ' also gives

$$\begin{aligned}\int_\Omega \psi'(|\widehat{\nu}|)\mathrm{div}\left(\mathrm{sgn}(\widehat{\nu})\frac{\nabla\widehat{\nu}^4}{\widehat{\nu}}\right) &= \int_\Omega \psi'(|\widehat{\nu}|)\mathrm{div}\left(\frac{\nabla\widehat{\nu}^4}{|\widehat{\nu}|}\right) = \\ &= -\int_\Omega \nabla\psi'(|\widehat{\nu}|)\cdot\frac{\nabla\widehat{\nu}^4}{|\widehat{\nu}|} = \\ &= -\int_\Omega g'(\widehat{\nu}^4)\frac{(\nabla\widehat{\nu}^4)^2}{|\widehat{\nu}|},\end{aligned}$$

where $g(x) = \psi'(x^{1/4})$, and so

$$\int_\Omega \psi'(|\widehat{\nu}|)\mathrm{div}\left(\mathrm{sgn}(\widehat{\nu})\frac{\nabla\widehat{\nu}^4}{\widehat{\nu}}\right) \leq 0. \tag{5.8.9}$$

Inserting (5.8.8) and (5.8.9) in (5.8.6), we find

$$\int_\Omega \varphi(|\widehat{\mu}|) - \varphi(|\widehat{\nu}|) \geq \tau\int_\Omega \psi''(h_\nu)|\nabla h_\nu|^2. \tag{5.8.10}$$

Since $\psi'' \geq 0$ (by convexity of ψ) we conclude. □

Corollary 5.8.2 *Let $\mu = \widehat{\mu} \in L^\infty(\Omega)$ and $K = \max\{1, \|\widehat{\mu}\|_\infty\}$. There exists a minimizer $\widehat{\mu}_\tau$ of* (5.7.4) *such that*

$$\|\widehat{\mu}_\tau\|_\infty \leq K, \quad |h_{\mu_\tau}| \leq K. \tag{5.8.11}$$

Proof. Since $K \geq 1$ we can construct a sequence φ_n of entropies converging monotonically to

$$\varphi(x) := \begin{cases} -\infty & \text{for } x < -K \\ x & \text{for } -K \leq x \leq K \\ +\infty & \text{for } x > K. \end{cases}$$

Let also ψ_n be such that $\psi_n''(x) = x\varphi_n''(x)$, with ψ_n'' converging monotonically to ∞ if $|x| > K$. By (5.8.10) we have

$$\int_\Omega \varphi_n(|\widehat{\mu}_\tau^\delta|) + \tau \int_\Omega \psi_n''(h_{\mu_\tau^\delta})|\nabla h_{\mu_\tau^\delta}|^2 \leq \int_\Omega \varphi_n(|\widehat{\mu}|).$$

Now we apply Lemma 5.7.1 to obtain a limit point μ_τ of μ_τ^δ, as $\delta \to 0$, such that μ_τ is a minimizer of (5.7.4). Then, the weak lower semicontinuity of $|\mu| \mapsto \int_\Omega \varphi(|\mu|)$ in L^p yields

$$\int_\Omega \varphi_n(|\widehat{\mu}_\tau|) + \tau \int_\Omega \psi_n''(h_{\mu_\tau})|\nabla h_{\mu_\tau}|^2 \leq \int_\Omega \varphi_n(|\widehat{\mu}|) \leq \int_\Omega \varphi(|\widehat{\mu}|).$$

From the convergence properties of φ_n and ψ_n we get $|\widehat{\mu}_\tau| \leq K$ a.e. in Ω and $|h_{\mu_\tau}| \leq K$. □

Corollary 5.8.3 *Let $\mu = \widehat{\mu} \in L^p(\Omega)$, $p \geq 4$. Then there exists a minimizer μ_τ of* (5.7.4) *with $\widehat{\mu}_\tau$ belonging to $L^p(\Omega)$ and such that*

$$\|\widehat{\mu}_\tau\|_p \leq \|\widehat{\mu}\|_p.$$

Proof. Let us make the following choice for the entropy.

$$\varphi(x) := \begin{cases} x & \text{for } 0 \leq x \leq 1, \\ x^p + (p-1)(1-x) & \text{for } x > 1, \end{cases}$$

extended by oddness to negative numbers. By Lemma 5.8.1, we have

$$\int_\Omega \varphi(|\widehat{\mu}_\tau^\delta|) \leq \int_\Omega \varphi(|\widehat{\mu}|).$$

On the other hand, by Lemma 5.7.1 we know we have a limit point μ_τ of μ_τ^δ, as $\delta \to 0$, which minimizes (5.7.4). But the above inequality gives weak compactness in $L^p(\Omega)$ for the sequence $(\widehat{\mu}_\tau^\delta)$. The weak lower semicontinuity in $L^p(\Omega)$ of $\nu \mapsto \int_\Omega \varphi(|\nu|)$ allows to conclude. □

Now we pass to the limit as $\delta \to 0$ in (5.7.8), taking advantage of the regularity result. We are lead to

Lemma 5.8.4 *Let $p \geq 4$ and $\mu = \widehat{\mu} \in L^p(\Omega)$. There exists a minimizer μ_τ of* (5.7.4) *satisfying*

$$-\nabla h_{\mu_\tau}\widehat{\mu}_\tau = \frac{1}{\tau}\pi^1_\#(\chi_\Omega(x)(x-y)\gamma_0^+) + \frac{1}{\tau}\pi^1_\#(\chi_\Omega(x)(x-y)\gamma_0^-) \quad \text{in } \mathcal{D}'(\mathbb{R}^2), \tag{5.8.12}$$

where $\gamma_0^+ \in \Gamma_0(\mu_\tau^+, \mu_0^+)$ and $\gamma_0^- \in \Gamma_0(\mu_\tau^-, \mu_0^-)$, with respect to Notation 5.6.5.

Proof. By the previous corollaries we know that (5.7.4) possesses a minimizer with $L^4(\Omega)$ interior part. Then, we can perform the same variational argument in Lemma 5.7.2 and Corollary 5.7.3 to deduce (5.8.12). □

5.9 Back to the continuous model

Let us fix the initial datum $\mu^0 \in \mathcal{M}_{\kappa, M}(\overline{\Omega})$, let $\mu_\tau^0 = \mu^0$. We define a sequence of discrete solutions μ_τ^k. At each step, we minimize starting from the interior part of the previous point, and then we simply add its boundary part. This way, more and more mass is cumulated on the boundary at each step, and never goes back in the interior of the domain. This is reminiscent of the analysis of [8], in the framework of probability measures. Indeed, in such context it is proven that no mass enters from the boundary, by means of energetic comparison. So, the recursive scheme will be the following. Given a time step $\tau > 0$ and $\mu_\tau^k \in \mathcal{M}_{\kappa, M}(\overline{\Omega})$, define ν_τ^{k+1} as a minimizer of the discrete problem

$$\min_{\nu \in \mathcal{M}_{\kappa', M}(\overline{\Omega}),\, |\nu|(\overline{\Omega}) \le |\widehat{\mu}_\tau^k|(\Omega)} \Phi_\lambda(\nu) + \frac{1}{2\tau}\mathcal{W}_2^2(\nu, \widehat{\mu}_\tau^k), \quad k \in \mathbb{N}, \tag{5.9.1}$$

where $\kappa' = \widehat{\mu}_\tau^k(\Omega)$. Since we minimize starting from the internal part of μ_τ^k, we can choose ν_τ^{k+1} satisfying the regularity properties obtained by virtue of the entropy argument in the previous section. Then we let

$$\mathcal{M}_{\kappa, M}(\overline{\Omega}) \ni \mu_\tau^{k+1} = \nu_\tau^{k+1} + \widetilde{\mu}_\tau^k, \quad k \in \mathbb{N}. \tag{5.9.2}$$

Also, we define the piecewise constant interpolation $\overline{\mu}_\tau(t) := \mu_\tau^{\lceil t/\tau \rceil}$ for any $t \ge 0$. The following result shows that a minimizing movement does exist, as pointwise limit of $\overline{\mu}_\tau(t)$.

Proposition 5.9.1 (Existence of a limit curve) *There exists a vanishing sequence τ_n such that $\overline{\mu}_{\tau_n}(t)$ converges to $\mu(t) \in \mathcal{M}_{\kappa, M}(\overline{\Omega})$ weakly in the sense of measures, for any $t \ge 0$.*

Proof. Since ν_τ^{k+1} is a minimizer starting from $\widehat{\mu}_\tau^k$, we have as usual, for any k,

$$\mathcal{W}_2^2(\nu_\tau^{k+1}, \widehat{\mu}_\tau^k) \le 2\tau\Phi_\lambda(\widehat{\mu}_\tau^k) - 2\tau\Phi_\lambda(\nu_\tau^{k+1}).$$

Since $\widehat{\mu}_\tau^{k+1} = \widehat{\nu}_\tau^{k+1}$, and since Φ_λ depends only on the interior part of measures, we find

$$\mathcal{W}_2^2(\nu_\tau^{k+1}, \widehat{\mu}_\tau^k) \le 2\tau\Phi_\lambda(\widehat{\mu}_\tau^k) - 2\tau\Phi_\lambda(\widehat{\mu}_\tau^{k+1}). \tag{5.9.3}$$

Let us insert (5.6.17) and take (5.6.7) into account, so

$$\mathbb{W}_1^2(\mu_\tau^{k+1}, \mu_\tau^k) = \mathbb{W}_1^2(\nu_\tau^{k+1}, \widehat{\mu}_\tau^k) \le 4M\tau\Phi_\lambda(\widehat{\mu}_\tau^k) - 4M\tau\Phi_\lambda(\widehat{\mu}_\tau^{k+1}). \tag{5.9.4}$$

Of course this also implies

$$\Phi_\lambda(\mu_\tau^k) \le \Phi_\lambda(\mu^0), \quad \forall k > 0. \tag{5.9.5}$$

We introduce interpolation of minimizers $\bar{\mu}_\tau(t) := \mu_\tau^{\lceil t/\tau \rceil}$. Let $t \in (k_1\tau, (k_1+1)\tau]$ and $s \in (k_2\tau, (k_2+1)\tau]$, for some $k_1, k_2 > 0$, with $k_2 > k_1$. Summing up in (5.9.4) and making use of the triangle inequality (mind that $\mathbb{W}_1$ is a distance), along with (5.9.5) and the positiveness of Φ_λ, we have

$$\begin{aligned}\mathbb{W}_1^2(\bar{\mu}_\tau(t),\bar{\mu}_\tau(s)) &= \mathbb{W}_1^2(\mu_\tau^{k_1+1},\mu_\tau^{k_2+1})\\ &\leq (k_2-k_1)\sum_{k=k_1+1}^{k_2}\mathbb{W}_1^2(\mu_\tau^k,\mu_\tau^{k+1})\\ &\leq 2\tau(k_2-k_1)\sum_{k=k_1+1}^{k_2}\big(\Phi_\lambda(\mu_\tau^{k_1+1})-\Phi_\lambda(\mu_\tau^{k_2+1})\big)\\ &\leq 2\tau(k_2-k_1)\Phi_\lambda(\mu^0).\end{aligned}$$

Hence

$$\mathbb{W}_1(\bar{\mu}_\tau(t),\bar{\mu}_\tau(s)) \leq \sqrt{2\Phi_\lambda(\mu^0)}\sqrt{|t-s|+\tau} \quad \forall\, s,t\in[0,+\infty).$$

The discrete $C^{0,1/2}$ estimate allows to perform the usual argument of the proof of Theorem 3.4.4 and find a subsequence $\tau_n \to 0$ such that in the sense of measures

$$\lim_{n\to\infty}\bar{\mu}_{\tau_n}(t) = \mu(t) \qquad \forall t\geq 0. \tag{5.9.6}$$

This concludes the proof. □

Notation 5.9.2 The tranportation is described by the cost $\mathcal{W}_2(\nu_\tau^{k+1},\widehat{\mu}_t^k)$, described through an optimal couple of measures $(\vartheta^+)_\tau^{k+1}$, $(\vartheta^-)_\tau^{k+1}$ as in (5.6.16). That is,

$$\mathcal{W}_2^2(\nu_\tau^{k+1},\widehat{\mu}_t^k) = W_2^2((\vartheta^+)_\tau^{k+1},(\widehat{\mu}^+)_\tau^k) + W_2^2((\vartheta^-)_\tau^{k+1},(\widehat{\mu}^-)_\tau^k). \tag{5.9.7}$$

With reference to Notation 5.6.5, we let $(\vartheta_1)_\tau^{k+1}$ be the common part of $(\vartheta^+)_\tau^{k+1}$ and $(\vartheta^-)_\tau^{k+1}$. The two terms in the right hand side of (5.9.7) correspond to optimal plans $(\gamma^+)_\tau^{k+1}$ and $(\gamma^-)_\tau^{k+1}$ and can be split as

$$(\gamma^+)_\tau^{k+1} = (\gamma_0^+)_\tau^{k+1} + (\gamma_1^+)_\tau^{k+1} \quad \text{and} \quad (\gamma^-)_\tau^{k+1} = (\gamma_0^-)_\tau^{k+1} + (\gamma_1^-)_\tau^{k+1}. \tag{5.9.8}$$

Here

$$(\gamma_0^+)_\tau^{k+1} \in \Gamma_0((\nu^+)_\tau^{k+1},(\widehat{\mu}_0^+)_\tau^k) \quad \text{and} \quad (\gamma_1^+)_\tau^{k+1} \in \Gamma_0((\vartheta_1)_\tau^{k+1},(\widehat{\mu}_1^+)_\tau^k),$$

where $(\widehat{\mu}_0^+)_\tau^k$ and $(\widehat{\mu}_1^+)_\tau^k$ are suitable positive submeasures of $(\widehat{\mu}^+)_\tau^k$, which is their sum. Similarly for the negative parts.

The discrete velocity of the scheme (5.7.4) (neglecting the common parts) could be defined by $(x-y)/\tau$ with $(x,y)\in\big(\mathrm{supp}(\gamma_0^+)_\tau^{k+1}\big)\cup\big(\mathrm{supp}(\gamma_0^-)_\tau^{k+1}\big)$. The characterization of the discrete velocity is crucial to interpret our recursive scheme as the discrete version of a differential equation. But we can no more invoke Theorem 3.4.4, since we do not have a standard continuity equation for the signed case. Instead, we will see how to obtain a partial result constructing 'by hand' the limiting differential equation.

In [52], the authors are able to produce solutions to the related model (5.1.5) by means of a discrete scheme. They take advantage of strong regularity hypotheses on the initial datum, which are preserved during the evolution, guaranteeing good compactness properties. Here we would like to address the case of mere L^∞ initial data.

We start by introducing a basic estimate. With the notation above, we have shown that

$$\mathcal{W}_2^2(\nu_\tau^{k+1}, \widehat{\mu}_\tau^k) = \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2 \, d\left((\gamma_0^+)_\tau^{k+1} + (\gamma_0^-)_\tau^{k+1} + (\gamma_1^+)_\tau^{k+1} + (\gamma_1^-)_\tau^{k+1}\right)(x,y). \quad (5.9.9)$$

From (5.9.3), summing the telescopic series, we immediately see that

$$\sum_{k=0}^{\infty} \mathcal{W}_2^2(\nu_\tau^{k+1}, \widehat{\mu}_\tau^k) \le 2\tau\Phi_\lambda(\mu^0). \quad (5.9.10)$$

Hence, each of the four terms in the right hand side of (5.9.9) satisfies the same bound.

The next proposition shows that there is no contribution from the transport plans γ_1^+ and γ_1^-, which can be thought as accounting for auto-annihilation of mass, in the subsequent limit process.

Proposition 5.9.3 *Let $\phi \in C_b^1(\overline{\Omega})$. Then*

$$\lim_{\tau\to 0} \sum_{k=0}^{+\infty} \int_{\overline{\Omega}\times\overline{\Omega}} (\phi(y) - \phi(x)) \; d(\gamma_1^+)_\tau^{k+1}(x,y) = 0,$$

$$\lim_{\tau\to 0} \sum_{k=0}^{+\infty} \int_{\overline{\Omega}\times\overline{\Omega}} (\phi(y) - \phi(x)) \; d(\gamma_1^-)_\tau^{k+1}(x,y) = 0. \quad (5.9.11)$$

Proof. By definition of $\mathcal{W}_2$, and taking into account the constraint in the discrete minimization problem $|\nu_\tau^{k+1}|(\overline{\Omega}) \le |\mu_\tau^k|(\Omega)$, we see that for any k there holds $(\vartheta^+)_\tau^{k+1}(\overline{\Omega}) + (\vartheta^-)_\tau^{k+1}(\overline{\Omega}) = |\mu_\tau^k|(\Omega)$. Then, recalling that $(\vartheta_1)_\tau^k$ is the common part of $(\vartheta^+)_\tau^k$ and $(\vartheta^-)_\tau^k$, it is clear that $\sum_{k=0}^{\infty}(\vartheta_1)_\tau^k \le |\mu^0|(\overline{\Omega}) \le M$, hence

$$\sum_{k=0}^{\infty} (\gamma_1^+)_\tau^{k+1}(\overline{\Omega}\times\overline{\Omega}) + (\gamma_1^-)_\tau^{k+1}(\overline{\Omega}\times\overline{\Omega}) \le M. \quad (5.9.12)$$

Now we compute

$$\sum_{k=0}^{\infty} \int_{\overline{\Omega}\times\overline{\Omega}} (\phi(y) - \phi(x)) \, d(\gamma_1^+)_\tau^{k+1}(x,y) \le \mathrm{Lip}\phi \sum_{k=0}^{\infty} \int_{\overline{\Omega}\times\overline{\Omega}} |x-y| \, d(\gamma_1^+)_\tau^{k+1}(x,y).$$

With the Cauchy-Schwarz inequality we see that the last term is controlled by

$$\left(\mathrm{Lip}^2\phi \sum_{k=0}^{\infty} \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2 \, d(\gamma_1^+)_\tau^{k+1}(x,y)\right)^{1/2} \left(\sum_{k=0}^{\infty} \int_{\overline{\Omega}\times\overline{\Omega}} d(\gamma_1^+)_\tau^{k+1}(x,y)\right)^{1/2}.$$

But making use of (5.9.10) we see that the first factor is bounded by $\mathrm{Lip}\phi\sqrt{2\tau\Phi_\lambda(\mu^0)}$, while by (5.9.12) the second is less than or equal to $\sqrt{M}$. Hence

$$\limsup_{\tau\to 0}\sum_{k=0}^{+\infty}\int_{\overline{\Omega}\times\overline{\Omega}}(\phi(y)-\phi(x))\,d(\gamma_1^+)_\tau^{k+1}(x,y)\le \lim_{\tau\to 0} M\mathrm{Lip}\phi\sqrt{2\tau\Phi_\lambda(\mu^0)}=0.$$

Similarly one shows that (5.9.11) holds. □

We will also need the following similar result.

Proposition 5.9.4 *Let $\phi\in C_b^1(\overline{\Omega})$. There holds*

$$\lim_{\tau\to 0}\sum_{k=0}^{+\infty}\int_{\partial\Omega\times\overline{\Omega}}(\phi(x)-\phi(y))\,d(\gamma_0^+)_\tau^{k+1}(x,y)=0,$$

and the same for the analogous sum involving $(\gamma_0^-)_\tau^k$.

Proof. Reasoning as in the previous proposition, one estimates the sum above by

$$\left(\mathrm{Lip}^2\phi\sum_{k=0}^{\infty}\int_{\partial\Omega\times\overline{\Omega}}|x-y|^2\,d(\gamma_0^+)_\tau^{k+1}(x,y)\right)^{1/2}\left(\sum_{k=0}^{\infty}(\gamma_0^+)_\tau^{k+1}(\partial\Omega\times\overline{\Omega})\right)^{1/2}. \tag{5.9.13}$$

Since no mass on the boundary goes back in the interior of the domain during the discrete steps, we have

$$\sum_{k=1}^{+\infty}(\gamma_0^+)_\tau^k(\partial\Omega\times\overline{\Omega})=\sum_{k=1}^{+\infty}(\nu_\tau^k)^+(\partial\Omega)\le\sum_{k=1}^{\infty}|\nu_\tau^k|(\partial\Omega)\le|\mu^0|(\overline{\Omega})\le M.$$

This shows that the second factor in (5.9.13) is uniformly bounded. The first one is controlled again by $\mathrm{Lip}\phi\sqrt{2\tau\Phi_\lambda(\mu^0)}$, as a consequence of (5.9.9) and (5.9.10). The same argument gives the thesis if $(\gamma_0^+)_\tau^k$ are replaced by $(\gamma_0^-)_\tau^k$. □

Lemma 5.9.5 (Convergence of the total variation) *Let (τ_n) be a vanishing sequence. Then there exist positive measures $\varrho^+(t)$, $\varrho^-(t)$ such that, possibly on a subsequence, there holds*

$$\overline{\mu}_{\tau_n}^+(t)\rightharpoonup\varrho^+(t),\quad \overline{\mu}_{\tau_n}^-(t)\rightharpoonup\varrho^-(t),\quad |\overline{\mu}_{\tau_n}(t)|\rightharpoonup\varrho^+(t)+\varrho^-(t). \tag{5.9.14}$$

Proof. We prove the convergence of positive parts. Then one reasons analogously for negative parts. Let φ be a bounded Lipschitz function over $\overline{\Omega}$. Possibly adding a constant, we can assume that φ is nonnegative. Let

$$\begin{aligned} a_\tau^k &:= \int_{\overline{\Omega}}\varphi\,d(\nu^+)_\tau^{k+1}-\int_{\overline{\Omega}}\varphi\,d(\widehat{\mu}_0^+)_\tau^k,\\ b_\tau^k &:= \int_{\overline{\Omega}}\varphi\,d(\widehat{\mu}_1^+)_\tau^k. \end{aligned}$$

We have, by (5.9.9),

$$\begin{aligned}
a_\tau^k &= \int_{\overline{\Omega}\times\overline{\Omega}} (\varphi(y)-\varphi(x))\, d(\gamma_0^+)_\tau^{k+1}(x,y)\\
&\le \mathrm{Lip}(\varphi)\left(\int_{\overline{\Omega}\times\overline{\Omega}} d(\gamma_0^+)_\tau^{k+1}\right)^{1/2}\left(\int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2\, d(\gamma_0^+)_\tau^{k+1}(x,y)\right)^{1/2}\\
&\le \mathrm{Lip}(\varphi)\sqrt{M}\,\mathcal{W}_2(\nu_\tau^{k+1},\widehat{\mu}_\tau^k),
\end{aligned}$$

which gives, making use of (5.9.10),

$$\sum_{k=0}^\infty \frac{(a_\tau^k)^2}{\tau} \le \frac{1}{\tau}\,\mathrm{Lip}(\varphi)\sqrt{M}\sum_{k=0}^\infty \mathcal{W}_2^2(\mu_\tau^{k+1},\mu_\tau^k) \le 2\mathrm{Lip}(\varphi)\sqrt{M}\,\Phi_\lambda(\mu^0).$$

This implies, as $\tau\to 0$, the convergence in measure (on a suitable sequence)

$$A_\tau := \sum_{k=0}^\infty a_\tau^k \delta_{\{k\tau\}} \rightharpoonup f\mathcal{L}^1,$$

where $f\in L^2(0,+\infty)$. Hence, on the suitable sequence, for any t there holds

$$\lim_{\tau\downarrow 0}\sum_{k=0}^\infty a_\tau^k\delta_{\{k\tau\}}([0,t]) \to \int_0^t f(y)\,dy.$$

Next, notice that, by (5.9.2),

$$\frac{d}{dt}\int_{\overline{\Omega}} \varphi\, d\overline{\mu}_\tau^+(t) = \sum_{k=0}^\infty (a_\tau^k - b_\tau^k)\delta_{\{k\tau\}}.$$

As $b_\tau^k \le 0$, we see that

$$\frac{d}{dt}\left(\int_{\overline{\Omega}} \varphi\, d\overline{\mu}_\tau^+(t) - A_\tau([0,t])\right) \le 0.$$

We have a family of monotone functions. We can apply Helly pointwise compactness theorem (see for instance [4, Lemma 3.3.3]), finding a vanishing sequence τ_n such that a pointwise, nonincreasing limit exists. We shall denote such limit (the part relative to the first term) by $L_\varphi(t)$, that is,

$$\int_{\overline{\Omega}} \varphi\, d\overline{\mu}_\tau^+(t) \to L_\varphi(t), \quad \forall t\ge 0.$$

The convergence holds for any fixed Lipschitz function φ. By a diagonal argument we can find a vanishing sequence, that we still denote by τ_n, such that

$$\int_{\overline{\Omega}} \varphi\, d\overline{\mu}_\tau^+(t) \to L_\varphi(t), \quad \forall t\ge 0 \text{ and } \varphi\in\mathcal{D},$$

where $\mathcal{D}$ is a countable dense subset of $C_b^0(\overline{\Omega})$. Then, for any fixed t, $L_\varphi(t)$ can be extended uniquely to a weakly continuous linear functional on $C_b^0(\overline{\Omega})$. By the Riesz representation theorem we conclude that $L_\varphi(t) = \int_{\overline{\Omega}} \varphi \, d\varrho^+(t)$, for some $\varrho^+ \in \mathcal{M}^+(\overline{\Omega})$, and for any t there holds $\overline{\mu}_\tau^+(t) \rightharpoonup \varrho^+(t)$.

Letting $\varrho^-(t)$ be the pointwise weak limit of $\overline{\mu}_\tau^-(t)$, we also infer the convergence of the total variation:

$$|\overline{\mu}_\tau(t)| \rightharpoonup \varrho(t) \quad \forall t \geq 0,$$

where $\varrho(t) = \varrho^+(t) + \varrho^-(t)$. □

Eventually, we are able to produce a limiting equation.

Theorem 5.9.6 (Equation in the limit) *The minimizing movement $\mu(t)$ given by* Proposition 5.9.1 *satisfies*

$$\frac{d}{dt}\mu(t) - \operatorname{div}\left(\chi_\Omega \nabla h_{\mu(t)} \varrho(t)\right) = 0 \quad \text{in } \mathcal{D}'((0,+\infty) \times \mathbb{R}^2), \tag{5.9.15}$$

where $\varrho(t)$ is a suitable positive measure satisfying $\widehat{\varrho}(t) \geq |\widehat{\mu}(t)|$.

Proof. Let $\phi \in C^2(\overline{\Omega})$. Let us compute the derivative of the step measure $\bar{\mu}_\tau(t)$ (this is similar to the proof of Theorem 3.5.8). We have, in the sense of distributions,

$$\frac{d}{dt}\int_{\overline{\Omega}} \phi \, d\bar{\mu}_\tau(t) = \sum_{k=0}^{+\infty} \left(\int_{\overline{\Omega}} \phi \, d\mu_\tau^{k+1} - \int_{\overline{\Omega}} \phi \, d\mu_\tau^k \right) \delta_{\{k\tau\}}.$$

But

$$\int_{\overline{\Omega}} \phi \, d\mu_\tau^{k+1} - \int_{\overline{\Omega}} \phi \, d\mu_\tau^k = \int_{\overline{\Omega}} \phi \, d\nu_\tau^{k+1} - \int_{\overline{\Omega}} \phi \, d\widehat{\mu}_\tau^k = \int_{\overline{\Omega}\times\overline{\Omega}} (\phi(x) - \phi(y)) \, d\gamma_\tau^{k+1},$$

where, using the notation introduced in (5.9.8),

$$\gamma_\tau^{k+1} := (\gamma_0^+)_\tau^{k+1} - (\gamma_0^-)_\tau^{k+1} + (\gamma_1^+)_\tau^{k+1} - (\gamma_1^-)_\tau^{k+1}.$$

So we can write

$$\begin{aligned} \frac{d}{dt}\int_{\overline{\Omega}} \phi \, d\bar{\mu}_\tau(t) &= \sum_{k=0}^{+\infty} \delta_{\{k\tau\}} \int_{\overline{\Omega}\times\overline{\Omega}} (\phi(x) - \phi(y)) \, d\gamma_\tau^{k+1}(x,y) \\ &= \sum_{k=0}^{+\infty} \delta_{\{k\tau\}} \left(\int_{\overline{\Omega}\times\overline{\Omega}} \langle \nabla\phi(x), x-y \rangle \, d\gamma_\tau^{k+1}(x,y) + \mathcal{R}_\tau^k \right) \end{aligned} \tag{5.9.16}$$

Let us estimate the remainder $\mathcal{R}_\tau^k$ by writing it in integral form. We have

$$\begin{aligned} \mathcal{R}_\tau^k &= \frac{1}{2}\int_0^1 \int_{\overline{\Omega}\times\overline{\Omega}} |\langle \nabla^2\phi((1-\theta)x + \theta y)(y-x), y-x \rangle| d\gamma_\tau^{k+1}(x,y) d\theta \\ &\leq \frac{1}{2} \sup |\nabla^2 \phi| \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2 \, d\gamma_\tau^{k+1}(x,y) \\ &\leq \frac{1}{2} \sup |\nabla^2\phi| \, \mathcal{W}_2^2(\nu_\tau^{k+1}, \widehat{\mu}_\tau^k) \end{aligned} \tag{5.9.17}$$

By (5.9.10) we see that

$$\lim_{\tau\downarrow 0}\sum_{k=0}^{\infty}\mathcal{R}_\tau^k = 0.$$

Together with Proposition 5.9.3 and Proposition 5.9.4 this shows that (5.9.16) can be written as

$$\frac{d}{dt}\int_{\overline{\Omega}}\phi\, d\bar{\mu}_\tau(t) = \sum_{k=0}^{+\infty}\delta_{\{k\tau\}}\left(\int_{\overline{\Omega}\times\overline{\Omega}}\langle\nabla\phi(x), x-y\rangle\chi_\Omega(x)\, d\left((\gamma_0^+)_\tau^{k+1} - (\gamma_0^-)_\tau^{k+1}\right)(x,y)\right) + o(\tau). \tag{5.9.18}$$

The Euler equation for discrete minimizers ν_τ^k of (5.7.4), since $\widehat{\nu}_\tau^k = \widehat{\mu}_\tau^k$, reads (see Lemma 5.8.4)

$$-\nabla h_{\mu_\tau^k}\left((\widehat{\mu}_\tau^k)^+ - (\widehat{\mu}_\tau^k)^-\right) = \frac{1}{\tau}\left(\pi_\#^1(\chi_\Omega(x)(x-y)(\gamma_0^+)_\tau^k) + \pi_\#^1(\chi_\Omega(x)(x-y)(\gamma_0^-)_\tau^k)\right),$$

but notice that the first term in the right hand side can be different from zero only on $\mathrm{supp}(\widehat{\mu}_\tau^k)^+$. Similarly for the second term. Hence we can split the equation in

$$\begin{aligned} -\nabla h_{\mu_\tau^k}(\widehat{\mu}_\tau^k)^+ &= \frac{1}{\tau}\pi_\#^1(\chi_\Omega(x)(x-y)(\gamma_0^+)_\tau^k), \\ \nabla h_{\mu_\tau^k}(\widehat{\mu}_\tau^k)^- &= \frac{1}{\tau}\pi_\#^1(\chi_\Omega(x)(x-y)(\gamma_0^-)_\tau^k). \end{aligned}$$

Substituting in (5.9.18), we find

$$\frac{d}{dt}\int_{\overline{\Omega}}\phi\, d\bar{\mu}_\tau(t) = -\sum_{k=0}^{+\infty}\tau\delta_{\{k\tau\}}\int_\Omega\left\langle\nabla\phi(x), \nabla h_{\mu_\tau^k}(x)\right\rangle\, d|\widehat{\mu}_\tau^k|(x) + o(\tau).$$

Passing to the limit as τ goes to zero, along a suitable sequence τ_n, we get (by the interior regularity of μ_τ^k, following from the entropy argument)

$$\frac{d}{dt}\int_{\overline{\Omega}}\phi\, d\mu(t) + \int_\Omega\langle\nabla\phi, \nabla h_{\mu(t)}\rangle\, d\varrho(t) = 0,$$

where $\varrho(t)$ is some positive measure, pointwise limit of the total variation of $\mu_\tau^{\lceil t/\tau\rceil}$ (see Lemma 5.9.5), hence satisfying $\widehat{\varrho}(t) \geq |\widehat{\mu}(t)|$ and $\varrho(t)(\overline{\Omega}) \leq |\mu^0|(\overline{\Omega})$. □

Chapter 6

Stationary configurations for the average distance functional

6.1 Introduction

In this chapter we consider functionals $\mathcal{F}(\Sigma)$ defined on the class of all closed connected subsets of $\mathbb{R}^n$ and the corresponding minimization problems

$$\min\left\{\mathcal{F}(\Sigma)\ :\ \Sigma \text{ closed connected subset of } \mathbb{R}^n\right\}. \tag{6.1.1}$$

Due to the fact that the class of closed connected sets has good compactness properties with respect to the Hausdorff convergence, mild coercivity assumptions on $\mathcal{F}$ give the existence of minimizers for problem (6.1.1). We are interested in finding "first order" necessary optimality conditions satisfied by the minimizers Σ of (6.1.1).

The case we consider is the *average distance functional*

$$\mathcal{F}(\Sigma) := \int_{\mathbb{R}^n} \operatorname{dist}(x,\Sigma)\, d\mu(x) + \lambda\mathcal{H}^1(\Sigma), \tag{6.1.2}$$

where μ is a given finite nonnegative Borel measure over $\mathbb{R}^n$ with compact support, and the penalization term $\lambda\mathcal{H}^1(\Sigma)$ with $\lambda > 0$ is added to give a suitable coercivity to $\mathcal{F}$ and to prevent minimizing sequences from spreading over all the space. A simple and standard argument involving Blaschke and Gołąb theorems gives the existence of minimizers of $\mathcal{F}$. Of particular interest for us will be situations when μ is a uniform measure over some open set $\Omega \subset \mathbb{R}^n$, i.e. $\mu = \mathcal{L}^n \llcorner \Omega$.

The average distance term in (6.1.2) comes from mass transport theory and describes for instance the total transportation cost to move a mass μ of residents to a public transport network Σ. This last is the unknown of the problem and has to be designed in order to minimize $\mathcal{F}$ also taking into account the construction costs which here are taken as proportional to $\mathcal{H}^1(\Sigma)$. The minimization problem (6.1.1), as well as some qualitative properties of its minimizers, have been studied in several recent papers (see e.g. [20, 22, 23, 24, 47, 61, 66, 68]) to which we refer the interested reader. Our goal is to find "first order" conditions of differential character satisfied by the minimizers of (6.1.2). Such conditions will open the way to define a natural notion of stationary (or critical) points of (6.1.2). The main difficulty, which is quite common in shape optimization problems, is that the domain of definition of this functional (i.e. the class of closed connected subsets of $\mathbb{R}^n$) does not possess any natural differentiable structure, and the usual "first variation" argument has to be intended in a suitable way.

In the last section we consider a similar case arising from the theory of elliptic equations:

$$\mathcal{F}(\Sigma) := \int_\Omega u_\Sigma(x) f(x)\, dx + \lambda \mathcal{H}^1(\Sigma), \tag{6.1.3}$$

where $\Omega \subset \mathbb{R}^2$ is a given bounded open subset, f is a given $L^2(\Omega)$ function, and u_Σ is the unique solution of the PDE

$$\begin{cases} -\Delta u &= f \quad \text{in } \Omega \setminus \Sigma, \\ u &= 0 \quad \text{on } \partial\Omega \cup \Sigma. \end{cases}$$

One has to remark that while a lot of properties are known for minimizers of the average distance functional (see [20, 23, 61, 68]), like partial regularity, absence of loops, topological properties (finite number of branching points, each of which is a regular tripod), no such property has been studied for minimizers of (6.1.3).

6.2 Euler equation for the average distance functional

For a compact set $\Sigma \subset \mathbb{R}^n$ we denote by π^Σ the projection map to Σ (i.e. such that $\pi^\Sigma(x) \in \Sigma$ is one of the nearest point in Σ to $x \in \mathbb{R}^n$). This map is uniquely defined everywhere except the *ridge* set $\mathfrak{R}_\Sigma$, which is defined as the set of all $x \in \mathbb{R}^n$ for which the minimum distance to Σ is attained at more than one point. It is well known that $\mathfrak{R}_\Sigma$ is the set of non differentiability points of the distance function to Σ (that is, of the map $x \in \mathbb{R}^n \mapsto \operatorname{dist}(x, \Sigma)$), and since the latter map is semiconcave, this set is an $(\mathcal{H}^{n-1}, n-1)$-rectifiable Borel set [51, proposition 3.7].

We will denote by $B_r(x) \subset \mathbb{R}^n$ the open ball with radius $r > 0$ and center $x \in \mathbb{R}^n$. The line segment with endpoints A and B will be denoted by $\overline{AB}$, the arc of a curve with the same endpoints will be denoted by $\widetilde{AB}$ (usually in this paper we will deal with arcs of circle).

To be begin with, we estimate the ascending local slope of (6.1.2), defined by

$$|\mathcal{F}'|(\Sigma) := \limsup_{d_H(\Sigma',\Sigma)\to 0} \frac{(\mathcal{F}(\Sigma') - \mathcal{F}(\Sigma))^+}{d_H(\Sigma', \Sigma)},$$

where d_H stands for Hausdorff distance between sets. The following simple assertion is valid.

Proposition 6.2.1 *If $\mu(\mathfrak{R}_\Sigma) = 0$, one has $|\mathcal{F}'|(\Sigma) \geq \lambda$.*

Proof. Let $x \in \Sigma$ be such that $\mu((\pi^\Sigma)^{-1}(\{x\})) = 0$ (all but a countable number of points of Σ have this property). Let then $\Sigma_\varepsilon := \Sigma \cup I_\varepsilon$, where I_ε stands for the line segment of length $\varepsilon > 0$, with one of the endpoints x and such that $\pi^\Sigma(I_\varepsilon) = x$. Then $d_H(\Sigma_\varepsilon, \Sigma) = \varepsilon$ and $\mathcal{H}^1(\Sigma_\varepsilon) = \mathcal{H}^1(\Sigma) + \varepsilon$. On the other hand, denoting

$$G_\varepsilon := \{z \in \mathbb{R}^n \, : \, \mathrm{dist}\,(z, \Sigma) \geq \mathrm{dist}\,(z, I_\varepsilon)\},$$

we have that

$$\int_{\mathbb{R}^n} \mathrm{dist}\,(x, \Sigma)\, d\mu(x) \geq \int_{\mathbb{R}^n} \mathrm{dist}\,(x, \Sigma_\varepsilon)\, d\mu(x) \geq \int_{\mathbb{R}^n} \mathrm{dist}\,(x, \Sigma)\, d\mu(x) - \varepsilon\mu(G_\varepsilon).$$

Thus

$$|\mathcal{F}'|(\Sigma) \geq \limsup_{d_H(\Sigma_\varepsilon, \Sigma)\to 0} \frac{(\mathcal{F}(\Sigma_\varepsilon) - \mathcal{F}(\Sigma))^+}{d_H(\Sigma_\varepsilon, \Sigma)} \geq \lim_{\varepsilon\to 0^+} \frac{(\lambda\varepsilon - \varepsilon\mu(G_\varepsilon))^+}{\varepsilon},$$

and to conclude the proof it suffices to mind that $\mu(G_\varepsilon) = o(1)$, because $G_\varepsilon \searrow \{x\}$ as $\varepsilon \to 0^+$. □

The above proposition in fact means that for the functional (6.1.2) no set Σ is stationary in the strong sense, i.e. is such that

$$\mathcal{F}(\Sigma') = \mathcal{F}(\Sigma) + o(d_H(\Sigma, \Sigma'))$$

as $\Sigma' \to \Sigma$ in Hausdorff distance. In other words there is no differentiability with respect to d_H. Therefore, in search for the natural notion of stationary points of $\mathcal{F}$ we have to restrict the set of admissible variations of Σ. For this purpose let $\phi_\varepsilon \colon \mathbb{R}^n \to \mathbb{R}^n$ be a one parameter group of diffeomorphisms satisfying

$$\phi_\varepsilon(x) = x + \varepsilon X(x) + o(\varepsilon), \tag{6.2.1}$$

as $\varepsilon \to 0$, where $X \in C_0^\infty(\mathbb{R}^n; \mathbb{R}^n)$. We will write Euler equation for the functional (6.1.2) by considering admissible variations of the type $\Sigma_\varepsilon := \phi_\varepsilon(\Sigma)$.

We recall the notion of a generalized mean curvature from [18]. The generalized mean curvature H_Σ of a $(\mathcal{H}^k, k)$-rectifiable set $\Sigma \subset \mathbb{R}^n$ (or, in terms of [18], of the measure $\mathcal{H}^k \llcorner \Sigma$) is the vector-valued distribution defined by the relationship

$$\langle X, H_\Sigma \rangle := -\int_\Sigma \mathrm{div}^{\,\Sigma} X \, d\mathcal{H}^k$$

for all $X \in C_0^\infty(\mathbb{R}^n, \mathbb{R}^n)$, where $\mathrm{div}^{\,\Sigma}$ stands for the tangential divergence operator (i.e. projection of the divergence to the approximate tangent space of Σ at $\mathcal{H}^k$-a.e. point of Σ). We have then the following result.

Theorem 6.2.2 *Let μ be a Borel measure such that $\mu(E) = 0$ whenever $\mathcal{H}^{n-1}(E) < +\infty$. Then for all $X \in C_0^\infty(\mathbb{R}^n; \mathbb{R}^n)$ one has*

$$\begin{aligned} \frac{\partial}{\partial\varepsilon} \mathcal{F}(\Sigma_\varepsilon)\Big|_{\varepsilon=0} &= \int_{\mathbb{R}^n} \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle d\mu - \lambda\langle H_\Sigma, X\rangle \\ &= \int_{\mathbb{R}^n} \left\langle X(\pi^\Sigma(x)), \nabla \mathrm{dist}\,(x, \Sigma) \right\rangle d\mu - \lambda\langle H_\Sigma, X\rangle. \end{aligned} \tag{6.2.2}$$

In particular, if Σ is a minimizer of $\mathcal{F}$, then

$$\int_{\mathbb{R}^n} \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle d\mu = \lambda \langle H_\Sigma, X\rangle \quad \forall X \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n). \tag{6.2.3}$$

Proof. First of all, we perform the variation for the first term. We adopt the method of calculation of the derivative of the distance function with respect to the variation of the set, used in [7, lemma 4.5]. Clearly, for $z := \phi_\varepsilon(\pi^\Sigma(x))$ one has

$$\begin{aligned} \mathrm{dist}\,(x,\Sigma) &= |\pi^\Sigma(x)-x|, \\ \mathrm{dist}\,(x,\Sigma_\varepsilon) &\le |z-x|. \end{aligned}$$

From (6.2.1) we get, for $\varepsilon \to 0$,

$$\begin{aligned} |z-x|^2 &= \left\langle \pi^\Sigma(x)-x+\varepsilon X(\pi^\Sigma(x)), \pi^\Sigma(x)-x+\varepsilon X(\pi^\Sigma(x))\right\rangle + o(\varepsilon) \\ &= |\pi^\Sigma(x)-x|^2 + 2\left\langle \pi^\Sigma(x)-x, \varepsilon X(\pi^\Sigma(x))\right\rangle + o(\varepsilon) \\ &= |\pi^\Sigma(x)-x|^2 \left(1+2\left\langle \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|^2}, \varepsilon X(\pi^\Sigma(x))\right\rangle + o(\varepsilon)\right) \end{aligned}$$

Then

$$\begin{aligned} \mathrm{dist}\,(x,\Sigma_\varepsilon) - \mathrm{dist}\,(x,\Sigma) &\le |z-x| - |\pi^\Sigma(x)-x| \\ &= \varepsilon \left\langle \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|}, X(\pi^\Sigma(x))\right\rangle + o(\varepsilon), \end{aligned}$$

and we deduce

$$\limsup_{\varepsilon\to 0} \frac{1}{\varepsilon}(\mathrm{dist}\,(x,\Sigma_\varepsilon) - \mathrm{dist}\,(x,\Sigma)) \le \left\langle \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|}, X(\pi^\Sigma(x))\right\rangle. \tag{6.2.4}$$

On the other hand, consider a sequence $\varepsilon_\nu \to 0^+$ for $\nu \to \infty$. The set of points $x \in \mathbb{R}^n$ for which both $\pi^\Sigma(x)$ and $\pi^{\Sigma_{\varepsilon_\nu}}(x)$ are singletons for any $\nu \in \mathbb{N}$ is of full measure μ in $\mathbb{R}^n$ (the complement is a countable union of μ-negligible sets). For all such x, since ϕ_ε is invertible for all sufficiently small ε, let $\zeta := \phi_{\varepsilon_\nu}^{-1}(\pi^{\Sigma_{\varepsilon_\nu}}(x))$, so that

$$\begin{aligned} \mathrm{dist}\,(x,\Sigma_{\varepsilon_\nu}) &= |\phi_{\varepsilon_\nu}(\zeta)-x|, \\ \mathrm{dist}\,(x,\Sigma) &\le |\zeta-x|. \end{aligned}$$

Again we have

$$\begin{aligned} |\phi_{\varepsilon_\nu}(\zeta)-x| - |\zeta-x| &= |\zeta-x|\left(\sqrt{1+2\left\langle \frac{\zeta-x}{|\zeta-x|^2}, \varepsilon_\nu X(\zeta)\right\rangle + o(\varepsilon_\nu)} - 1\right) \\ &= \varepsilon_\nu \left\langle \frac{\zeta-x}{|\zeta-x|}, X(\zeta)\right\rangle + o(\varepsilon_\nu). \end{aligned}$$

Therefore,

$$\operatorname{dist}(x,\Sigma_{\varepsilon_\nu}) - \operatorname{dist}(x,\Sigma) \geq \varepsilon_\nu \left\langle \frac{\zeta - x}{|\zeta - x|}, X(\zeta) \right\rangle + o(\varepsilon_\nu).$$

Passing to the limit as $\nu \to \infty$, we get

$$\left\langle \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|}, X(\pi^\Sigma(x)) \right\rangle \leq \liminf_{\nu\to\infty} \frac{1}{\varepsilon_\nu}(\operatorname{dist}(x,\Sigma_{\varepsilon_\nu}) - \operatorname{dist}(x,\Sigma). \tag{6.2.5}$$

Combining (6.2.4) with (6.2.5), we get for μ-a.e. $x \in \mathbb{R}^n$,

$$\lim_{\nu\to\infty} \frac{1}{\varepsilon_\nu}(\operatorname{dist}(x,\Sigma_{\varepsilon_\nu}) - \operatorname{dist}(x,\Sigma)) = \left\langle \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|}, X(\pi^\Sigma(x)) \right\rangle,$$

so that, by the dominated convergence theorem,

$$\lim_{\nu\to\infty} \frac{1}{\varepsilon_\nu} \int_\Omega (\operatorname{dist}(x,\Sigma_{\varepsilon_\nu}) - \operatorname{dist}(x,\Sigma))\, d\mu = \int_\Omega \left\langle \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|}, X(\pi^\Sigma(x)) \right\rangle d\mu.$$

Since the sequence ε_ν is arbitrary, one has

$$\lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon} \int_\Omega (\operatorname{dist}(x,\Sigma_{\varepsilon}) - \operatorname{dist}(x,\Sigma))\, d\mu = \int_\Omega \left\langle \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|}, X(\pi^\Sigma(x)) \right\rangle d\mu.$$

Finally, we observe that according to Theorem 1.6.6 one has

$$\frac{d}{d\varepsilon}\mathcal{H}^k(\Sigma_\varepsilon)\Big|_{\varepsilon=0} = \int_\Sigma \operatorname{div}^\Sigma X\, d\mathcal{H}^k = -\langle H_\Sigma, X\rangle,$$

which concludes the proof. □

We are in a position to give the following definition.

Definition 6.2.3 *A closed connected set $\Sigma \subset \mathbb{R}^n$ will be called stationary for the functional $\mathcal{F}$, if* (6.2.3) *holds.*

Clearly, every stationary point depends on the problem data, which in this case is the measure μ. To emphasize this dependence, we will further sometimes say for stationary points for the functional $\mathcal{F}$ that they are stationary with respect to μ. In the most important particular case we will be interested in μ is a uniform measure over some open $\Omega \subset \mathbb{R}^n$ (i.e. $\mu = \mathcal{L}^n \llcorner \Omega$) with $\Sigma \subset \Omega$. In such a situation we will be speaking of stationary points with respect to the set Ω.

6.3 Examples of regular stationary points

We will first show that, in sharp contrast with minimizers, stationary points may contain closed loops (i.e. homeomorphic images of S^1).

Proposition 6.3.1 *Let $\mu := \mathcal{L}^2 \llcorner B_1(0)$. There exist $r < 1$ such that the circumference $\partial B_r(0)$ is a stationary point for functional (6.1.2) if and only if $\lambda < \frac{1}{2}$. Nevertheless, no circumference is a minimizer of (6.1.2), since minimizers cannot contain closed loops.*

Proof. We set $\Sigma := \partial B_r(0)$ and impose (6.2.3). We choose X to be normal to Σ without loss of generality, since only the normal part plays a role in (6.2.3). If we write the integral term in polar coordinates, the integrand depends only on the angle. Setting $A = B_r(0)$ and $B = B_1(0)\backslash B_r(0)$, and letting $\nu(x)$ be the outward unit normal to $\partial B_r(0)$, we get

$$\begin{aligned}\int_\Omega \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|}\right\rangle dx &= \int_A \langle X(\pi^\Sigma(x)), \nu(\pi^\Sigma(x))\rangle\, dx \\ &\quad - \int_B \langle X(\pi^\Sigma(x)), \nu(\pi^\Sigma(x))\rangle\, dx,\end{aligned}$$

and we can compute

$$\begin{aligned}\int_A \langle X(\pi^\Sigma(x)), \nu(\pi^\Sigma(x))\rangle\, dx &= \int_0^r \int_0^{2\pi} |X(\theta)|\, \rho\, d\rho d\theta \\ &= \frac{1}{2} r^2 \int_0^{2\pi} |X(\theta)| d\theta,\end{aligned}$$

and similarly for the integral over B. Moreover,

$$\langle X, H_\Sigma\rangle = -\int_{\partial B_r(0)} |H_\Sigma(x)| \langle X(x), \nu(x)\rangle d\mathcal{H}^1(x) = -\frac{1}{r}\int_0^{2\pi} |X(\theta)| r d\theta.$$

So the Euler equation reads

$$\left(r^2 - \frac{1}{2} + \lambda\right) \int_0^{2\pi} |X(\theta)| d\theta = 0. \tag{6.3.1}$$

This equation is identically satisfied, if and only if $\lambda < 1/2$, for $r = \sqrt{1/2 - \lambda}$ (of course, $\lambda = 1/2$ would also be suited for (6.3.1), but it corresponds to a degenerate case when the circumference reduces to a point).

To show that minimizers of (6.1.2) cannot contain closed loops, and hence the above stationary points are not minimizers, we may proceed as in the proof of absence of loops in minimizers of average distance functionals with length constraint (see e.g. [61] or [20, 23]). In fact, suppose that Σ is a minimizer containing a closed loop. Then there is a set of positive length $C \subset \Sigma$ such that for every $x \in C$ and for every $\varepsilon > 0$ there is a closed connected subset $D_\varepsilon \subset \Sigma$ such that $x \in D_\varepsilon$, $\operatorname{diam} D_\varepsilon = \varepsilon$ (hence $\mathcal{H}^1(D_\varepsilon) \geq \varepsilon$) and $\Sigma_\varepsilon :=$

$\Sigma \setminus D_\varepsilon$ is connected. We may suppose without loss of generality that $\mu((\pi^\Sigma)^{-1}(\{x\})) = 0$ for all $x \in C$ (since the set of atoms of the latter measure is clearly at most countable). One has then by triangle inequality

$$\int_{\mathbb{R}^n} \mathrm{dist}\,(x, \Sigma_\varepsilon)\, d\mu(x) \leq \int_{\mathbb{R}^n} \mathrm{dist}\,(x, \Sigma)\, d\mu(x) + \varepsilon\mu((\pi^\Sigma)^{-1}(D_\varepsilon)),$$

and hence

$$\mathcal{F}(\Sigma_\varepsilon) \leq \mathcal{F}(\Sigma) + \varepsilon\mu((\pi^\Sigma)^{-1}(D_\varepsilon)) - \lambda\varepsilon.$$

Minding that $D_\varepsilon \searrow \{x\}$ as $\varepsilon \to 0^+$, we get

$$\mu((\pi^\Sigma)^{-1}(D_\varepsilon)) \to \mu((\pi^\Sigma)^{-1}(\{x\})) = 0,$$

and thus

$$\mathcal{F}(\Sigma_\varepsilon) \leq \mathcal{F}(\Sigma) + o(\varepsilon) - \lambda\varepsilon$$

as $\varepsilon \to 0^+$, which means that $\mathcal{F}(\Sigma_\varepsilon) < \mathcal{F}(\Sigma)$ for small $\varepsilon > 0$ concluding the proof. □

Let us now consider another example of a stationary point for (6.1.2) given by Figure 6.1, where the radii of the semicircles are equal to $\sqrt{\lambda}$. Here, as well as in all the other figures, the arrows starting at the endpoints of Σ indicate the directions of $-H_\Sigma$ in these points.

Proposition 6.3.2 *There exists a line segment which is stationary for the region Ω shown on* Figure 6.1*.*

Proof. In the example of Figure 6.1, points belonging to regions A and B are projected to the line segment Σ along the perpendicular, and it is clear that the symmetry of the domain yields

$$\int_A \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx + \int_B \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx = 0$$

for any vector field $X \in C_0^\infty(\mathbb{R}^n; \mathbb{R}^n)$.

Set $X_1 := \langle X, \mathbf{e}_1 \rangle$ and $X_2 := \langle X, \mathbf{e}_2 \rangle$, where $\mathbf{e}_1$, $\mathbf{e}_2$ stand for the basis vectors in $\mathbb{R}^2$. Let us compute the contribution of the right unit semicircle:

$$\begin{aligned} \int_D \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx &= -\int_0^{\sqrt{\lambda}} \int_{-\pi/2}^{\pi/2} X_1(F) \cos\theta \rho d\rho d\theta \\ &= -2X_1(F) \int_0^{\sqrt{\lambda}} \rho d\rho \\ &= -\lambda X_1(F). \end{aligned}$$

In the same way, the contribution of the semicircle C is given by

$$\int_C \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx = \lambda X_1(E).$$

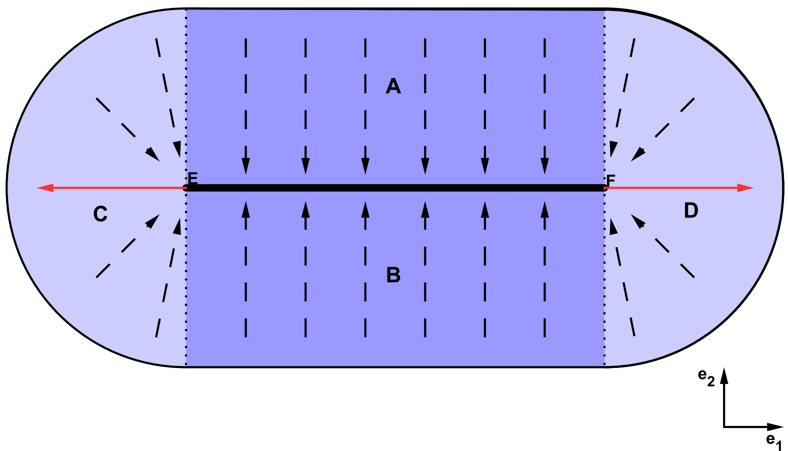

Figure 6.1: Construction of the proof of Proposition 6.3.2

Therefore,

$$\int_\Omega \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx = -\lambda X_1(F) + \lambda X_1(E).$$

On the other hand, at the endpoints E and F of the segment, the distributional curvature is given by $\delta_E \mathbf{e}_1$, $-\delta_F \mathbf{e}_1$ where δ_x stands for the Dirac mass concentrated at the point x (see [18, 38]), while at all the other points of the segment the curvature is zero. Thus the curvature term of the Euler equation reduces to $\lambda(X_1(F) - X_1(E))$, and hence (6.2.3) is satisfied. □

We now show an example of a set which is never stationary (i.e. it is not stationary for any ambient set Ω).

Proposition 6.3.3 *The line Σ made of two segments (not reduced to a single segment), is not stationary for any open set $\Omega \subset \mathbb{R}^2$.*

Proof. Let P be the common vertex of the two segments (with the aperture $2\varphi < \pi$), R be a point on one of the two edges, with $z := |P - R|$. Let moreover S be a point on the normal to the same segment passing through R, with $y := |S - R|$ located in the region B in Figure 6.2. Since the whole polygonal line Σ, and hence P, is in the interior of Ω, it is clear that the rectangle $B := PRST$ (with sidelengths z and y), is all contained in Ω for all sufficiently small y and z. Let finally Q be a point of the intersection of the line passing through S and R, with the bisector of the angle formed by the two segments of Σ (see Figure 6.2). Choose now a regular vector field X compactly supported in the open

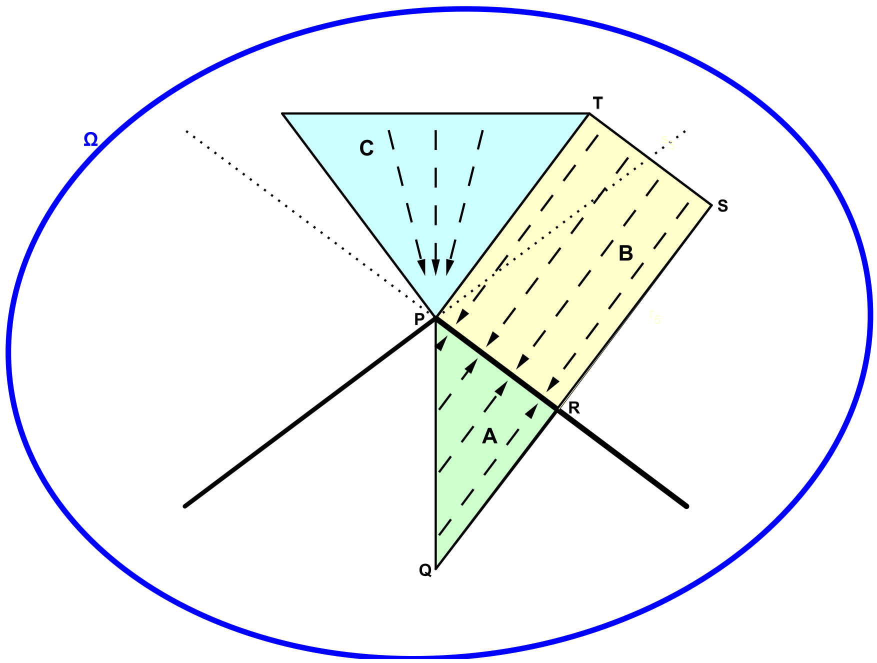

Figure 6.2: Construction of the proof of Proposition 6.3.3

segment $\overline{PR}$, and normal to it, pointing towards the region B in Figure 6.2. It is clear that there is no contribution from the curvature term in the Euler equation, since the curvature of the line segment is zero outside its endpoints. So it remains to check the integral term. Since $|Q-R| = z\tan\varphi$, an easy computation in the suitable coordinate system yields

$$\int_B \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx = -\int_B |X(\pi^\Sigma(x))|\, dx = -y\int_0^z |X(\zeta)| d\zeta$$

and

$$\int_A \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx = \int_A |X(\pi^\Sigma(x))|\, dx = z\tan\varphi\int_0^z |X(\zeta)| d\zeta.$$

Notice that z can be chosen small enough such that the sum of the above terms is strictly negative, while

$$\int_\Omega \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx \le \int_A \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx + \int_B \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx,$$

that is, for sufficiently small z the equation (6.2.3) is not satisfied. □

It is worth emphasizing that it is still quite easy to find a measure μ such that the given polygonal line is stationary with respect to μ.

6.4 Examples of irregular stationary points

In this section we will show that there exist Ω and Σ stationary in Ω such that Σ has angular points.

From now on, we will consider sets Σ made of two arcs of circumference with a common end point O. We will refer to such sets simply as curved corners. We will say that a curved corner is convex, if it is a convex curve (i.e. it intersects every line in at most two points).

Proposition 6.4.1 *There exists a convex curved corner Σ stationary with respect to some open $\Omega \subset \mathbb{R}^2$.*

Proof. Let $\lambda > 0$ be fixed. Our construction is that shown on Figure 6.3. Namely, the set Σ is made by two arcs $\widetilde{QO}$ and $\widetilde{PO}$ of circumferences with the same radius R and with centers C_1 and C_2 respectively. The points P and Q are chosen in such a way that both belong to the line v containing the centers of the circumferences. We denote by $2\varphi \in [0, \pi]$ the angle between the normals in O to the respective arcs, pointing away from v. Then $\alpha = \pi/2 - \varphi$ is the angle between v and the ray C_1O (and also, by symmetry, between v and the ray C_2O). We also assume the unit coordinate vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ to be directed as in Figure 6.3.

Now let

$$\begin{aligned} b &:= \sqrt{R^2 + 2\lambda} - R, \\ f(\theta) &:= \sqrt{2R^2 + 2\lambda - \left(\frac{R\cos\alpha}{\cos\theta}\right)^2}, \quad \theta \in [0, \alpha], \\ r &:= \sqrt{2\lambda}. \end{aligned}$$

Notice that $r > b$. Moreover, fix a $k \in (0, R(1 - \cos\alpha))$ and an $h > 0$ such that

$$-\int_{-k}^{k} \left(\int_{-h}^{0} y(z^2 + y^2)^{-1/2} dy \right) dz = \lambda. \tag{6.4.1}$$

Consider now the region bounded by Σ and the segment $\overline{PQ}$. It is divided symmetrically in two regions A and B by the line u passing through O perpendicular to v. Let C indicate the region identified by the arc $\widetilde{QO}$, the ray C_1O, the ray C_1Q and the curve defined by the equation $\rho = f(\theta)$ in polar coordinates with center C_1 and the angle θ counted counterclockwise increasing from 0 to α. Define D to be the region symmetric to C with respect to u. Let E and G be equal rectangles with an edge on v of length k, centered in P and Q respectively, with another edge of length h, and belonging to the half space bounded by v and not containing O. Finally, let F stand for the circular sector with center O and with the radius r bounded by the normals to $\widetilde{QO}$ and $\widetilde{PO}$ as in Figure 6.3.

Define now $\Omega := A \cup B \cup C \cup D \cup E \cup F \cup G$. We will show that Σ is optimal with respect to such Ω. Let ν be the outward normal to $\widetilde{QO}$. Points in B and C are projected on Σ to the arc $\widetilde{QO}$, and since $f(\alpha) = R + b$ and $f(\theta) > R + b$ for $\theta \in [0, \alpha)$, we have

$$\begin{aligned}\int_C \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx &= -\int_0^\alpha \int_R^{f(\theta)} \langle X(\theta), \nu(\theta) \rangle \rho d\rho d\theta \\ &= -\int_0^\alpha \int_{R+b}^{f(\theta)} \langle X(\theta), \nu(\theta) \rangle \rho d\rho d\theta + \\ &\quad - \int_0^\alpha \int_R^{R+b} \langle X(\theta), \nu(\theta) \rangle \rho d\rho d\theta,\end{aligned}$$

but, by the definition of b and f,

$$\begin{aligned}\int_0^\alpha \int_R^{R+b} \langle X(\theta), \nu(\theta) \rangle \rho d\rho d\theta &= \left(\frac{1}{2}(R + b^2) - \frac{1}{2}R^2 \right) \int_0^\alpha \langle X(\theta), \nu(\theta) \rangle d\theta \\ &= \lambda \int_0^\alpha \langle X(\theta), \nu(\theta) \rangle d\theta,\end{aligned}$$

$$\int_{R+b}^{f(\theta)} \rho d\rho = \frac{1}{2}(f(\theta))^2 - \frac{1}{2}(R + b)^2 = \frac{1}{2}R^2 - \frac{1}{2}\left(\frac{R \cos \alpha}{\cos \theta} \right)^2.$$

For the computation of the integral in the region B, it is easily seen that

$$B = \left\{ (\rho, \theta) \colon 0 \le \theta \le \alpha \ , \ \frac{R \cos \alpha}{\cos \theta} \le \rho \le R \right\}, \tag{6.4.2}$$

so it follows that

$$\begin{aligned}\int_B \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx &= \int_0^\alpha \langle X(\theta), \nu(\theta) \rangle \int_{\frac{R \cos \alpha}{\cos \theta}}^R \rho d\rho d\theta \\ &= \frac{1}{2} \int_0^\alpha \langle X(\theta), \nu(\theta) \rangle \left(R^2 - \left(\frac{R \cos \alpha}{\cos \theta} \right)^2 \right) d\theta.\end{aligned}$$

Hence one obtains

$$\int_{B \cup C} \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx = -\lambda \int_0^\alpha \langle X(\theta), \nu(\theta) \rangle d\theta. \tag{6.4.3}$$

Now consider the curvature term of the Euler equation. Let $H_\Sigma(\widetilde{QO})$ indicate the nonatomic part of the curvature of the arc $\widetilde{QO}$, i.e. the part not involving the contribution of endpoints. The term $\langle H_\Sigma(\widetilde{QO}), X \rangle$ is clearly equal to

$$-\int_0^\alpha \langle X(\theta), \nu(\theta) \rangle d\theta.$$

We end up with

$$\int_{B\cup C}\left\langle X(\pi^\Sigma(x)),\frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|}\right\rangle dx-\lambda\langle H_\Sigma(\widetilde{QO}),X\rangle=0 \tag{6.4.4}$$

By symmetry, the integral over region $A\cup D$ can be computed in polar coordinates with respect to C_2 and v, with angle θ' counted clockwise increasing from 0 to α, and has exactly the same form. Reasoning in the same way, one sees the analogy between the terms $\langle H_\Sigma(\widetilde{PO}),X\rangle$ and $\langle H_\Sigma(\widetilde{QO}),X\rangle$. It follows that

$$\int_{A\cup D}\left\langle X(\pi^\Sigma(x)),\frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|}\right\rangle dx-\lambda\langle H_\Sigma(\widetilde{PO}),X\rangle=0. \tag{6.4.5}$$

Let us now compute the integrals over E and G. These two regions are disjoint thanks to the choice of k. By (6.4.1) we get

$$\begin{aligned}\int_E&\left\langle X(\pi^\Sigma(x)),\frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|}\right\rangle dx\\&=-X_2(P)\int_{-k}^{k}\left(\int_{-h}^{0}y(z^2+y^2)^{-1/2}dy\right)dz\\&=\lambda X_2(P).\end{aligned} \tag{6.4.6}$$

Analogously the integral over G is given by

$$\int_G\left\langle X(\pi^\Sigma(x)),\frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|}\right\rangle dx=\lambda X_2(Q). \tag{6.4.7}$$

For the integral over F, we consider polar coordinates referred to the center O with the angle θ measured counterclockwise starting from the direction parallel to the ray C_1Q, so that

$$\frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|}=-(\cos\theta,\sin\theta),\quad x\in F.$$

Then, since in F the minimum distance from Σ is always attained in the point O, we get

$$\begin{aligned}\int_F&\left\langle X(\pi^\Sigma(x)),\frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|}\right\rangle dx\\&=-\int_{\frac{\pi}{2}-\varphi}^{\frac{\pi}{2}+\varphi}\int_0^r\langle X(O),(\cos\theta,\sin\theta)\rangle\rho d\rho d\theta\\&=-X_2(O)\int_{\frac{\pi}{2}-\varphi}^{\frac{\pi}{2}+\varphi}\int_0^r\sin\theta\rho d\rho d\theta\\&=-X_2(O)r^2\sin\varphi=-2\lambda X_2(O)\sin\varphi\end{aligned} \tag{6.4.8}$$

Finally consider the curvature terms at the endpoints P and Q. We have respectively

$$\langle H_\Sigma(P),X\rangle=X_2(P),\quad\langle H_\Sigma(Q),X\rangle=X_2(Q). \tag{6.4.9}$$

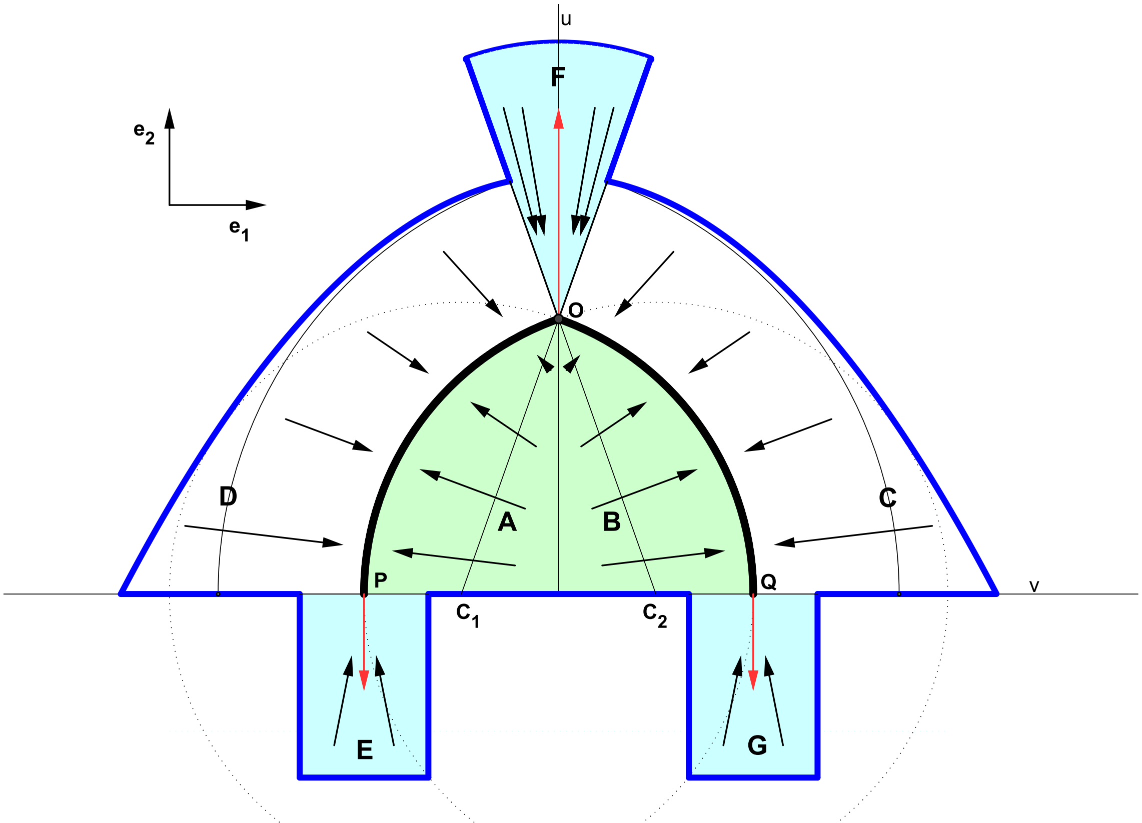

Figure 6.3: Construction of the proof of Proposition 6.4.1

For the point O, we have

$$H_\Sigma(O) = -2\cos\alpha\,\delta_O \mathbf{e}_2,$$

yielding

$$\langle H_\Sigma(O), X\rangle = -2\sin\varphi X_2(O). \tag{6.4.10}$$

Since $\Omega = A \cup B \cup C \cup D \cup E \cup F \cup G$ and

$$\langle H_\Sigma, X\rangle = \langle H_\Sigma(\widetilde{QO}), X\rangle + \langle H_\Sigma(\widetilde{PO}), X\rangle + \langle H_\Sigma(P), X\rangle + \langle H_\Sigma(O), X\rangle + \langle H_\Sigma(Q), X\rangle,$$

combining (6.4.4), (6.4.5), (6.4.6), (6.4.7), (6.4.8), (6.4.9) and (6.4.10) we see that the Euler equation (6.2.3) is identically satisfied. □

Next we will show that, for a convex domain Ω, if the amplitude of the corner is not too large, then a set composed of two arcs of circle is not stationary.

We first introduce the notation similar to that used in the proof of Proposition 6.4.1, but for a generic curved corner Σ made by two arcs $\widetilde{QO}$ and $\widetilde{PO}$ of circumferences with different radii R_1 and R_2 and with centers C_1 and C_2 respectively. Again $2\varphi \in [0, \pi]$ is the angle between the normals in O which bound the set of points (we will call the bisector ray of the latter angle u) in $\mathbb{R}^2$ having O as the unique point of minimum distance to Σ. Let v'

be a ray starting at C_1 forming the angle $\alpha \leq \pi/2 - \varphi$ with the ray C_1O. We assume that α is sufficiently small so that v' meets $\widetilde{QO}$ in some point M. In this way the rays v', C_1O and the arc $\widetilde{QO}$ form a sector of area $\alpha R_1^2/2$. We also assume the unit coordinate vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ to be directed as in Figures 6.4 and 6.5.

Fix α small enough such that v' meets the continuation of u. Consider the ridge set $\mathfrak{R}_\Sigma$ of Σ (i.e. the set of points of equal distance from the two arcs). Note that there exists a segment $\overline{OZ}$ with $Z \in v'$ which intersects $\mathfrak{R}_\Sigma$ only in O (this assertion is implied by the fact that $\mathfrak{R}_\Sigma$ is a regular curve tangent in O to u). Denote by W the curvilinear triangle bounded by $\widetilde{MO}$, $\overline{ZO}$ and $\overline{ZM}$. Clearly it contains the set of points T having the projection to Σ on $\widetilde{MO}$. Moreover, they are all projected to $\widetilde{MO}$ from the same side (i.e. either from outside of the circle $B_{R_1}(C_1)$ as in Figure 6.4, or from the inner part of the circle $B_{R_1}(C_1)$ as in Figure 6.5). It is important to observe that there are no points with such a property outside of W. We denote by C the set of points having the projection to Σ on $\widetilde{MO}$ but from the different side with respect to T.

In this section we will consider a vector field X supported in a small neighborhood of a subset of $\widetilde{MO}$ (in polar coordinates with respect to C_1 and v', the points of the support are contained in the set with angular coordinate $\theta \in [\theta_0, \alpha]$). We assume that X be vanishing in O and have restriction to $\widetilde{MO}$ directed towards the outward normal ν to the circle $B_{R_1}(C_1)$. Thus in the first member of (6.2.3) the only nonzero terms are the integrals in the regions T and C and the curvature term restricted to $\widetilde{MO}$.

Proposition 6.4.2 *A non convex curved corner is not stationary, for any $\Omega \subset \mathbb{R}^2$.*

Proof. Let Σ be a non convex curved corner. In this case one of the centers belongs to one of the rays bounding the cone of points for which the projection to Σ coincides with O (let it be C_1). So the region C is inside the sector bounded by the arc $\widetilde{QO}$ (see Figure 6.4). If β is the angle formed by $\overline{OZ}$ and $\overline{OC_1}$, it is easily seen that one can choose the point Z so that $\beta \in (\pi/2, \pi)$. In polar coordinates with respect to C_1 and v' for small α one has then

$$W = \left\{(\rho, \theta) \colon 0 < \theta < \alpha, R_1 < \rho < \frac{R_1 \sin\beta}{\sin(\beta + \alpha - \theta)}\right\} \tag{6.4.11}$$

(observe that $\sin\beta / \sin(\beta + \alpha - \theta) > 1$ since $\beta \in (\pi/2, \pi)$, and $\alpha - \theta > 0$ is small enough). We obtain also

$$-\lambda\langle H_\Sigma, X\rangle = \lambda \int_{\theta_0}^{\alpha} \langle X(\theta), \nu(\theta)\rangle d\theta = \lambda \int_{\theta_0}^{\alpha} |X(\theta)| d\theta.$$

Moreover, thanks to (6.4.11), we have

$$\begin{aligned}\left|\int_T \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|}\right\rangle dx\right| &\leq \int_W \left|\left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|}\right\rangle\right| dx \\ &= \frac{1}{2}\int_{\theta_0}^{\alpha} |X(\theta)| \left(\frac{R_1^2 \sin^2\beta}{\sin^2(\beta + \alpha - \theta)} - R_1^2\right) d\theta.\end{aligned}$$

But for $\theta \to \alpha$, with β fixed, we get

$$\frac{R_1^2 \sin^2 \beta}{\sin^2(\beta + \alpha - \theta)} - R_1^2 = o(1),$$

implying that

$$\left| \int_T \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx \right| \le o\left(\int_{\theta_0}^{\alpha} |X(\theta)|\, d\theta \right).$$

Therefore it is clear that, for θ_0 close enough to α, the Euler equation (6.2.3) is never satisfied for Σ. In fact, since

$$\int_C \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle \ge 0,$$

we have that

$$\int_\Omega \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle - \lambda \langle H_\Sigma, X \rangle \ge \int_T \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle - \lambda \langle H_\Sigma, X \rangle.$$

Hence, for $\theta_0 \to \alpha$ one has

$$\int_\Omega \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle - \lambda \langle H_\Sigma, X \rangle \ge \lambda \int_{\theta_0}^{\alpha} |X(\theta)|\, d\theta - o\left(\int_{\theta_0}^{\alpha} |X(\theta)|\, d\theta \right),$$

that is, the right hand side of the above inequality is always strictly positive once θ_0 is sufficiently close to α. □

Finally, we show that the condition for a curved corner to be stationary with respect to a *convex* Ω is even more restrictive.

Proposition 6.4.3 *Let Ω be convex. A curved corner is not stationary with respect to Ω if*

$$\frac{4\lambda}{b_1^2 + b_2^2} \ge h(\varphi), \tag{6.4.12}$$

where $b_i := \sqrt{R_i^2 + 2\lambda} - R_i$, $i = 1, 2$,

$$h(\varphi) := \frac{1}{\sin \varphi} \int_0^{\varphi} \frac{\cos(\varphi - \theta)}{\cos^2 \theta} d\theta.$$

In particular, there are no curved corners of amplitude less than or equal to 2γ, where $\gamma \in (0, \pi/2)$ is the angle that satisfies

$$\int_0^{\gamma} \frac{\cos(\gamma - \theta)}{\cos^2 \theta} d\theta = \sin \gamma,$$

so $\gamma \simeq 54°$.

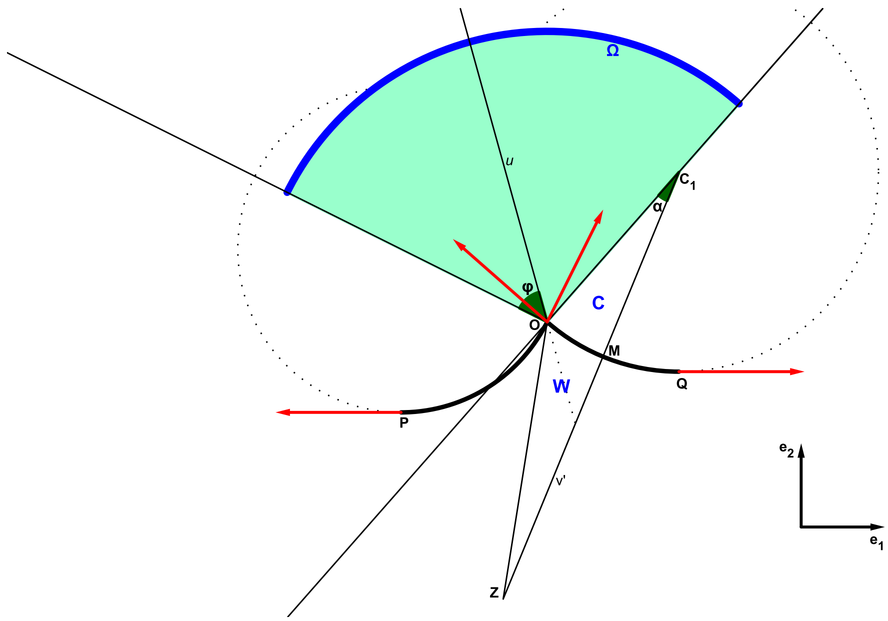

Figure 6.4: Construction of the proof of Proposition 6.4.2

Proof. If the curved corner is not convex, we refer to the previous Proposition 6.4.2. Otherwise, let β be the angle between $\overline{OZ}$ and $\overline{OC_1}$. This time $\beta < \varphi$, so that once α is sufficiently small, one has $\beta + \alpha < \pi/2$ (see Figure 6.5). In polar coordinates with respect to C_1 and v', we have

$$W = \left\{ (\rho, \theta) \colon 0 < \theta < \alpha, \frac{R_1 \sin\beta}{\sin(\beta + \alpha - \theta)} < \rho < R_1 \right\}. \tag{6.4.13}$$

Notice that $\alpha - \theta > 0$ is small and the bound on β gives $\sin\beta / \sin(\beta + \alpha - \theta) < 1$. The curvature term in the Euler equation (6.2.3) is given by

$$-\lambda \langle H_\Sigma, X \rangle = \lambda \int_{\theta_0}^{\alpha} \langle X(\theta), \nu(\theta) \rangle d\theta = \lambda \int_{\theta_0}^{\alpha} |X(\theta)| \, d\theta. \tag{6.4.14}$$

Using (6.4.13), we get

$$\begin{aligned} \int_T \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx &\le \int_W \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx \\ &= \int_{\theta_0}^{\alpha} \frac{|X(\theta)|}{2} \left(R_1^2 - \frac{R_1^2 \sin^2\beta}{\sin^2(\beta + \alpha - \theta)} \right) d\theta, \end{aligned}$$

and again for $\theta \to \alpha$, with β fixed, we have

$$R_1^2 - \frac{R_1^2 \sin^2\beta}{\sin^2(\beta + \alpha - \theta)} = o(1),$$

and hence

$$\int_T \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx \le o\left(\int_{\theta_0}^{\alpha} |X(\theta)|d\theta\right). \tag{6.4.15}$$

Notice that C contains a region formed by v', the ray C_1O, the arc $\widetilde{MO}$ and some arc concentric to $\widetilde{MO}$ but of bigger radius. We express the subset of the boundary of Ω bounding C in polar coordinates (ρ,θ) with respect to C_1 and v' by the equation $\rho = b_1 + R_1 + g_1(\theta)$, where $g_1(\theta) \to 0$ as $\theta \to \alpha$, and b_1 is the distance between O and the intersection between $\partial\Omega$ and the ray C_1O, which we denote by S. Then

$$\begin{aligned}\int_C \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx &= -\int_{\theta_0}^{\alpha}\int_{R_1}^{R_1+b_1+g_1(\theta)} |X(\theta)|\rho d\rho d\theta \\ &= -\frac{1}{2}\int_{\theta_0}^{\alpha} |X(\theta)|\left(2R_1b_1 + b_1^2 + g_1(\theta)^2 + 2(R_1+b_1)g_1(\theta)\right)d\theta \\ &= -\frac{1}{2}(2R_1b_1+b_1^2)\int_{\theta_0}^{\alpha}|X(\theta)|d\theta \\ &\quad -\frac{1}{2}\int_{\theta_0}^{\alpha}|X(\theta)|\left(g_1(\theta)^2 + 2(R_1+b_1)g_1(\theta)\right)d\theta.\end{aligned} \tag{6.4.16}$$

Suppose now that the Euler equation (6.2.3) holds, that is,

$$\begin{aligned}\int_T \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx + \int_C \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx \\ = \lambda\langle H_\Sigma, X\rangle.\end{aligned} \tag{6.4.17}$$

Combining (6.4.14), (6.4.15) and (6.4.16) in the above relationship, by comparison of the first order terms with respect to $\int_{\theta_0}^{\alpha}|X(\theta)|\,d\theta$ as $\theta_0 \to \alpha$, we obtain that

$$b_1 = \sqrt{R_1^2 + 2\lambda} - R_1. \tag{6.4.18}$$

Moreover, by this choice of b_1 we have

$$-\frac{1}{2}(2R_1b_1+b_1^2)\int_{\theta_0}^{\alpha}|X(\theta)|d\theta = \lambda\langle H_\Sigma, X\rangle.$$

From (6.4.17) and (6.4.16), we conclude that

$$\begin{aligned}\int_T \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx \\ -\frac{1}{2}\int_{\theta_0}^{\alpha}|X(\theta)|\left(g_1(\theta)^2 + 2(R_1+b_1)g_1(\theta)\right)d\theta = 0;\end{aligned}$$

but since

$$\int_T \left\langle X(\pi^\Sigma(x)), \frac{\pi^\Sigma(x)-x}{|\pi^\Sigma(x)-x|} \right\rangle dx > 0,$$

it follows

$$\int_{\theta_0}^{\alpha} |X(\theta)| \left(g_1(\theta)^2 + 2(R_1 + b_1)g_1(\theta)\right) d\theta > 0.$$

Minding that X has an arbitrary support in $[\theta_0, \alpha]$, this means that

$$g_1(\theta)^2 + 2(R_1 + b_1)g_1(\theta) > 0$$

and implies $g_1(\theta) > 0$ for all $\theta \in [\theta_0, \alpha]$ whenever α is small enough (since otherwise $g_1(\theta) < -R_1 - b_1$ which contradicts the fact that g should be vanishing as $\theta \to \alpha$). Hence, the part of $\partial\Omega$ corresponding to the angular coordinate $\theta \in [\theta_0, \alpha]$ is, for small α, more distant from C_1 than the arc σ of the circumference with center C_1 passing through S, thus satisfying the equation $\rho(\theta) = R_1 + b_1$. Thanks to convexity of Ω we have then that any ray starting in S, directed inside the cone of points with projection to Σ in O, and belonging to a support line to $\partial\Omega$ in S, forms an angle not greater than $\pi/2$ with the segment $\overline{SO}$ (mind that the angle of $\pi/2$ corresponds to the case when the ray is tangent to σ). As a consequence, the part of Ω which lies in the angle (of value φ) bounded by u and the ray OS, is contained in the triangle V_1, formed by u, $\overline{OS}$ and the tangent in S to σ.

Now fix a new vector field $\hat{X}$, compactly supported in a small neighborhood of O and such that $\hat{X}(O)$ is directed along u. One has

$$H_\Sigma(O) = \delta_O(\tau_Q + \tau_P),$$

where τ_Q and τ_P are the unit vectors tangent in O to the arcs $\widetilde{PO}$ and $\widetilde{QO}$ respectively and directed towards P and Q respectively. Since

$$\langle \hat{X}, \delta_O \tau_Q \rangle = \langle \hat{X}, \delta_O \tau_P \rangle = -|\hat{X}(O)| \sin\varphi,$$

we get

$$-\lambda \langle \hat{X}, H_\Sigma(O) \rangle = 2\lambda |\hat{X}(O)| \sin\varphi. \tag{6.4.19}$$

Now compute the contribution given by triangle V_1 to the first term of the Euler equation (6.2.3). For this purpose we use polar coordinates with respect to O and the ray OS, with $\theta \in [0, \varphi]$. It is clear that

$$V = \left\{ (\rho, \theta) \colon 0 \le \theta \le \varphi, 0 < \rho \le \frac{b}{\cos\theta} \right\}. \tag{6.4.20}$$

Therefore,

$$\begin{aligned}
\int_{V_1} \left\langle \hat{X}(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx &= -|\hat{X}(O)| \int_0^\varphi \cos(\varphi - \theta) \int_0^{\frac{b_1}{\cos\theta}} \rho d\rho d\theta \\
&= -\frac{1}{2} b_1^2 |\hat{X}(O)| \int_0^\varphi \frac{\cos(\varphi - \theta)}{\cos^2\theta} d\theta \\
&= -\frac{1}{2} b_1^2 |\hat{X}(O)| h(\varphi) \sin\varphi.
\end{aligned} \tag{6.4.21}$$

Reasoning in the same way with arc $\widetilde{PO}$ instead of the arc $\widetilde{PQ}$, we obtain the analogous triangle V_2 with

$$\int_{V_2} \left\langle \hat{X}(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx = -\frac{1}{2} b_2^2 |\hat{X}(O)| h(\varphi) \sin\varphi, \tag{6.4.22}$$

with $b_2 := \sqrt{R_2^2 + 2\lambda} - R_2$. But since Ω is convex, and one of the sides of V_1 (resp. V_2) is in the support line to Ω, we have

$$(\pi^\Sigma)^{-1}(O) \cap \Omega \subset V_1 \cup V_2. \tag{6.4.23}$$

Let us write the Euler equation (6.2.3) with respect to the vector field $\hat{X}$. Letting $\Gamma := (\Sigma \backslash \{O\}) \cap \operatorname{supp} \hat{X}$, thanks to (6.4.19) we get

$$\int_{(\pi^\Sigma)^{-1}(O)} \left\langle \hat{X}(\pi^\Sigma(x)), \frac{\pi^\Sigma(x) - x}{|\pi^\Sigma(x) - x|} \right\rangle dx + 2\lambda |\hat{X}(O)| \sin\varphi + c_\Gamma = 0, \tag{6.4.24}$$

where by c_Γ we denoted the sum of all the terms in the Euler equation which involve integrals over Γ. Minding the strict inclusion (6.4.23), and using (6.4.21) and (6.4.22), we obtain

$$-\frac{1}{2}(b_1^2 + b_2^2)|\hat{X}(O)| h(\varphi) \sin\varphi + 2\lambda |\hat{X}(O)| \sin\varphi + c_\Gamma < 0. \tag{6.4.25}$$

Since c_Γ contains only integral terms, we have that c_Γ can be made arbitrarily small by choosing a sufficiently small support of $\hat{X}$, and hence (6.4.25) may be satisfied, only if

$$\frac{4\lambda}{b_1^2 + b_2^2} < h(\varphi), \tag{6.4.26}$$

or, in other words, when $h(\varphi)$ is as in the statement being proven, then the Euler equation is not satisfied. Finally, to prove the second claim, it remains to observe that $4\lambda/(b_1^2 + b_2^2) > 1$, and hence with $h(\varphi) \leq 1$ the respective curved corner is not stationary. □

6.5 Examples of $C^{1,1}$ minimizers

In this section we mention an interesting result obtained in [69]: explicit examples of minimizers of the average distance functional are exhibited therein, for the related constrained problem (that is, with length constraint, rather than a penalization term).

Let Σ be a $C^{1,1}$ simple curve (homeomorphic image of $[0,1]$), and let Ω_ε be its ε-neighborhood. For ε small enough it is proven in [69] that, under suitable bounds on the curvature and the length of Σ (see Theorem 1.1 therein), the area of Ω_ε is given by $2\varepsilon\mathcal{H}^1(\Sigma) + \pi\varepsilon^2$. Making use of this fact one can prove the following.

Theorem 6.5.1 *Let Σ and Ω_ε be chosen as above. Let $\mathcal{H}^1(\Sigma) = \ell$. If ε is small enough, then Σ is a minimizer for*

$$\int_{\Omega_\varepsilon} \operatorname{dist}(x, \Sigma)\, d\mathcal{L}^2(x)$$

among all closed connected subsets of Ω_ε with $\mathcal{H}^1$ measure equal to ℓ.

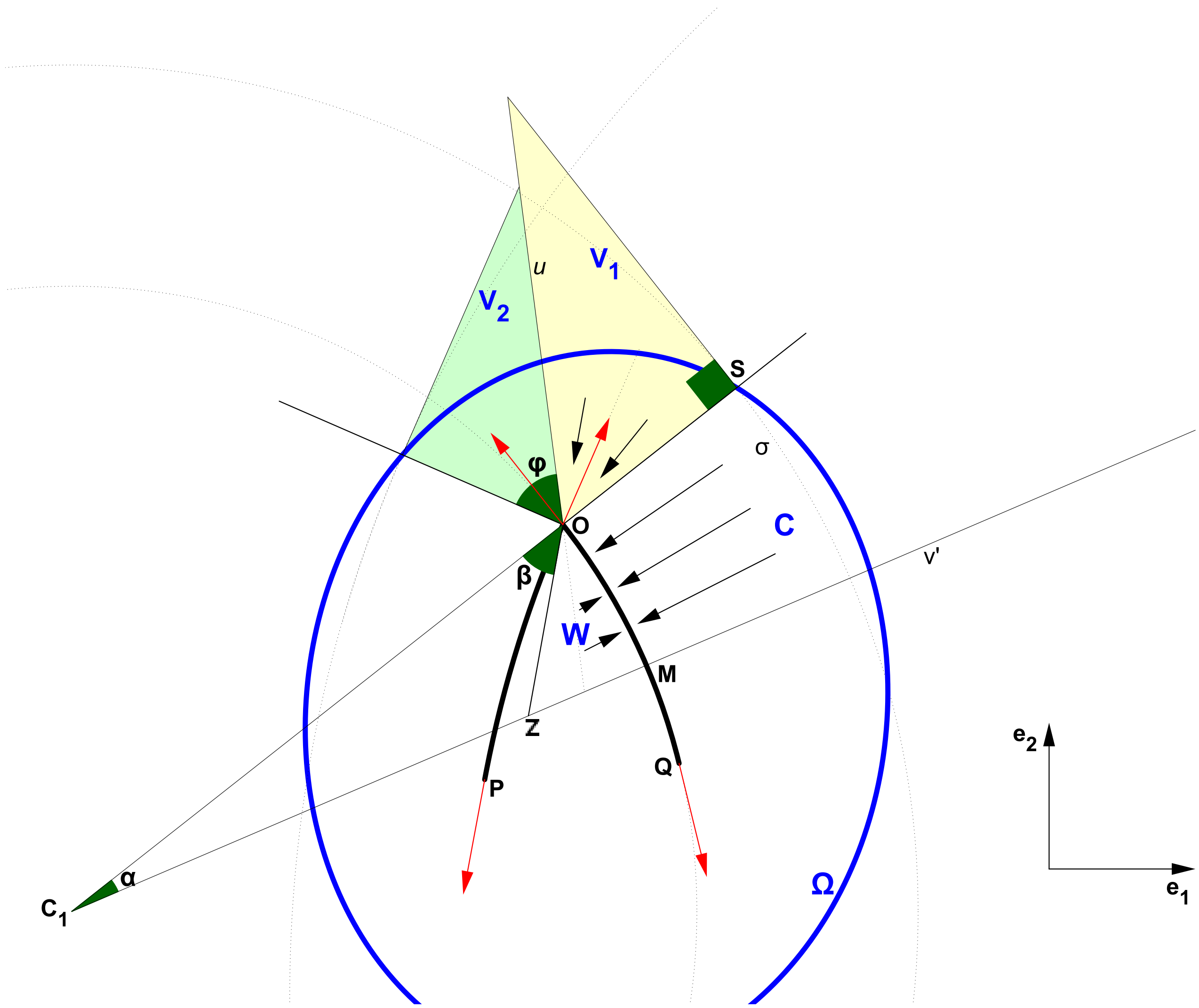

Figure 6.5: Construction of the proof of Proposition 6.4.3

We give a brief sketch of the proof. It is shown in [56] that for any $\Omega \subset \mathbb{R}^2$ and any compact connected $\Gamma \subset \mathbb{R}^2$ with $\mathcal{H}^1(\Gamma) = \ell$ there holds

$$\mathcal{L}^2(\{x \in \Omega : \mathrm{dist}\,(x, \Gamma) \le s\}) \le \min\{\mathcal{L}^2(\Omega), 2s\ell + \pi s^2\} \quad \forall s > 0. \tag{6.5.1}$$

Let Ω and Γ be as above, with $\Gamma \subset \Omega$. By the slicing formula we have

$$\begin{aligned} \int_\Omega \mathrm{dist}\,(x, \Gamma)\, d\mathcal{L}^2(x) &= \int_0^{+\infty} \mathcal{L}^2(\{x \in \Omega : \mathrm{dist}\,(x, \Gamma) > s\})\, ds \\ &= \int_0^{+\infty} \big(\mathcal{L}^2(\Omega) - \mathcal{L}^2(\{x \in \Omega : \mathrm{dist}\,(x, \Gamma) \le s\})\big)\, ds. \end{aligned} \tag{6.5.2}$$

Now let $\varepsilon > 0$. Choosing $\Omega = \Omega_\varepsilon$ and applying (6.5.1) we get

$$\int_{\Omega_\varepsilon} \mathrm{dist}\,(x, \Gamma)\, d\mathcal{L}^2(x) \ge \int_0^{+\infty} \big(\mathcal{L}^2(\Omega_\varepsilon) - \min\{\mathcal{L}^2(\Omega_\varepsilon), 2s\ell + \pi s^2\}\big) \tag{6.5.3}$$

for any compact connected subset Γ of Ω_ε satisfying $\mathcal{H}^1(\Gamma) = \ell$. The crucial point here is that, as previously remarked, (6.5.1) is an equality, for any $s \le \varepsilon$, if Ω is the ε-neighborhood of Γ (for small enough ε). Hence the set Σ realizes equality in (6.5.3) and the thesis follows.

6.6 Euler equation for the compliance functional

In this final section we derive another first order equation, for a related problem. We consider the case of a functional arising from the theory of elliptic equations:

$$\mathcal{F}(\Sigma) := \int_\Omega u_\Sigma(x) f(x)\, dx + \lambda \mathcal{H}^1(\Sigma).$$

Here $\Omega \subset \mathbb{R}^2$ is a given bounded open subset, f is a given function, and u_Σ is the unique solution of the PDE

$$\begin{cases} -\Delta u = f & \text{in } \Omega \setminus \Sigma, \\ u = 0 & \text{on } \partial\Omega \cup \Sigma. \end{cases}$$

An integration by parts in the PDE above gives that the *compliance* term $\int_\Omega u_\Sigma f\, dx$ appearing in the functional $\mathcal{F}$ can be expressed in an equivalent way:

$$\int_\Omega u_\Sigma(x) f(x)\, dx = \max\Big\{ \int_\Omega \big(2f(x)u - |\nabla u|^2\big)\, dx \ : \ u \in W_0^{1,2}(\Omega \setminus \Sigma) \Big\}.$$

For simplicity we assume that Ω has a Lipschitz boundary and that $f \in W^{1,2}(\mathbb{R}^2)$. In fact, we could also consider the case of a p-Laplace operator, and the similarity with the average distance functional consists in the fact (shown in [22]) that as $p \to +\infty$ the p-compliance problem converges to the one with the average distance functional. Here we limit ourselves to the case $p = 2$. Also for simplicity we have taken the Dirichlet condition $u = 0$ on $\partial\Omega$; all the arguments can be repeated for the Neumann case $\frac{\partial u}{\partial \nu} = 0$ on $\partial\Omega$.

The existence of a solution to the minimum problem

$$\min\left\{\mathcal{F}(\Sigma)\ :\ \Sigma \text{ closed connected subset of } \Omega\right\}$$

follows by an application of the Šverák compactness theorem (see [22]). Here we are interested, as before, in the first order necessary conditions of optimality.

Following [42, theorem 5.3.2], if ϕ_ε is a one parameter group of diffeomorphisms satisfying (6.2.1), setting $\Sigma_\varepsilon := \phi_\varepsilon(\Sigma)$, $u := u_\Sigma$ and $u_\varepsilon = u_{\Sigma_\varepsilon}$, we have as $\varepsilon \to 0$ that $\frac{u_\varepsilon - u}{\varepsilon} \to u'$ in $L^2(\Omega)$, where u' satisfies the PDE

$$\begin{cases} -\Delta u' = 0 \text{ in } \Omega \setminus \Sigma, \\ u' = 0 \text{ on } \partial\Omega,\ u' = -\nabla u \cdot X \text{ on } \Sigma. \end{cases}$$

Note that the boundary conditions in the above equation are understood in the weak sense, i.e. $u' + \nabla u \cdot X \in W_0^{1,2}(\mathbb{R}^2)$. Therefore, the first variation argument applied to the functional $\mathcal{F}$ gives

$$\frac{\partial}{\partial\varepsilon}\mathcal{F}(\Sigma_\varepsilon)\Big|_{\varepsilon=0} = \int_\Omega u' f\, dx - \lambda\langle H_\Sigma, X\rangle.$$

Suppose now that $\Omega = \Omega^+ \cup \Omega^-$ with $\Sigma \subset \partial\Omega^+ \cap \partial\Omega^-$. Then, if Σ, $\partial\Omega$ and f provide sufficient regularity for u and u' so that the Green formula can be applied, we have

$$\begin{aligned}\int_{\Omega^+} u' f\, dx = -\int_{\Omega^+} u'\Delta u\, dx &= \int_{\Omega^+} \nabla u'\nabla u\, dx - \int_{\partial\Omega^+} u'\frac{\partial u}{\partial n}\, d\mathcal{H}^1 \\ &= \int_{\Omega^+} \nabla u'\nabla u\, dx + \int_\Sigma \nabla u\cdot X\frac{\partial u}{\partial n}\, d\mathcal{H}^1 \\ &\qquad - \int_{\partial\Omega^+\setminus(\partial\Omega\cup\Sigma)} u'\frac{\partial u}{\partial n}\, d\mathcal{H}^1,\end{aligned}$$

where n stands for the external normal to Ω^+. But

$$\begin{aligned}\int_{\Omega^+} \nabla u'\nabla u\, dx &= -\int_{\Omega^+} u\Delta u'\, dx + \int_{\partial\Omega^+} u'\frac{\partial u}{\partial n}\, d\mathcal{H}^1 \\ &= -\int_{\partial\Omega^+\setminus(\partial\Omega\cup\Sigma)} u\frac{\partial u'}{\partial n}\, d\mathcal{H}^1.\end{aligned}$$

Thus,

$$\int_{\Omega^+} u' f\, dx = \int_\Sigma \nabla u^+\cdot X\frac{\partial u^+}{\partial n}\, d\mathcal{H}^1 - \int_{\partial\Omega^+\setminus(\partial\Omega\cup\Sigma)} \left(u\frac{\partial u'}{\partial n} + u'\frac{\partial u}{\partial n}\right)\, d\mathcal{H}^1, \qquad (6.6.1)$$

where ∇u^+ stands for the trace on Σ of the gradient of u restricted to Ω^+, and $\frac{\partial u^+}{\partial n}$ stands for the trace of the respective normal derivative. Analogously, minding that the external normal to Ω^- over $\partial\Omega^+ \cap \partial\Omega^-$ is given by $-n$, we get

$$\int_{\Omega^-} u' f\, dx = -\int_\Sigma \nabla u^-\cdot X\frac{\partial u^-}{\partial n}\, d\mathcal{H}^1 + \int_{\partial\Omega^+\setminus(\partial\Omega\cup\Sigma)} \left(u\frac{\partial u'}{\partial n} + u'\frac{\partial u}{\partial n}\right)\, d\mathcal{H}^1, \qquad (6.6.2)$$

where ∇u^- stands for the trace on Σ of the gradient of u restricted to Ω^-, and $\frac{\partial u^-}{\partial n}$ stands for the trace of the respective normal derivative. From (6.6.1) and (6.6.2) we obtain

$$\int_\Omega u' f \, dx = \int_\Sigma \nabla u^+ \cdot X \frac{\partial u^+}{\partial n} \, d\mathcal{H}^1 - \int_\Sigma \nabla u^- \cdot X \frac{\partial u^-}{\partial n} \, d\mathcal{H}^1.$$

Recalling that

$$\nabla u^\pm = \frac{\partial u^\pm}{\partial n} n,$$

since the tangential derivatives of $u^\pm$ over Σ vanish (because $u^\pm = u = 0$ on Σ), we get

$$\int_\Omega u' f \, dx = \int_\Sigma \left(\left(\frac{\partial u^+}{\partial n} \right)^2 - \left(\frac{\partial u^-}{\partial n} \right)^2 \right) X \cdot n \, d\mathcal{H}^1.$$

Hence,

$$\frac{\partial}{\partial \varepsilon} \mathcal{F}(\Sigma_\varepsilon) \Big|_{\varepsilon=0} = \int_\Sigma \left(\left(\frac{\partial u^+}{\partial n} \right)^2 - \left(\frac{\partial u^-}{\partial n} \right)^2 \right) X \cdot n \, d\mathcal{H}^1 - \lambda \langle H_\Sigma, X \rangle.$$

Since this holds for every vector field X, we deduce the Euler equation that must hold for every minimizer of $\mathcal{F}$:

$$\left(\frac{\partial u^+}{\partial n} \right)^2 - \left(\frac{\partial u^-}{\partial n} \right)^2 = \lambda \langle H_\Sigma, n \rangle.$$

Remark 6.6.1 In the 1-dimensional case, with $\Omega = [0,1]$ and Σ made by a finite number of points (so $H_\Sigma = 0$) in the interior of the interval, the Euler equation says that u_Σ has symmetric left and right tangents at each element of the minimal Σ.

Bibliography

[1] M. Agueh, *Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory*, Adv. Differential Equations **10** (2005), no.3, 309–360.

[2] L. Ambrosio, *Lecture notes on optimal tranport problem*, in Mathematical aspects of evolving interfaces, CIME summer school in Madeira (Pt), P. Colli and J. Rodrigues, eds., vol. 1812, Springer (2003), 1–52.

[3] L. Ambrosio, N. Fusco, D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York (2000).

[4] L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows in metric spaces and in the spaces of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel (2005).

[5] L. Ambrosio, E. Mainini, *Infinite-dimensional porous media equations and optimal transportation*, J. Evol. Equ. **10** (2010), 217–246.

[6] L. Ambrosio, E. Mainini, S. Serfaty, *Gradient flow of the Chapman-Rubinstein-Schatzman model for signed vortices*, preprint.

[7] L. Ambrosio, C. Mantegazza, *Curvature and distance function from a manifold*, J. Geom. Anal. **8** (1998), no.5, 723–748.

[8] L. Ambrosio, S. Serfaty, *A gradient flow approach to an evolution problem arising in superconductivity*, Comm. Pure Appl. Math. **LXI** (2008), no.11, 1495–1539.

[9] L. Ambrosio, G. Savaré, L. Zambotti, *Existence and Stability for Fokker-Planck equations with log-concave reference measure*, Probab. Theory Related Fields **145** (2009), 517–564.

[10] L. Ambrosio, P. Tilli, *Selected Topics on "Analysis in Metric Spaces",* Scuola Normale Superiore, Pisa (2000).

[11] J.-D. Benamou, Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (1995), 375–393.

[12] P. Bénilan, L. Boccardo, T. Gallouët, R. Gariepy, M. Pierre, J.L. Vázquez, *An L^1-theory of existence and uniqueness of solutions of nonlinear elliptic equations,* Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **22** (1995), no.2, 241–273.

[13] V.I. Bogachev, *Gaussian measures,* Mathematical Surveys and Monographs, 62. American Mathematical Society, Providence, RI (1998).

[14] V.I. Bogachev, *Measure Theory* (2 vols.). Springer-Verlag, Berlin (2007).

[15] Y. Brenier, *Décomposition polaire et réarrangement monotone des champs de vecteurs,* C. R. Acad. Sci. Paris, Sér. I Math. **305** (1987), 805–808.

[16] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions,* Comm. Pure Appl. Math. **44** (1991), 375–417.

[17] H.Brezis, *Convergence in $\mathcal{D}'$ and in L^1 under strict convexity.* In: Boundary value problems for partial differential equations and applications, RMA Res. Notes Appl. Math. **29** (1993), 43–52.

[18] G. Bouchitté, G. Buttazzo, I. Fragalà, *Mean curvature of a measure and related variational problems,* Ann. Scuola Norm. Sup. Cl. Sci. (4) **25** (1997), 179–196.

[19] G. Buttazzo, E. Mainini, E. Stepanov, *Stationary configurations for the average distance functional and related problems,* Control Cybernet. **38** (2009), no.4A, 1107–1130.

[20] G. Buttazzo, E. Oudet, E. Stepanov, *Optimal transportation problems with free Dirichlet regions.* In: Variational methods for discontinuous structures, volume 51, pages 41–65. Birkhäuser (2002).

[21] G. Buttazzo, A. Pratelli, S. Solimini, E. Stepanov, *Optimal Urban Networks via Mass Transportation*, Springer-Verlag (2009).

[22] G. Buttazzo, F. Santambrogio, *Asymptotical compliance optimization for connected networks,* Netw. Heterog. Media **2** (2007), no.4, 761–777.

[23] G. Buttazzo, E. Stepanov, *Optimal transportation networks as free Dirichlet regions in the Monge-Kantorovich problem, Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **II** (2003), no.4, 631–678.

[24] G. Buttazzo, E. Stepanov, *Minimization problems for average distance functionals.* In: Calculus of variations: topics from the mathematical heritage of E. De Giorgi, volume 14 of Quad. Mat., pages 48–83. Dept. Math., Seconda Univ. Napoli, Caserta (2004).

[25] J. Carrillo Menéndez, *Entropy solutions for nonlinear degenerate problems,* Arch. Ration. Mech. Anal. **147** (1999), no.4, 269–361.

[26] E. A. Carlen, W. Gangbo, *Constrained steepest descent in the 2-Wasserstein metric,* Ann. of Math. **157** (2003), no.3, 807–846.

[27] E. A. Carlen, W. Gangbo, *Solution of a model Boltzmann equation via steepest descent in the 2-Wasserstin metric,* Arch. Ration. Mech. Anal. **172** (2004), no.1, 21–64.

[28] J. S. Chapman, J. Rubinstein, M. Schatzman, *A mean-field model for superconducting vortices,* Eur. J. Appl. Math. **7** (1996), no. 2, 97-111.

[29] G. Dal Maso, F. Murat, L. Orsina, A. Prignet, *Renormalized solutions of elliptic equations with general measure data, Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **28** (1999), no. 4, 741–808.

[30] G. Da Prato, *An introduction to infinite-dimensional analysis,* Springer (2006).

[31] G. Da Prato, J. Zabczyk, *Second order partial differential equations in Hilbert spaces,* London Mathematical Society Lecture Note Series, **293**, Cambridge U. P. (2002).

[32] E. De Giorgi, *New problems on minimizing movements,* in Boundary Value Problems for PDE and Applications, C. Baiocchi and J.L. Lions, eds., Masson (1993), 81–98.

[33] C. Dellacherie, P.-A. Meyer, *Probabilities and potential*, vol. 29 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam (1978).

[34] L.C. Evans, R.F. Gariepy, *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL (1992).

[35] K.J. Falconer, *The geometry of fractal sets.* Cambridge Tracts in Mathematics, 85. Cambridge University Press, Cambridge (1986).

[36] H. Federer, *Geometric measure theory.* Springer-Verlag New York Inc., New York (1969).

[37] S. Fang, J. Shao, K.-T. Sturm, *Wasserstein space over the Wiener space*, Probab. Theory Relat. Fields, in press.

[38] I. Fragalà, *Tangential Calculus and Variational Integrals with Respect to a Measure.* PhD Thesis (2000).

[39] W. Gangbo, R. J. McCann, *The geometry of optimal transportation*, Acta Math. **177** (1996), no.2, 113–161.

[40] N. Gigli, *On the geometry of the space of probability measures in* $\mathbb{R}^n$ *endowed with the quadratic optimal transportation distance.* PhD Thesis (2004).

[41] N. Gigli, *Second order analysis in* $(\mathscr{P}_2(M), W_2)$*,* Memoirs of the American Mathematical Society, in press.

[42] A. Henrot, M. Pierre, *Variation et optimisation de formes: une analyse géométrique*, volume 48 of Mathématiques et Applications. Springer-Verlag, Berlin (2005).

[43] R. Jordan, D. Kinderlehrer, F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), 1–17.

[44] L. V. Kantorovich, *On the transfer of masses.* Dokl. Akad. Nauk. SSSR **37** (1942), 227-229.

[45] L. V. Kantorovich, *On a problem of Monge.* Uspekhi Mat.Nauk. **3** (1948), 225-226.

[46] L. LeCam, *Convergence in distribution of stochastic processes,* Univ. Calif. Publ. Statist., **2** (1957), 207–236.

[47] A. Lemenant, *About the regularity of average distance minimizers in* $\mathbb{R}^2$*,* preprint (2009).

[48] F.H. Lin, P. Zhang, *On the hydrodynamic limit of Ginzburg-Landau vortices.* Discrete cont. dyn. systems **6** (2000), 121–142.

[49] J.Maas, *Analysis of infinite-dimensional diffusions*, PhD thesis (2009).

[50] E. Mainini, *A global uniqueness result for an evolution problem arising in superconductivity*, Boll. Unione Mat. Ital. (9) **II** (2009), no.2, 509–528.

[51] C. Mantegazza, A. Mennucci, *Hamilton-Jacobi equations and distance functions on Riemannian manifolds,* Appl. Math. Optim. **47** (2003), no.1, 1–25.

[52] N. Masmoudi, P. Zhang, *Global solutions to vortex density equations arising from sup-conductivity.* Ann. Inst. H. Poincaré Anal. Non Linéaire **22** (2005), no.4, 441–458.

[53] R. J. McCann, *A convexity principle for interacting gases,* Adv. Math. **128** (1997), no.1, 153–179.

[54] T. Mikami, *Dynamical systems in the variational formulation of the Fokker-Planck equation by the Wasserstein metric,* Appl. Math. Optim. **42** (2000), 203–227.

[55] G. Monge, *Memoire sur la Theorie des Déblais et des Remblais,* Hist. de l'Acad. des Sciences de Paris (1781).

[56] S. Mosconi, P. Tilli, Γ*-convergence for the irrigation problem.* Journal of Convex Analysis, **12** (2005), no.1, 145–158.

[57] F. Otto, *Doubly degenerate diffusion equations as steepest descents,* Manuscript (1996).

[58] F. Otto, *Evolution of microstructure in unstable porous media flow: a relaxational approach,* Comm. Pure Appl. Math. **52** (1999), 873–915.

[59] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation,* Comm. Partial Differential Equations **26** (2001), 101–174.

[60] F. Otto, *Dynamics of Labyrinthine Pattern Formation in Magnetic Fluids: A Mean-Field Theory,* Arch.Rational Mech. Anal. **141** (1998), 63–103.

[61] E. Paolini, E. Stepanov, *Qualitative properties of maximum distance minimizers and average distance minimizers in* $\mathbf{R}^n$*,* J. Math. Sciences (N.Y.) **122** (2004), no.3, 105–122.

[62] S.T. Rachev, L. Rüschendorf, *Mass transportation problems* (vol. 1: Theory). Probability and its Applications, Springer-Verlag, New York (1998).

[63] W. Rudin, *Real and complex analysis.* Third edition. McGraw-Hill Book Co., New York (1987).

[64] E. Sandier, S. Serfaty, *A Rigorous Derivation of a Free-Boundary Problem Arising in Superconductivity,* Ann. Scien. Ecole Normale Supérieure, 4e ser **33** (2000), 561–592.

[65] E. Sandier, S. Serfaty, *Limiting Vorticities for the Ginzburg-Landau equations Duke Math. J.* **117** (2003), 403–446.

[66] F. Santambrogio, P. Tilli, *Blow-up of optimal sets in the irrigation problem,* J. Geom. Anal. **15** (2005), no.2, 343–362.

[67] L. Schwartz, *Radon measures on arbitrary topological spaces and cylindrical measures.* Published for the Tata Institute of Fundamental Research, Bombay, by Oxford University Press, London (1973). Tata Institute of Fundamental Research Studies in Mathematics, No. 6.

[68] E. Stepanov, *Partial geometric regularity of some optimal connected transportation networks*, J. Math. Sciences (N.Y.) **132** (2006), no.4, 522–552.

[69] P. Tilli, *Some explicit examples of minimizers for the irrigation problem,* J. Conv. Anal. **17** (2010), no.2.

[70] M. Tinkham, *Introduction to Superconductivity,* second Edition, McGraw Hill (1996).

[71] J.L.Vázquez, *The porous medium equation. Mathematical Theory.* Oxford Mathematical Monographs, Oxford (2007).

[72] C. Villani, *Optimal transport, old and new.* Springer-Verlag (2008).

[73] C. Villani, *Topics in optimal transportation,* vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI (2003).

[74] A. Visintin, *Strong convergence results related to strict convexity,* Comm. PDE. **9** (1984), 439–466.

[75] V. Yudovich, *Nonstationary flow of an ideal incompressible liquid,* Zhurn. Vych. Mat. **3** (1963), 1032–1066.

[76] V. Yudovich, *Some bounds for solutions of elliptic equations,* Mat. Sb. **59** (1962), 229–244. English transl. in Amer. Mat. Soc. Transl. (2) **56** (1962).