Harmonic mappings valued in the Wasserstein space

Hugo Lavenant

HARMONIC MAPPINGS VALUED IN THE WASSERSTEIN SPACE

HUGO LAVENANT

ABSTRACT. We propose a definition of the Dirichlet energy (which is roughly speaking the integral of the square of the gradient) for mappings $\boldsymbol{\mu} : \Omega \to (\mathcal{P}(D), W_2)$ defined over a subset Ω of $\mathbf{R}^p$ and valued in the space $\mathcal{P}(D)$ of probability measures on a compact convex subset D of $\mathbf{R}^q$ endowed with the quadratic Wasserstein distance. Our definition relies on a straightforward generalization of the Benamou-Brenier formula (already introduced by Brenier) but is also equivalent to the definition of Koorevaar, Schoen and Jost as limit of approximate Dirichlet energies, and to the definition of Reshetnyak of Sobolev spaces valued in metric spaces.

We study harmonic mappings, i.e. minimizers of the Dirichlet energy provided that the values on the boundary $\partial\Omega$ are fixed. The notion of constant-speed geodesics in the Wasserstein space is recovered by taking for Ω a segment of $\mathbf{R}$. As the Wasserstein space $(\mathcal{P}(D), W_2)$ is positively curved in the sense of Alexandrov we cannot apply the theory of Koorevaar, Schoen and Jost and we use instead arguments based on optimal transport. We manage to get existence of harmonic mappings provided that the boundary values are Lipschitz on $\partial\Omega$, uniqueness is an open question.

If Ω is a segment of $\mathbf{R}$, it is known that a curve valued in the Wasserstein space $\mathcal{P}(D)$ can be seen as a superposition of curves valued in D. We show that it is no longer the case in higher dimensions: a generic mapping $\Omega \to \mathcal{P}(D)$ cannot be represented as the superposition of mappings $\Omega \to D$.

We are able to show a Ishihara-type property: the composition $F \circ \boldsymbol{\mu}$ of a function $F : \mathcal{P}(D) \to \mathbf{R}$ convex along generalized geodesics and a harmonic mapping $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ is a subharmonic real-valued function.

We also study the special case where we restrict ourselves to a given family of elliptically contoured distributions (a finite-dimensional and geodesically convex submanifold of $(\mathcal{P}(D), W_2)$ which generalizes the case of Gaussian measures) and show that it boils down to harmonic mappings valued in the Riemannian manifold of symmetric matrices endowed with the distance coming from optimal transport.

CONTENTS

LABORATOIRE DE MATHÉMATIQUES D'ORSAY, UNIV. PARIS-SUD, CNRS, UNIVERSITÉ PARIS-SACLAY, 91405 ORSAY CEDEX, FRANCE
E-mail address: hugo.lavenant@u-psud.fr.
Date: December 20, 2017.

1. Introduction

1.1. **Harmonic mappings.** If $f : \Omega \to \mathbf{R}$ is a real-valued function defined on a subset Ω of $\mathbf{R}^p$, one says that f is *harmonic* if

$$\Delta f = 0, \tag{1.1}$$

where $\Delta = \sum_{\alpha=1}^{p} \partial_{\alpha\alpha}$ denotes the Laplacian operator. Although this equation can be traced back to physics (for instance it corresponds to the equation satisfied by the electric potential in the absence of charge, or the one satisfied by the temperature in some homogeneous and isotropic medium when the permanent regime is reached), it has revealed to have its own mathematical interest [14]. In particular it is associated to a concept of *equilibrium*, as for an harmonic function f, the value of f at a point $\xi \in \Omega$ is always equal to the mean of the values of f on a ball centered at ξ. A whole line of research has been devoted to define harmonic mappings $f : X \to Y$ where X and Y are spaces without a structure as strong as the Euclidean one. If X and Y are Riemannian manifolds, one can define an analogue of (1.1) which involves the metric tensors of both X and Y (see for instance [11] or, for a modern presentation, [17, 14]). The standard hypothesis to get existence results and nice properties of harmonic mappings is that X has a positive curvature and Y has a negative curvature. In the 90s, Koorevaar and Schoen [19] on one side and Jost [16] on the other side, presented independently a more general setting and showed that one can define harmonic mappings $f : \Omega \to Y$ provided that Ω is a compact Riemannian manifold and Y is a metric space with negative curvature in the sense of Alexandrov [19, Section 2.1].

The most robust point of view for the definition of harmonic mappings valued in metric spaces is related to the Dirichlet problem. Indeed, if we go back to the case where $Y = \mathbf{R}$, a function

$f : \Omega \to \mathbf{R}$ is harmonic if and only if it is a minimizer of the Dirichlet energy

$$\mathrm{Dir}(g) := \int_\Omega \frac{1}{2}|\nabla g(\xi)|^2 \mathrm{d}\xi$$

among all functions $g : \Omega \to \mathbf{R}$ having the same values as f on $\partial\Omega$ the boundary of Ω. The main advantage of this formulation is that it involves only first order derivatives, and most of the concepts involving first order derivatives can be defined on metric spaces even without any vectorial structure [3]. Koorevaar, Schoen and Jost proved that for every separable metric space Y, one can define the analogue of the Dirichlet energy of any mapping $f : \Omega \to Y$. Then under the assumption that Y has a negative curvature in the sense of Alexandrov, they proved existence and uniqueness of a minimizer of the Dirichlet energy (provided that the values at the boundary $\partial\Omega$ are fixed), interior and boundary regularity of the minimizer and lots of other properties similar to harmonic mappings between manifolds. Most of the proofs mimic the ones in the Euclidean case and rely only on the curvature properties of the target space Y. To quote Koorevaar and Schoen: "*We find the generality, elegance, and simplicity of the proofs presented here to be an indication that we have found the proper framework for their expression*" [19, p. 614].

In this article, our goal is to define and to study harmonic mappings defined over a compact domain Ω of $\mathbf{R}^p$ and valued in the space of probability measures endowed with the distance coming from optimal transport, the so-called quadratic Wasserstein space [31, 4, 27]. If D is a convex compact domain of $\mathbf{R}^q$, and if μ, ν are two probability measures on D (the set of probability measures on D is denoted by $\mathcal{P}(D)$) then the (quadratic) Wasserstein distance $W_2(\mu, \nu)$ between the two is defined as

$$W_2(\mu, \nu) := \inf_\pi \sqrt{\int_{D\times D} |x - y|^2 \mathrm{d}\pi(x, y)},$$

where the infimum is taken over all transport plans $\pi \in \mathcal{P}(D \times D)$ whose marginals are μ and ν. We will define the Dirichlet energy for mappings $\boldsymbol{\mu} : \Omega \to (\mathcal{P}(D), W_2)$ and study its minimizers under the constraint that the values at the boundary $\partial\Omega$ are fixed. It is known [4, Section 7.3] that $(\mathcal{P}(D), W_2)$ is a positively curved space in the sense of Alexandrov, hence the whole theory of Koorevaar, Schoen and Jost does not apply: we have to leave the world of "*generality, elegance and simplicity*". Though we manage to develop a fairly satisfying theory of Dirichlet energy and harmonic mappings valued in the Wasserstein space, it is *ad hoc*: it intensively relies on specific properties of $(\mathcal{P}(D), W_2)$ and is hardly generalizable to other positively curved spaces.

1.2. **Related works.** This work can be seen as an extension of an article written by Brenier [9] almost 15 years ago. Recently, few articles [28, 29, 32, 21] have been published on related topics even though none of them seems aware of Brenier's work.

In Section 3 of [9], Brenier proposed a definition of what he called *generalized harmonic functions* which is the same thing as our harmonic mappings valued in the Wasserstein space. He defined the Dirichlet energy for such mappings; proved the existence of harmonic mappings in some special cases and gave an explicit solution in the very special case where all measures on $\partial\Omega$ are Dirac masses; indicated the formulation of the dual problem; and formulated some conjectures. In the present article, we will rely on the same definition of Dirichlet energy than in Brenier's article, but we push the analysis much further: we provide a rigorous functional analysis framework; link the Dirichlet energy with already known notions of analysis in metric spaces (in particular with the definition of Koorevaar, Schoen and Jost; prove the existence of harmonic mappings in a more general context; and answer to Brenier's conjectures.

In [28], the authors study *soft maps* (which are nothing more than maps $\Omega \to \mathcal{P}(D)$ except that Ω and D are surfaces, i.e. Riemannian manifolds of dimension 2) and define a Dirichlet energy in the same way as Koorevaar, Schoen and Jost. These maps are seen as relaxations of "classical"

maps $\Omega \to D$, and they focus on numerical computation and visualization of theses soft maps, see also [29] for applications to supervised learning. On the other hand, they do not analyze in detail the theoretical properties of the Dirichlet energy and harmonic mappings, which in contrast is the main topic of the present article. In [21], the author provides some theoretical analysis of soft maps by focusing on the cases where the boundary measures on $\partial\Omega$ are either Dirac masses or Gaussian measures.

Finally, in [32] the authors also study mappings valued in the space of probability measures, but are rather interested in the bounded variation norm (the integral of the norm of the gradient) than in the Dirichlet energy. Their provide applications to the denoising of measure-valued images.

Apart from these articles, let us underline the interest of our work by relating it to other already known concepts:

- It is well known that harmonic mappings defined over an interval of $\mathbf{R}$ and valued in a geodesic space are precisely the constant-speed geodesics, and it is the case with our definition. Thus our work can be seen as extending the definition of geodesics in the Wasserstein space, the latter being an object which is now well understood.
- As we said above, our definition of Dirichlet energy coincides with the one of Koorevaar, Schoen and Jost. In particular, our work shows that their definition can be applied to positively curved spaces and still get some non trivial result, even though we rely on the very special structure of the Wasserstein space.
- To study the regularity of minimal surfaces, Almgren proposed the notion of Q-valued functions (see [5] or [10] for a clear and self-contained reference), which can be seen (up to renormalization) as mappings defined on $\Omega \subset \mathbf{R}^p$ and valued in the subset $\mathcal{A}_Q(D)$ (where $Q \geqslant 1$ is an integer) of the Wasserstein space $(\mathcal{P}(D), W_2)$ defined as

$$\mathcal{A}_Q(D) := \left\{ \frac{1}{Q} \sum_{i=1}^{Q} \delta_{x_i} \ : \ (x_1, x_2, \dots, x_Q) \in D^Q \right\}.$$

  In other words, $\mathcal{A}_Q(D)$ is the set of probability measures which are combinations of at most Q Dirac masses with weights which are multiples of $1/Q$, and is endowed with the Wasserstein distance W_2. To put it shortly, a Q-function is a function which in every point takes Q unordered different values (counted with multiplicity). There exists a beautiful existence and regularity theory for harmonic Q-functions. As $\bigcup_{Q\geqslant 1} \mathcal{A}_Q(D)$ is dense in $\mathcal{P}(D)$, it would be tempting to see the Dirichlet problem for mappings valued in the Wasserstein space $\mathcal{P}(D)$ as the limit as $Q \to +\infty$ of the Dirichlet problem for Q-functions. However, it is not so obvious that this limit really holds, and most of the results in the theory of Q-functions are proved by induction on Q through clever decompositions and combinatorial arguments, hence they depend heavily on Q and not much can be passed to the limit $Q \to +\infty$. Notice that the space $\mathcal{A}_Q(D)$ is also positively curved in the sense of Alexandrov (the example in [4, Section 7.3] lives in $\mathcal{A}_2(D)$), hence the theory of Q-functions is a theory of harmonic mappings valued in a positively curved space. However, it is known [10, Theorem 2.1] that $\mathcal{A}_Q(D)$ is in a bilipschitz bijection with a subset of $\mathbf{R}^N$ for some large N: with Q-functions we stay in the finite-dimensional world. On the contrary, in the present article, the target space $(\mathcal{P}(D), W_2)$ will be both positively curved and genuinely infinite-dimensional.

1.3. **Main definitions and results.** Let us go into the details and summarize the content of this article as well as the key insights. In this discussion we will stay informal, with sometimes sloppy or non rigorous statements.

In Section 2, we give our notations and collect some known facts about the Wasserstein space, which can be found in standard textbooks. In particular, we take Ω and D two compact domains of respectively $\mathbf{R}^p$ and $\mathbf{R}^q$ and assume that D is convex.

Section 3 is devoted to the definition and properties of the Dirichlet energy of a mapping $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$.

The idea is to start from curves valued in the Wasserstein space and the so-called Benamou-Brenier formula [7]. If I is a segment of $\mathbf{R}$ and $\boldsymbol{\mu} : I \to \mathcal{P}(D)$ is an absolutely continuous curve, then its Dirichlet energy, which is nothing else than the integral of the square of its metric derivative [4, Section 1.1] is equal to

$$\mathrm{Dir}(\boldsymbol{\mu}) = \inf \left\{ \int_I \left(\int_D \frac{1}{2} |\mathbf{v}(t,x)|^2 \boldsymbol{\mu}(t, \mathrm{d}x) \right) \mathrm{d}t \ : \ \mathbf{v} : I \times D \to \mathbf{R}^q \ \text{ and } \partial_t \boldsymbol{\mu} + \nabla \cdot (\boldsymbol{\mu} \mathbf{v}) = 0 \right\},$$

which means that one minimizes the integral over time of the kinetic energy among all velocity fields $\mathbf{v}$ such that the continuity equation $\partial_t \boldsymbol{\mu} + \nabla \cdot (\boldsymbol{\mu} \mathbf{v}) = 0$ is satisfied. What Benamou and Brenier understood is that the correct variable is the momentum $E = \mathbf{v}\boldsymbol{\mu}$. Indeed, the continuity equation $\partial_t \boldsymbol{\mu} + \nabla \cdot E = 0$ becomes a linear constraint and

$$\int_I \left(\int_D \frac{1}{2} |\mathbf{v}(t,x)|^2 \boldsymbol{\mu}(t, \mathrm{d}x) \right) \mathrm{d}t = \int_{I \times D} \frac{|E|^2}{2\boldsymbol{\mu}}$$

is a convex function of the pair $(\boldsymbol{\mu}, E)$. In particular, to find the constant-speed geodesic between μ and $\nu \in \mathcal{P}(D)$, assuming that $I = [0,1]$, one minimizes the convex Dirichlet energy over the pairs $(\boldsymbol{\mu}, E)$ with linear constraints given by the continuity equation, that $\boldsymbol{\mu}(0) = \mu$ and that $\boldsymbol{\mu}(1) = \nu$.

As noticed in [9, Section 3], this formulation can be directly extended to the case where the starting space is no longer of dimension 1: if Ω is a subset of $\mathbf{R}^p$, one can define a (generalized) continuity equation for the pair $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ and $E : \Omega \times D \to \mathbf{R}^{pq}$ by

$$\nabla_\Omega \boldsymbol{\mu} + \nabla_D \cdot E = 0, \tag{1.2}$$

where ∇_Ω stands for the gradient w.r.t. variables in Ω and $\nabla_D \cdot$ stands for the divergence w.r.t. variables in D. More precisely if $(E^{\alpha i})_{1 \leqslant \alpha \leqslant p, 1 \leqslant i \leqslant q}$ denote the components of E, and if the derivatives w.r.t. Ω (resp. D) are denoted by $(\partial_\alpha)_{1 \leqslant \alpha \leqslant p}$ (resp. $(\partial_i)_{1 \leqslant i \leqslant q}$) then the the continuity equation reads: for any $\alpha \in \{1, 2, \dots, p\}$,

$$\partial_\alpha \boldsymbol{\mu} + \sum_{i=1}^{q} \partial_i E^{\alpha i} = 0.$$

The Dirichlet energy of the pair $(\boldsymbol{\mu}, E)$ is defined as

$$\int_{\Omega \times D} \frac{|E|^2}{2\boldsymbol{\mu}} = \int_{\Omega \times D} \sum_{\alpha=1}^{p} \sum_{i=1}^{q} \frac{|E^{i\alpha}|^2}{2\boldsymbol{\mu}},$$

and $\mathrm{Dir}(\boldsymbol{\mu})$, the Dirichlet energy of $\boldsymbol{\mu}$, is the minimal Dirichlet energy of the pairs $(\boldsymbol{\mu}, E)$ among all E such that the continuity equation is satisfied (Definition 3.7). It is a straightforward copy of the classical proofs of optimal transport to show that there exists a unique optimal momentum E (which we call the tangent momentum) which is written $E = \mathbf{v}\boldsymbol{\mu}$ for some velocity field $\mathbf{v} : \Omega \times D \to \mathbf{R}^{pq}$, and that Dir is convex and lower semi-continuous (l.s.c.).

We will prove that for $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$, one has $\mathrm{Dir}(\boldsymbol{\mu}) < +\infty$ if and only if for any $u : \mathcal{P}(D) \to \mathbf{R}$ which is 1-Lipschitz, one has that $u \circ \boldsymbol{\mu}$ belongs to $H^1(\Omega)$ with $|\nabla(u \circ \boldsymbol{\mu})| \leqslant g$, where $g \in L^2(\Omega)$ is independent of u. Moreover, the minimal g will be shown to be equal to

$$g(\xi) = \sqrt{\int_D |\mathbf{v}(\xi, x)|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x)}$$

for a.e. $\xi \in \Omega$, where $E = \mathbf{v}\boldsymbol{\mu}$ is the tangent momentum (Theorem 3.17). This precisley shows that the space $\{\boldsymbol{\mu} : \Omega \to \mathcal{P}(D) \ : \ \mathrm{Dir}(\boldsymbol{\mu}) < +\infty\}$ coincides with the set $H^1(\Omega, \mathcal{P}(D))$, where the latter is defined in the sense of Reshetnyak [25], and that the gradient of $\boldsymbol{\mu}$ in the sense of Reshetnyak (the minimal g above) is related to the tangent $\mathbf{v}$. For instance, it implies that Lipschitz mappings $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ (i.e. such that $W_2(\boldsymbol{\mu}(\xi), \boldsymbol{\mu}(\eta)) \leqslant C|\xi - \eta|$ for any $\xi, \eta \in \Omega$) have a finite Dirichlet energy.

We will also prove that our Dirichlet energy coincides with the one of Koorevaar and Schoen, as well as Jost. The idea of theses authors goes as follows: if $f : \Omega \to \mathbf{R}$ is smooth, then for any $\xi \in \mathbf{R}^p$,

$$|\nabla f(\xi)|^2 = \lim_{\varepsilon \to 0} C_p \int_{B(\xi,\varepsilon)} \frac{|f(\eta) - f(\xi)|^2}{\varepsilon^{p+2}} \mathrm{d}\eta,$$

for some constant C_p which depends on the dimension p of Ω, where $B(\xi, \varepsilon)$ is the ball of center ξ and radius ε. Thus, if $\varepsilon > 0$ is small, a good approximation of the Dirichlet energy of f would be

$$\mathrm{Dir}(f) = \int_\Omega \frac{1}{2} |\nabla f(\xi)|^2 \mathrm{d}\xi \simeq C_p \int_{\Omega \times \Omega} \frac{|f(\xi) - f(\eta)|^2}{2\varepsilon^{p+2}} \mathbb{1}_{|\xi - \eta| \leqslant \varepsilon} \mathrm{d}\xi \mathrm{d}\eta.$$

Notice that the right hand side (r.h.s.) involves only metric quantities, thus its definition can be extended if $f : \Omega \to Y$ where (Y, d) is an arbitrary metric space by replacing $|f(\xi) - f(\eta)|^2$ by $d(f(\xi), f(\eta))^2$: this is what is done and extensively studied in [19, Section 1] (curvature assumptions on Y are not required for the definition of the Dirichlet energy, but are used to derive existence, uniqueness and properties of the minimizers). The counterpart in our case is to define the ε-Dirichlet energy of a mapping $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ by

$$\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}) := C_p \int_{\Omega \times \Omega} \frac{W_2^2(\boldsymbol{\mu}(\xi), \boldsymbol{\mu}(\eta))}{2\varepsilon^{p+2}} \mathbb{1}_{|\xi - \eta| \leqslant \varepsilon} \mathrm{d}\xi \mathrm{d}\eta.$$

We are able show that Dir_ε converges to Dir as $\varepsilon \to 0$: it holds pointwisely but also in the sense of Γ-convergence (Theorem 3.24). For both the equivalence with the definition of Koorevaar, Schoen and Jost, or with the one of Reshetnyak, the difficulty is not to guess them (they are fairly simple at the formal level) but to conduct careful approximation arguments.

To conclude the section, we will show how one can define values on $\partial\Omega$ for mappings $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ with finite Dirichlet energy. There already exists a trace theory in [19], however in view of the dual formulation for the Dirichlet problem, we prefer to define trace values by extending the continuity equation up to the boundary of Ω. Indeed, multiplying (1.2) by a test function $\varphi \in C^1(\Omega \times D, \mathbf{R}^p)$ valued in $\mathbf{R}^p$, we get the following weak formulation:

$$\int_{\Omega \times D} \nabla_\Omega \cdot \varphi \mathrm{d}\boldsymbol{\mu} + \int_{\Omega \times D} \nabla_D \varphi \cdot \mathrm{d}E = \int_{\partial\Omega} \left(\int_D \varphi(\xi, x) \cdot \mathbf{n}_\Omega(\xi) \boldsymbol{\mu}(\xi, \mathrm{d}x) \right) \sigma(\mathrm{d}\xi),$$

where $\mathbf{n}_\Omega$ is the outward normal to $\partial\Omega$ and σ the surface measure. We will show that, if $\mathrm{Dir}(\boldsymbol{\mu}) < +\infty$, then the r.h.s. can always be defined as a finite vector-valued measure acting on φ called $\mathrm{BT}_{\boldsymbol{\mu}}$ (Theorem 3.25). Two mappings will have the same values on the boundary $\partial\Omega$ if, by definition, they have the same boundary term.

In Section 4 we define the Dirichlet problem and establish its dual formulation. This is fairly classic in optimal transport theory, our proofs do not bring any new ideas.

To define the Dirichlet problem, we assume that a mapping $\boldsymbol{\mu}_b : \Omega \to \mathcal{P}(D)$ with finite Dirichlet energy is given and we study

$$\min\{\mathrm{Dir}(\boldsymbol{\mu}) \ : \ \boldsymbol{\mu} = \boldsymbol{\mu}_b \text{ on } \partial\Omega\}.$$

Thanks to the Benamou-Brenier formulation, existence of a solution is a straightforward application of the direct method of calculus of variations (Theorem 4.3). As we discuss it in the core of the article, we do not know if uniqueness holds. Only in some particular case where the boundary

values belong to a family of elliptically contoured distributions we are able to prove uniqueness (see below in the introduction).

In the formulation of the Dirichlet problem, we define the boundary conditions through a mapping $\boldsymbol{\mu}_b$ defined on the whole Ω. A natural question arises: if $\boldsymbol{\mu}_b : \partial\Omega \to \mathcal{P}(D)$ is given, is it possible to extend it on Ω in such a way that $\mathrm{Dir}(\boldsymbol{\mu}_b) < +\infty$? We will show that the answer to this question is positive if $\boldsymbol{\mu}_b$ is Lipschitz on $\partial\Omega$, indeed in this case one can extend it as a Lipschitz mapping on Ω. The question of the existence of a Lipschitz extension for mappings $f : Z \to Y$, where $Z \subset X$ and X, Y are metric spaces has been intensively studied, see for instance [20, 24] and references therein. The general philosophy is that lower bounds on the curvature are required for the starting space X, whereas upper bounds on the curvature are required for the target space Y. In our case, there are no upper bounds for the curvature of the target space $\mathcal{P}(D)$, hence we cannot apply classical results. However, we use the fact that we want to extend Lipschitz mappings defined not on an arbitrary closed subset of Ω, but on the boundary $\partial\Omega$ which has some regularity. By some *ad hoc* construction, we are able to treat the case where Ω is a ball, but we cannot control the Lipschitz constant of the extension on Ω by the Lipschitz constant of the mapping on $\partial\Omega$. Nevertheless, we can conclude for smooth domains, as they can be cut in a finite number of pieces, each piece being in a bilipschitz bijection with a ball (Theorem 4.4).

Let us establish here the dual formulation *via* a formal $\inf - \sup$ exchange, it was already done in [9]. Indeed, given the definition of Dir and the weak formulation of the continuity equation,

$$\begin{aligned}
&\min\{\mathrm{Dir}(\boldsymbol{\mu}) \ : \ \boldsymbol{\mu} = \boldsymbol{\mu}_b \text{ on } \partial\Omega\} \\
&\qquad = \inf_{\boldsymbol{\mu},\mathbf{v}} \left[\int_{\Omega\times D} \frac{1}{2}|\mathbf{v}|^2 \boldsymbol{\mu} + \sup_{\varphi\in C^1(\Omega\times D,\mathbf{R}^p)} \left(\mathrm{BT}_{\boldsymbol{\mu}_b}(\varphi) - \int_{\Omega\times D} \nabla_\Omega\cdot\varphi \mathrm{d}\boldsymbol{\mu} - \int_{\Omega\times D} \nabla_D\varphi\cdot\mathbf{v}\boldsymbol{\mu} \right) \right] \\
&\qquad = \sup_{\varphi\in C^1(\Omega\times D,\mathbf{R}^p)} \left[\mathrm{BT}_{\boldsymbol{\mu}_b}(\varphi) + \inf_{\boldsymbol{\mu},\mathbf{v}} \int_{\Omega\times D} \left(\frac{1}{2}|\mathbf{v}|^2 - \nabla_D\varphi\cdot\mathbf{v} - \nabla_\Omega\cdot\varphi \right) \boldsymbol{\mu} \right].
\end{aligned}$$

Optimizing in $\mathbf{v}$, we have that $\mathbf{v} = \nabla_D\varphi$, and then the infimum in $\boldsymbol{\mu}$ is translated into the constraint $\nabla_\Omega\cdot\varphi + \frac{1}{2}|\nabla_D\varphi|^2 \leqslant 0$. Hence, we have (formally, and it is proved rigorously in the core of the article, see Theorem 4.7) the following identity:

$$\sup\left\{ \mathrm{BT}_{\boldsymbol{\mu}_b}(\varphi) \ : \ \varphi\in C^1(\Omega\times D,\mathbf{R}^p) \text{ and } \nabla_\Omega\cdot\varphi + \frac{|\nabla_D\varphi|^2}{2} \leqslant 0 \right\} = \min\{\mathrm{Dir}(\boldsymbol{\mu}) \ : \ \boldsymbol{\mu} = \boldsymbol{\mu}_b \text{ on } \partial\Omega\}.$$

We do not have an existence result for solutions φ of the dual problem. Notice that φ is a vector-valued function, but there is only a scalar constraint on it: the dual problem looks harder than in the case where Ω is a segment of $\mathbf{R}$. Formally, as it is done in [9], one can get optimality conditions out of the dual formulation. Indeed, we have that $\mathbf{v} = \nabla_D\varphi$ and, from the optimization in $\boldsymbol{\mu}$, that $\nabla_\Omega\cdot\varphi + \frac{1}{2}|\nabla_D\varphi|^2 = 0$ $\boldsymbol{\mu}$-a.e. If we assume that $\boldsymbol{\mu}$ is strictly positive a.e., we end up with the following system for $\mathbf{v}$ (the first equation is just a rewriting of the fact that $\mathbf{v}$ is a gradient, the second one is obtained by differentiating $\nabla_\Omega\cdot\varphi + \frac{1}{2}|\nabla_D\varphi|^2 = 0$ w.r.t. D):

$$\begin{cases}
\partial_i\mathbf{v}^{\alpha j} = \partial_j\mathbf{v}^{\alpha i} & \text{for } \alpha\in\{1,2,\dots,p\} \text{ and } i,j\in\{1,2,\dots,q\}, \\
\displaystyle\sum_{\alpha=1}^{p} \partial_\alpha\mathbf{v}^{\alpha i} + \sum_{\alpha=1}^{p}\sum_{j=1}^{q} \mathbf{v}^{\alpha j}\partial_j\mathbf{v}^{\alpha i} = 0 & \text{for } i\in\{1,2,\dots,q\}.
\end{cases}$$

However, we will not push the analysis further and try to derive a rigorous version of theses optimality conditions, it might be the topic of an other study.

In Section 5, we answer to a a problem formulated by Brenier [9, Problem 3.5]. The question is the following: if $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$, does there exists a probability Q over functions $f : \Omega \to D$ such that

$\boldsymbol{\mu}$ is represented by f, i.e.

$$\int_D a(x)\boldsymbol{\mu}(\xi, \mathrm{d}x) = \int a(f(\xi))Q(\mathrm{d}f)$$

for all $a \in C(D)$ continuous and $\xi \in \Omega$; and such that the Dirichlet energy is the mean of the Dirichlet energy of the f:

$$\mathrm{Dir}(\boldsymbol{\mu}) = \int \left(\int_\Omega \frac{1}{2} |\nabla f(\xi)|^2 \mathrm{d}\xi \right) Q(\mathrm{d}f)?$$

If Ω is a segment of $\mathbf{R}$ the answer is positive as shown in [4, Section 8.2] (it is known as the probabilistic representation or the superposition principle). However, as soon as Ω is more than 2-dimensional (in fact it already fails if Ω is a circle), the answer becomes negative (Corollary 5.7). We will provide a counterexample and explain the obstruction.

The main consequence is the following: there is no Lagrangian formulation for mappings $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$. There can be no static formulation of the Dirichlet problem analogue to transport plans or multimarginal formulation. One is forced to work only with the Eulerian formulation, namely the Benamou-Brenier formula. It explains why it is substantially more difficult to study mappings $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ as soon as the dimension of Ω is larger than 2, as most of the difficult results of optimal transport are proved thanks to the Lagrangian point of view.

In Section 6, we prove a Ishihara type property (see below) for harmonic mappings valued in the Wasserstein space, which in particular allows us to give a positive answer to [9, Conjecture 3.3]. This can be considered as the main result of this article, and many properties proved before this section are there to prepare this proof.

If $f : \Omega \to \mathbf{R}$ is a real-valued harmonic function, then $(F \circ f) : \Omega \to \mathbf{R}$ is a subharmonic function for every $F : D \to \mathbf{R}$ convex, which means that $\Delta(F \circ f) \geqslant 0$. It can be checked by a direct computation using the chain rule. If we take $f : X \to Y$, where X and Y are two Riemannian manifolds, then the result still holds (provided that harmonicity, subharmonicity and convexity are properly defined through the Riemannian structures) and it is even a characterization of harmonic mappings: this was first remark by Ishihara [15] (hence the denomination "Ishihara type property"), one can find a statement and a proof in [17, Corollary 8.2.4]. Extensions of this result when the target is a metric space with negative curvature are available, see for instance [30, Section 7].

In the Wassertein space, mappings which are convex w.r.t. the metric structure, which means convex along geodesics, are well understood (see for instance [4, Chapter 9] or [27, Chapter 7]). Actually, we will need something a little stronger, which is convexity along generalized geodesics [4, Section 9.2]. In our case the Ishihara property reads: if $F : \mathcal{P}(D) \to \mathbf{R}$ is convex along generalized geodesics and if $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ is a solution of the Dirichlet problem, then $(F \circ \boldsymbol{\mu}) : \Omega \to \mathbf{R}$ is subharmonic (Theorem 6.3).

The proof of geodesic convexity usually relies on the Lagrangian formulation, which, as we said above, is not available in our case. To overcome this difficulty, we use the approximate Dirichlet energies Dir_ε as a substitute for Dir. Indeed, as explained by Jost [16], if $\boldsymbol{\mu}_\varepsilon$ is a minimizer of Dir_ε (with for instance fixed values around the boundary $\partial\Omega$), then for a.e. $\xi \in \Omega$, $\boldsymbol{\mu}_\varepsilon(\xi)$ is a minimizer of

$$\nu \mapsto \int_{B(\xi,\varepsilon)} W_2^2(\nu, \boldsymbol{\mu}_\varepsilon(\eta)) \mathrm{d}\eta,$$

in other words $\boldsymbol{\mu}_\varepsilon(\xi)$ is a barycenter of the $\boldsymbol{\mu}_\varepsilon(\eta)$, for $\eta \in B(\xi, \varepsilon)$ (for barycenters in the Wasserstein space, see [1] for the finite case and [8, 18] for the infinite case). Notice that if $f : \Omega \to \mathbf{R}$ is real-valued and harmonic, then for any $\varepsilon > 0$ $f(\xi)$ is the barycenter of $f(\eta)$ for $\eta \in B(\xi, \varepsilon)$, while in the metric case this property only holds asymptotically as $\varepsilon \to 0$. For barycenters in the Wasserstein space, there exists a generalized Jensen inequality: it was already proved in the finite case by Agueh and Carlier [1, Proposition 7.6] under the assumption that F is convex along generalized

geodesics, and in a more general case (in particular with an infinite numbers of measures defined on a compact manifold, whereas Agueh and Carlier worked in the Euclidan space) by Kim and Pass [18, Section 7], but with rather strong regularity assumptions on the measures. We reprove this Jensen inequality in a case adapted to our context by letting the barycenter $\boldsymbol{\mu}_\varepsilon(\xi)$ follow the gradient of the functional F (for gradients flows in the Wasserstein space see [4]) and use the result as a competitor: through arguments first advanced in [22] in a very different context under the name of *flow interchange*, one can show (estimating the derivative of the Wasserstein distance along the flow of F with the so-called (EVI) inequality) that for a.e. $\xi \in \Omega$

$$\int_{B(\xi,\varepsilon)} [F(\boldsymbol{\mu}_\varepsilon(\eta)) - F(\boldsymbol{\mu}_\varepsilon(\xi))] \mathrm{d}\eta \geqslant 0. \tag{1.3}$$

Then, as Dir_ε Γ-converges to Dir, one knows that $\boldsymbol{\mu}_\varepsilon$ converges to $\boldsymbol{\mu}$ a solution of the Dirichlet problem. Passing in the limit (1.3), one concludes that $(F \circ \boldsymbol{\mu})$ is subharmonic in the sense of distributions.

Let us make a few comments. The main drawback of the proof, as we proceed by approximation and that uniqueness in the Dirichlet problem is not known, is that we are only able to show subharmonicity of $F \circ \boldsymbol{\mu}$ for one solution of the Dirichlet problem (which moreover depends on F), and not for all. To overcome this limitation, the best thing to do would be to prove uniqueness in the Dirichlet problem. Let us also discuss the regularity that we need on F. Either we require F to be continuous (which is very restrictive: it excludes the internal energies); or, if F is only lower semi-continuous, we need F to be bounded on bounded subsets of $L^\infty(D) \cap \mathcal{P}(D)$ (which is not very restrictive), but we also need the weak lower semi-continuity of

$$\boldsymbol{\mu} \mapsto \int_\Omega F(\boldsymbol{\mu}(\xi)) \mathrm{d}\xi.$$

More precisely, a mapping $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ can be seen as an element of $\mathcal{P}(\Omega \times D)$ (by "fubinization") and we require lower semi-continuity of $\boldsymbol{\mu} \mapsto \int_\Omega (F \circ \boldsymbol{\mu})$ w.r.t. the weak convergence on $\mathcal{P}(\Omega \times D)$. This weak lower semi-continuity holds heuristically if F is convex for the usual (and not geodesic) convexity on $\mathcal{P}(D)$. Indeed, even if the Dirichlet energy has a nice behavior w.r.t. geodesic convexity, the approximate Dirichlet energies Dir_ε behave well w.r.t. usual convexity. At the end of the day, the Ishihara property works for potential energies (for a convex, L^1 and lower semi-continuous potential), for internal energies (which have a super linear growth and satisfy McCann's conditions) and for the interaction energies (but only for a convex continuous interaction potential). Indeed, for a generic lower semi-continuous potential, the interaction energy W is itself lower semi-continuous on $(\mathcal{P}(D), W_2)$, but $\boldsymbol{\mu} \mapsto \int_\Omega (W \circ \boldsymbol{\mu})$ is not. Finally, notice that we do not have the converse statement: we do not know if the fact that $F \circ \boldsymbol{\mu}$ is subharmonic for any F convex along generalized geodesics is enough to prove that $\boldsymbol{\mu}$ is harmonic. To prove such a result, one would need a better understanding of the optimality conditions of the Dirichlet problem.

Finally, we conclude in Section 7 with two examples.

The first one is the case where the set D, on which the target space $\mathcal{P}(D)$ is modeled, is a segment of $\mathbf{R}$. In this case, the Wasserstein space $(\mathcal{P}(D), W_2)$ is in an isometric bijection with a convex subset of the Hilbert space $L^2([0,1])$. Hence, the Dirichlet problem reduces to the study of the Dirichlet problem for mappings valued in a Hilbert space, which is fairly simple.

The second one is the case where we restrict our attention to a family of *elliptically contoured distributions*. This terminology comes from [13] and denotes a generalization of the family of Gaussian measures. In statistics this type of family is sometimes called a *location-scatter family*. More precisely, we take $\rho \in L^1(\mathbf{R}^q)$ a positive and compactly supported function such that the measure $\rho(x)\mathrm{d}x$ has a unit mass, zero mean, and the identity matrix as covariance matrix. The family of

elliptically contoured distributions built on ρ is nothing else than the sets of measures obtained as image measures from $\rho(x)\mathrm{d}x$ by symmetric positive linear transformations. For instance, if ρ is the indicator function of a ball, the family of elliptically contoured distributions built on ρ consists in probability measures uniformly distributed on centered ellipsoids (in general the level sets of the density are ellipsoids, hence the terminology). The Gaussian case would be obtained by taking for $\rho(x)\mathrm{d}x$ a centered standard Gaussian, but this probability measure is not compactly supported (recall that we work in $\mathcal{P}(D)$ where $D \subset \mathbf{R}^q$ is compact). As in the Gaussian case, the elements of the family of elliptically contoured distributions are parametrized by their covariance matrix. Notice that it is already known that the geodesic between Gaussian measures and more generally the barycenter of Gaussian measures stay in the Gaussian family [1, Section 6.3]. If the boundary values $\boldsymbol{\mu} : \partial\Omega \to \mathcal{P}(D)$ are valued in a family of elliptically contoured distributions, we show that there exists at least one solution of the Dirichlet problem which takes values in the same family everywhere on Ω (Theorem 7.9).

Under the additional assumption that the covariance matrices on the boundary $\partial\Omega$ are non singular we are able to show much more (Theorem 7.10). It implies that there is a solution of the Dirichlet problem with covariance matrices non singular everywhere in Ω: to prove it we use the fact that the Boltzmann entropy is subharmonic because of the Ishihara type property. From this we are able to derive the Euler-Lagrange equation satisfied by the covariance matrix. Moreover we can show the uniqueness of the solution to the Dirichlet problem even in the class of mappings not necessarily valued in the family of elliptically contoured distributions. Let us give the structure of the proof as it is almost the only case where we know how to prove uniqueness. The observation is that all solutions of the Dirichlet problem must have the same tangent velocity field. Indeed, if φ is a solution of the dual problem, from optimality the tangent velocity field to any solution must be equal to $\nabla_D\varphi$. Now, if the velocity field $\nabla_D\varphi$ is regular enough (namely Lipschitz w.r.t. D), then the solution of the (1-dimensional) continuity equation with velocity field $\nabla_D\varphi$ is unique. As the (generalized) continuity equation implies the 1-dimensional one, and as all solutions of the Dirichlet problem coincide on $\partial\Omega$ they must be equal everywhere. In the case of a family of elliptically contoured distributions the tangent velocity field is linear w.r.t. D with some uniform bounds which allow us to make this argument rigorous. If we leave the world of families of elliptically contoured distributions, we do not think that we could get enough regularity on the tangent velocity field for this strategy to work. In summary, under the assumption that the covariance matrices on the boundary $\partial\Omega$ are non singular we are able to give a full solution to the problem: existence, uniqueness and Euler-Lagrange equation.

Let us comment the somehow restrictive framework that we have chosen. The compactness assumption of Ω and D allows to simplify proofs by avoiding tails estimates: we believe that there is enough technical difficulties and non trivial statements even in this case, and that the key features of the Dirichlet problem are captured, which is the reason why we have restricted ourselves to the compact case. Although we have stick to the Euclidan case, we see no deep reason which would prevent our definitions and results to be applied to the case where Ω and D are compact Riemannian manifolds. In particular, our regularization procedures rely on heat flows which are available in Riemannian manifolds. Finally, we have stick to the quadratic Wasserstein distance. We believe that if $p \in (1, +\infty)$ is given, the machinery that we use can be adapted in a straightforward way to define

$$\int_\Omega \frac{1}{p}|\nabla\boldsymbol{\mu}|^p,$$

where $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ but $\mathcal{P}(D)$ is endowed with the p-Wassertsein distance. However the Ishihiara type property is related to the Riemannian framework; also the explicit computations in the case of a family of elliptically contoured distributions are no longer avalaible. The case $p = 1$ which

corresponds the *total variation* of $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ (where $\mathcal{P}(D)$ is equipped with the 1-Wasserstein distance) has been defined and studied very recently [32] in the context of image denoising.

To conclude the introduction, let us explain the connection between the different parts of the paper. If one just wants to understand the definition of the Dirichlet problem, then Subsections 3.1, 3.5 and Section 4 are enough. Section 5 can be read independently from the rest of the article (except for Subsection 3.1 to get the definition of the objects involved). To have the full proof of the Ishihara property in Section 6, one needs also to read entirely Section 3 and Subsections 4.1, 4.2 as some necessary results are proved there. To understand the examples in Section 7, the reading of Section 3 is advised.

2. Preliminaries

2.1. **Notations.** Let p and q be two integers larger than 1. The space $\mathbf{R}^p$ and $\mathbf{R}^q$ are endowed with their Euclidean structure: the scalar product is denoted by $\cdot$ and the norm by $|\cdot|$. The closed ball of center ξ and radius r is denoted by $B(\xi, r)$. We will take $\Omega \subset \mathbf{R}^p$ and $D \subset \mathbf{R}^q$ two *compact* domains, their interior, assumed to be non empty, are denoted by $\mathring{\Omega}$ and $\mathring{D}$. The outward normal vector to $\partial\Omega$ (resp. ∂D) is denoted by $\mathbf{n}_\Omega$ (resp. $\mathbf{n}_D$). In general, all elements related to Ω will be denoted with Greek letters, and those related to D with Latin ones. For instance, points in Ω (resp. D) will be denoted by ξ, η (resp. x, y), and $(e_\alpha)_{1\leqslant\alpha\leqslant p}$ (resp. $(e_i)_{1\leqslant i\leqslant q}$) is the canonical basis of $\mathbf{R}^p$ (resp. $\mathbf{R}^q$). We make the following regularity assumptions:

Assumption. *We assume that Ω is a connected compact subset of $\mathbf{R}^p$. Moreover, $\partial\Omega$ is assumed to be Lipschitz, which means that around any point of $\partial\Omega$, up to a rotation, Ω is the epigraph of a Lipschitz function.*

We assume that D is a convex compact subset of $\mathbf{R}^q$.

Notice that we assume much regularity on D than on Ω. We will consider mappings $\Omega \to \mathcal{P}(D)$ with prescribed values on $\partial\Omega$, the regularity of the latter is important. On the contrary, we assume that D is convex, which translates in the fact that $(\mathcal{P}(D), W_2)$ is a geodesic space: in some sense, the boundary ∂D of D will be invisible.

The restriction of the Lebesgue measure on $\mathbf{R}^p$ (resp. $\mathbf{R}^q$) to Ω (resp. D) will be denoted by $\mathcal{L}_\Omega$ (resp. $\mathcal{L}_D$). To avoid normalization constants, we assume that Ω has unit mass, thus $\mathcal{L}_\Omega$ is a probability measures.

If X is a polish space (metric, complete and separable), it is endowed with its Borel σ-algebra. We define $\mathcal{P}(X)$ as the space of Borel positive measure with unit mass. It is endowed with the topology of weak convergence, which means convergence in duality with $C(X)$ the space of continuous bounded and real-valued functions defined on X. We also define $\mathcal{M}(X, \mathbf{R}^n)$, for $n \geqslant 1$ as the space of Borel (vectorial) measures valued in $\mathbf{R}^n$ with finite mass, still endowed with the topology of weak convergence. In the case $n = 1$, we use the shortcut $\mathcal{M}(X) := \mathcal{M}(X, \mathbf{R})$. If $\mu \in \mathcal{P}(X)$ or $\mathcal{M}(X, \mathbf{R}^n)$, integration w.r.t. μ is denoted by $\mathrm{d}\mu$, or by $\mu(\mathrm{d}x)$ if the variable cannot be omitted. If no measure is specified or simply $\mathrm{d}\xi$ or $\mathrm{d}x$ is used, it means that the integration is performed w.r.t. the Lebesgue measure. If $x \in X$, the Dirac mass at point x is denoted by δ_x. The indicator function of a set X will be denoted by $\mathbb{1}_X$.

If $T : X \to Y$ is a measurable application between two measurable spaces X and Y and μ is a measure on X, then the image measure of μ by T, denoted by $T\#\mu$, is the measure defined on Y by $(T\#\mu)(B) = \mu(T^{-1}(B))$ for any measurable set $B \subset Y$. It can also be defined by

$$\int_Y a(y)(T\#\mu)(\mathrm{d}y) := \int_X a(T(x))\mu(\mathrm{d}x),$$

this identity being valid as soon as $a : Y \to \mathbf{R}$ is an integrable function [4, Section 5.2].

If (X,μ) is a measured space and (Y,d) is any metric separable space, $L^2_\mu(X,Y)$ will denote the space of measurable mappings $f : X \to Y$ for which $d(f,y)^2$ integrable w.r.t. μ for some $y \in Y$. If $Y = \mathbf{R}$, then the letter Y is omitted, and if μ is the Lebesgue measure, then the letter μ is omitted. If Y is an Euclidean space, then we set

$$\|f\|^2_{L^2_\mu(X,Y)} := \int_X |f(x)|^2 \mu(\mathrm{d}x).$$

The space $H^1(\Omega)$ is the set of functions $f : \Omega \to \mathbf{R}$ such that both f and $\partial_\alpha f$, for $1 \leqslant \alpha \leqslant p$ are in $L^2(\Omega)$.

If X and Y are two subsets of Euclidean spaces, the L^∞ norm of a measurable function $f : X \to Y$ is defined as $\|f\|_\infty := \operatorname{ess\,sup}_{x\in X} |f(x)|$, where the essential supremum is taken w.r.t. the Lebesgue measure.

If X and Y are two subsets of Euclidean spaces, $C(X,Y)$ and $C^1(X,Y)$ will denote respectively the continuous and C^1 functions defined on X and valued in Y. If $Y = \mathbf{R}$, then the target space is omitted and we use $C(X)$ or $C^1(X)$. On the space $C^1(\Omega \times D, Y)$ the following differential operators can be defined. The derivatives w.r.t. Ω will be denoted by ∇_Ω, or simply $(\partial_\alpha)_{1\leqslant\alpha\leqslant p}$, and those w.r.t. D by ∇_D, or simply $(\partial_i)_{1\leqslant i\leqslant q}$. (If X is of dimension 1, the derivative of a function f will be denoted $\dot f$). The gradient will be denoted by ∇, and the divergence by $\nabla\cdot$. As an example, if $\varphi \in C^1(\Omega \times D, \mathbf{R}^p)$, with components $(\varphi^\alpha)_{1\leqslant\alpha\leqslant p}$, then $\nabla_\Omega \cdot \varphi \in C(\Omega \times D)$ is defined as

$$\nabla_\Omega \cdot \varphi(\xi,x) = \sum_{\alpha=1}^p \partial_\alpha \varphi^\alpha(\xi,x),$$

for all $\xi \in \Omega$ and $x \in D$; and $\nabla_D\varphi \in C(\Omega \times D, \mathbf{R}^{pq})$ is defined as, for any $\alpha \in \{1,2,\dots,p\}$ and $i \in \{1,2,\dots,q\}$,

$$(\nabla_D\varphi)^{\alpha i}(\xi,x) = \partial_i\varphi^\alpha(\xi,x) \in \mathbf{R}.$$

The notation $C^1_c(\mathring\Omega \times D, Y)$ will stand for the smooth functions which are compactly supported in $\mathring\Omega$ but not necessarily in D (and valued in Y): if $\varphi \in C^1_c(\mathring\Omega \times D, Y)$, it means that there exists a compact set $X \subset \mathring\Omega$ such that $\varphi(\xi,x) = 0$ as soon as $\xi \notin X$.

2.2. **The Wasserstein space.** We recall well known facts about the Wasserstein space. All these results can be found in classical books like [31, 27, 4].

We endow the space $\mathcal{P}(D)$ with the L^2-Wasserstein distance W_2. If μ and ν are elements of $\mathcal{P}(D)$, then

$$W_2(\mu,\nu) := \sqrt{\min\left\{\int_{D\times D} |x-y|^2 \pi(\mathrm{d}x,\mathrm{d}y) \ : \ \pi \in \Pi(\mu,\nu)\right\}},$$

where $\Pi(\mu,\nu)$ is the set of transport plans, i.e. of probability measures on $D \times D$ which have μ and ν as marginals. There exists at least one $\pi \in \Pi(\mu,\nu)$ realizing the infimum, it is called an optimal transport plan. The Wasserstein distance admits a dual formulation which reads

$$W_2^2(\mu,\nu) = \max\left\{\int_D \varphi(x)\mu(\mathrm{d}x) + \int_D \psi(x)\nu(\mathrm{d}x) \ : \ \varphi,\psi \in C(D) \text{ and } \forall x,y \in D, \varphi(x)+\psi(y) \leqslant |x-y|^2\right\},$$

There exists at least one solution of the dual problem, and any pair (φ,ψ) which is a solution is called a pair of Kantorovicth potentials. Notice that thanks to the dual formulation, we see that $W_2^2 : \mathcal{P}(D) \times \mathcal{P}(D) \to \mathbf{R}$ is the supremum of continuous affine functionals, hence it is convex for the affine structure on $\mathcal{P}(D)$ (and continuous by definition). In particular, there is a Jensen's inequality:

if $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ are measurable mappings defined on Ω and valued in $\mathcal{P}(D)$, and if $f : \Omega \to \mathbf{R}$ is a positive measurable function whose integral is 1, then

$$W_2^2\left(\int_\Omega \boldsymbol{\mu}(\xi)f(\xi)\mathrm{d}\xi, \int_\Omega \boldsymbol{\nu}(\xi)f(\xi)\mathrm{d}\xi\right) \leqslant \int_\Omega W_2^2(\boldsymbol{\mu}(\xi),\boldsymbol{\nu}(\xi))f(\xi)\mathrm{d}\xi.$$

In the formula above the integral $\int_\Omega \boldsymbol{\mu} f \in \mathcal{P}(D)$ is defined according to the affine structure on $\mathcal{P}(D)$ for instance by duality: for any $a \in C(D)$,

$$\int_D a\mathrm{d}\left[\int_\Omega \boldsymbol{\mu}(\xi)f(\xi)\mathrm{d}\xi\right] := \int_\Omega \left(\int_D a\mathrm{d}\boldsymbol{\mu}(\xi)\right) f(\xi)\mathrm{d}\xi. \tag{2.1}$$

The space $(\mathcal{P}(D), W_2)$ is a metric space whose topology is the one of weak convergence. In particular, according to Prokhorov's theorem, it is a compact separable space. The space $(\mathcal{P}(D), W_2)$ is a geodesic space. If $\mu, \nu \in \mathcal{P}(D)$ and $\pi \in \Pi(\mu,\nu)$ is an optimal transport plan between μ and ν, then a constant speed geodesic $\boldsymbol{\mu} : [0,1] \to \mathcal{P}(D)$ joining μ to ν is given by $\boldsymbol{\mu}(t) := f_t\#\pi$ where $f_t : (x,y) \in D \times D \mapsto (1-t)x + ty \in D$ (Notice that we have assumed D to be convex).

We will briefly use the 1-Wasserstein distance W_1 in the proof of Proposition 6.5. The definition by duality will be enough: if μ and ν are probability measures on D,

$$W_1(\mu,\nu) := \max\left\{\int_D \varphi(x)\mu(\mathrm{d}x) - \int_D \varphi(x)\nu(\mathrm{d}x) \ : \ \varphi \in C(D) \text{ and } \varphi \text{ is } 1\text{-Lipschitz}\right\}.$$

Moreover, as D is compact, there exists a constant C such that $W_2 \leqslant C\sqrt{W_1}$.

2.3. **Absolutely continuous curves in the Wasserstein space.** A central tool when one is studying the infinitesimal properties of the Wasserstein space is the concept of (2-)absolutely continuous curves valued in the Wasserstein space. Let I be a segment of $\mathbf{R}$. A curve $\boldsymbol{\mu} : I \to \mathcal{P}(D)$ is said absolutely continuous if there exists $g \in L^2(I)$ such that for any $s < t$ elements of I,

$$W_2(\boldsymbol{\mu}(t),\boldsymbol{\mu}(s)) \leqslant \int_s^t g(r)\mathrm{d}r. \tag{2.2}$$

Let us recall the following result, which holds in fact for absolutely continuous curves valued in arbitrary metric spaces, see [4, Theorem 1.1.2].

Theorem 2.1. *If $\boldsymbol{\mu} : I \to \mathcal{P}(D)$ is an absolutely continuous curve, then the quantity*

$$|\dot{\boldsymbol{\mu}}|(t) := \lim_{h\to 0} \frac{W_2(\boldsymbol{\mu}(t+h),\boldsymbol{\mu}(t))}{|h|}$$

exists and is finite for a.e. $t \in I$. Moreover, $|\dot{\boldsymbol{\mu}}| \leqslant g$ a.e. on I for all g such that (2.2) *holds.*

In the Wasserstein space, absolutely continuous curves are related to solutions of the continuity equation: see [4, Chapter 8].

Theorem 2.2. *Let $\boldsymbol{\mu} : I \to \mathcal{P}(D)$ an absolutely continuous curve. If $(\mathbf{v}_t)_{t\in I}$ is a measurable family such that $\int_I \|\mathbf{v}_t\|_{L^2_{\boldsymbol{\mu}(t)}(D,\mathbf{R}^q)}\mathrm{d}t < +\infty$ and the equation $\partial_t\boldsymbol{\mu} + \nabla_D \cdot (\mathbf{v}\boldsymbol{\mu}) = 0$ is satisfied in a weak sense on $I \times D$ with no-flux boundary conditions on D, then one has*

$$|\dot{\boldsymbol{\mu}}|(t) \leqslant \sqrt{\int_D |\mathbf{v}_t|^2\mathrm{d}\boldsymbol{\mu}(t)} \tag{2.3}$$

for a.e. $t \in I$.

Moreover, there exists a unique (for a.e. $t \in I$, $\mathbf{v}_t$ is unique $\boldsymbol{\mu}(t)$ a.e.) family $(\mathbf{v}_t)_{t\in I}$ for which equality holds in (2.3) *for a.e. $t \in I$. This optimal family is characterized by the fact that for a.e. $t \in I$, there exists a sequence $(\psi_n)_{n\in\mathbf{N}}$ of elements of $C^1(D)$ such that $(\nabla\psi)_{n\in\mathbf{N}}$ converges to $\mathbf{v}_t$ in $L^2_{\boldsymbol{\mu}(t)}(D,\mathbf{R}^q)$.*

2.4. **Gradient flows.** At some point (in Section 6) we will need the notion of gradient flow in the Wassertsein space. Roughly speaking, if $F : \mathcal{P}(D) \to \mathbf{R} \cup \{+\infty\}$ is a given functional, a gradient flow is a curve $\boldsymbol{\mu} : [0, +\infty) \to \mathcal{P}(D)$ along which F decreases "the most" w.r.t. the Wasserstein distance, in a formal way it can be written

$$\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t}(t) = -\nabla F(\boldsymbol{\mu}(t)). \tag{2.4}$$

Of course nor the notion of gradient or of time derivative make sense as vectors in the Wassertsein space. In [4] (see also [27, Chapter 8]), it is shown how the notion of gradient flow can still be defined through the use of metric quantities only.

A standard assumption to ensure the existence and uniqueness of a gradient flow with a given value $\boldsymbol{\mu}(0)$ is that F is convex along generalized geodesic. If μ, ν and μ_0 are three probability measures on D, one can always build a transport plan $\pi \in \Pi(\mu_0, \mu, \nu) \subset \mathcal{P}(D \times D \times D)$ such that the 2-marginals are optimal transport plans between μ_0, μ on the one hand and μ_0, ν on the other hand (notice that in general the last 2-marginal is not an optimal plan between μ and ν). Then, the generalized geodesic $\boldsymbol{\mu} : [0, 1] \to \mathcal{P}(D)$ between μ and ν with base point μ_0 is defined as $\boldsymbol{\mu}(t) := f_t \# \pi$, with $f_t : (x, y, z) \in D^3 \mapsto (1 - t)y + tz \in D$. A functional $F : \mathcal{P}(D) \to \mathbf{R} \cup \{+\infty\}$ is said convex along generalized geodesics if for any points μ_0, μ and ν, there exists a generalized geodesic $\boldsymbol{\mu}$ joining μ to ν with base point μ_0 such that $F \circ \boldsymbol{\mu} : [0, 1] \to \mathbf{R} \cup \{+\infty\}$ is a convex function.

The only result that we will need is called the Evolution Variational Inequality (EVI) formulation of gradient flows (which is a way to make sense of (2.4) in the metric framework). It is summarized in the following theorem, whose proof can be found in [4, Theorem 11.2.1].

Theorem 2.3. *Let $F : \mathcal{P}(D) \to \mathbf{R} \cup \{+\infty\}$ a functional l.s.c. and convex along generalized geodesics. Then, for any $\mu \in \mathcal{P}(D)$ such that $F(\mu) < +\infty$, there exists an absolutely continuous curve $t \in [0, +\infty) \mapsto S_t^F \mu \in \mathcal{P}(D)$ such that $S_0^F \mu = \mu$ and for any $t \geqslant 0$ and any ν such that $F(\nu) < +\infty$*

$$\limsup_{h \to 0,\ h > 0} \frac{W_2^2(S_{t+h}^F \mu, \nu) - W_2^2(S_t^F \mu, \nu)}{2h} \leqslant F(\nu) - F(S_t^F \mu).$$

Moreover, the function $t \mapsto F\left(S_t^F \mu\right)$ is decreasing.

The curve $S^F \mu$ (which can be shown to be unique) is nothing else than the gradient flow of F starting form μ.

2.5. **Heat flows.** To regularize measures the main tool will be the heat flow. We recall in this subsection classical results that we will use in the sequel. We will denote by $\Phi^D : [0, +\infty) \times \mathcal{P}(D) \to \mathcal{P}(D)$ the heat flow with Neumann boundary conditions acting on D. For a proper definition, one can view Φ^D as the gradient flow of the Boltzmann entropy, which is convex along generalized geodesics because D is convex, see [4]. If $\mu \in \mathcal{P}(D)$ and $t > 0$, then $\Phi_t^D \mu \in \mathcal{P}(D)$ is defined as the measure $u(t, x)\mathrm{d}x$ with a density $u : (0, +\infty) \times D \to \mathbf{R}$ which is the solution of the Cauchy Problem

$$\begin{cases} \partial_s u(s, x) = \Delta u(s, x) & \text{if } (s, x) \in (0, +\infty) \times \mathring{D}, \\ \nabla u(s, x) \cdot \mathbf{n}_D(x) = 0 & \text{if } (s, x) \in (0, +\infty) \times \partial D, \\ \lim\limits_{s \to 0} [u(s, x)\mathrm{d}x] = \mu & \text{in } \mathcal{P}(D), \end{cases}$$

where $\mathbf{n}_D$ is the outward normal to D. Thanks to the gradient flow interpretation, we know that Φ^D is a contraction in the Wasserstein space [4, Theorem 11.2.1]: for any $t > 0$ and any $\mu, \nu \in \mathcal{P}(D)$, one has $W_2(\Phi_t^D \mu, \Phi_t^D \nu) \leqslant W_2(\mu, \nu)$.

As far as Ω is concerned, we will also use the heat flow but only through the heat kernel. We denote by $K^\Omega : (0, +\infty) \times \Omega \times \Omega \to \mathbf{R}$ the heat kernel associated to the Laplacian on Ω with Neumann

boundary conditions [6, Section 7]. For $u_0 \in L^1(\Omega)$, if we define for any $t \geqslant 0$

$$u(t,\xi) = (\Phi_t^\Omega u_0)(\xi) := \int_\Omega K_t^\Omega(\xi,\eta) u_0(\eta) \mathrm{d}\eta,$$

then u is the solution of the heat equation on Ω with Neumann boundary conditions such that $u(t,\cdot) \to u_0$ in $L^1(\Omega)$ as $t \to 0$. Notice, as a constant function is preserved by the heat flow, that the integral of $K_t^\Omega(\xi,\cdot)$ is 1 for a.e. $\xi \in \Omega$.

Useful properties of the heat flow and the heat kernel are summarized in the next proposition.

Proposition 2.4. *The heat flow Φ^D and the heat kernel K^Ω satisfy the following properties:*

(i) *For any $\mu \in \mathcal{P}(D)$ and any $t > 0$, the measure $\Phi_t^D \mu$ has a density w.r.t. $\mathcal{L}_D$ which is bounded from below by a strictly positive constant and belongs to $C^1(\mathring{D})$.*
For any $t > 0$, the heat kernel K_t^Ω is bounded from below by a strictly positive constant and belongs to $C^1(\mathring{\Omega} \times \mathring{\Omega})$.

(ii) *For any $t > 0$, the density of $\Phi_t^D \mu$ w.r.t. $\mathcal{L}_D$ is bounded in $L^\infty(D)$ independently of $\mu \in \mathcal{P}(D)$. There exists two constants $C_1 > 0$ and $C_2 > 0$ such that, for any $0 < t < 1$ and a.e. $\xi, \eta \in \Omega$,*

$$K_t^\Omega(\xi,\eta) \leqslant \frac{C_1}{t^{p/2}} \exp\left(-\frac{|\xi-\eta|^2}{C_2 t}\right).$$

(iii) *For a fixed $t > 0$ and for any $\mu \in \mathcal{P}(D)$ and $a \in C(D)$, one has*

$$\int_D a \left(\Phi_t^D \mu\right) = \int_D \left(\Phi_t^D a\right) \mathrm{d}\mu.$$

For a fixed $t > 0$ and all $\xi, \eta \in \mathring{\Omega}$, one has $K_t^\Omega(\xi,\eta) = K_t^\Omega(\eta,\xi)$.

Proof. Point (i) is standard interior parabolic regularity. Point (ii) comes from $L^\infty - L^1$ estimates and Gaussian estimates for the Neumann Laplacian, see [6, Section 7]. Point (iii) just states that the heat flow is self-adjoint. □

Let us collect some other results on heat kernel, which are not really difficult but which we did not find in standard textbooks.

Proposition 2.5. *There exists $\omega : \mathbf{R}_+ \to \mathbf{R}$, with $\lim_{t\to 0} \omega(t) = 0$, such that for every $\mu \in \mathcal{P}(D)$ and every $t \geqslant 0$, one has $W_2(\Phi_t^D \mu, \mu) \leqslant \omega(t)$.*

In other words, the convergence $\Phi_t^D \mu \to \mu$ as $t \to 0$ holds uniformly in $\mu \in \mathcal{P}(D)$.

Proof. Let us assume by contradiction that this is not the case. Then there exists $\varepsilon > 0$, a sequence $(t_n)_{n\in\mathbf{N}}$ which goes to 0, and a sequence $(\mu_n)_{n\in\mathbf{N}}$ of probability measures such that $W_2(\Phi_{t_n}^D \mu_n, \mu_n) \geqslant \varepsilon$. As $\mathcal{P}(D)$ is compact, up to extraction we can assume that μ_n converges to $\mu \in \mathcal{P}(D)$. By the triangular inequality,

$$\varepsilon \leqslant W_2(\Phi_{t_n}^D \mu_n, \mu_n) \leqslant W_2(\Phi_{t_n}^D \mu_n, \Phi_{t_n}^D \mu) + W_2(\Phi_{t_n}^D \mu, \mu) + W_2(\mu, \mu_n).$$

As we know that, for any t, Φ_t^D is a contraction in the Wasserstein space, one has $W_2(\Phi_{t_n}^D \mu_n, \Phi_{t_n}^D \mu) \leqslant W_2(\mu_n, \mu)$, thus the r.h.s. of the equation above goes to 0, which is a contradiction. □

Proposition 2.6. *Let $a \in C(\Omega \times \Omega)$ a positive continuous function which vanishes on the diagonal of $\Omega \times \Omega$. Then*

$$\lim_{t\to 0} \int_{\Omega\times\Omega} a(\xi,\eta) K_t^\Omega(\xi,\eta) \mathrm{d}\xi \mathrm{d}\eta = 0.$$

Proof. We fix $\varepsilon > 0$. By continuity of a and compactness of $\Omega \times \Omega$, we can find a δ such that, if $|\xi - \eta| \leqslant \delta$, then $a(\xi, \eta) \leqslant \varepsilon/2$. As the heat kernel is positive and its integral is 1, by cutting the integral in the parts where $|\xi - \eta| \leqslant \delta$ and those where $|\xi - \eta| \geqslant \delta$, we see that it is enough to show that

$$\int_{|\xi-\eta|\geqslant\delta} K_t^\Omega(\xi, \eta)\mathrm{d}\xi\mathrm{d}\eta \leqslant \varepsilon,$$

at least if t is small enough. But this is straightforward to get thanks to the Gaussian estimate, see (ii) in Proposition 2.4. □

3. The Dirichlet energy and the space $H^1(\Omega, \mathcal{P}(D))$

In this section, we define the Dirichlet energy of a function $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ following the idea of [9, Section 3]. We relate the space of $\boldsymbol{\mu}$ with finite Dirichlet energy with $H^1(\Omega, \mathcal{P}(D))$ using the theory of Sobolev spaces valued into metric spaces of Reshetnyak [25, 26], and we also prove that this Dirichlet energy coincides with the limit of ε-Dirichlet energies introduced by Korevaar, Schoen and Jost [19, 16].

Let us first define the space $L^2(\Omega, \mathcal{P}(D))$. As $\mathcal{P}(D)$ is bounded, it coincides with the measurable mappings valued in $\mathcal{P}(D)$.

Definition 3.1. *We denote by $L^2(\Omega, \mathcal{P}(D))$ the quotient space of measurable mappings $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ by the equivalence relation of being equal $\mathcal{L}_\Omega$-a.e. This space is endowed with the distance d_{L^2} defined by: for any $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ in $L^2(\Omega, \mathcal{P}(D))$,*

$$d_{L^2}^2(\boldsymbol{\mu}, \boldsymbol{\nu}) := \int_\Omega W_2^2(\boldsymbol{\mu}(\xi), \boldsymbol{\nu}(\xi))\mathrm{d}\xi.$$

If $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$, we can define a probability measure on $\Omega \times D$, that we will call temporary $\bar{\boldsymbol{\mu}}$, in the following way: for any $a \in C(\Omega \times D)$,

$$\int_{\Omega\times D} a\mathrm{d}\bar{\boldsymbol{\mu}} := \int_\Omega \left(\int_D a(\xi, \cdot)\mathrm{d}\boldsymbol{\mu}(\xi) \right) \mathrm{d}\xi. \tag{3.1}$$

If we take a function $a \in C(\Omega)$ which depends only on Ω, one can see that

$$\int_{\Omega\times D} a\mathrm{d}\bar{\boldsymbol{\mu}} = \int_\Omega a(\xi)\mathrm{d}\xi. \tag{3.2}$$

We will denote by $\mathcal{P}_0(\Omega \times D)$ the subspace of $\mathcal{P}(\Omega \times D)$ such that (3.2) is satisfied for all $a \in C(\Omega)$. Thanks to the disintegration Theorem [4, Theorem 5.3.1], one can see that, reciprocally, to each $\bar{\boldsymbol{\mu}} \in \mathcal{P}_0(\Omega \times D)$, one can associate a unique element $\boldsymbol{\mu}$ of $L^2(\Omega, \mathcal{P}(D))$ such that (3.1) holds. In all the sequel, we will drop the "bar" on $\bar{\boldsymbol{\mu}}$ and use the same letter $\boldsymbol{\mu}$ to denote an element of $L^2(\Omega, \mathcal{P}(D))$ and its counterpart in $\mathcal{P}_0(\Omega \times D)$ through the bijection that we have just described. Any $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ can be seen in two different ways: either as a mapping $\Omega \to \mathcal{P}(D)$, or as a probability measure on $\Omega \times D$, and we will very often switch between the two points of view. To clarify the notations:

- if $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$, then $\boldsymbol{\mu}(\xi)$ or $\boldsymbol{\mu}(\xi, \mathrm{d}x)$, which is an element of $\mathcal{P}(D)$, will denote the function $\boldsymbol{\mu}$ evaluated at ξ;
- $\boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x)$ will indicate that we consider $\boldsymbol{\mu}$ as an element of $\mathcal{P}_0(\Omega \times D)$, integration on $\Omega \times D$ will be denoted by $\mathrm{d}\boldsymbol{\mu}$ or $\boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x)$, notice that we have the following relation: $\boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x) = \boldsymbol{\mu}(\xi, \mathrm{d}x)\mathrm{d}\xi$;
- the mapping $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ is said continuous (resp. Lipschitz) if there is one representative of $\boldsymbol{\mu}$ such that $W_2(\boldsymbol{\mu}(\xi), \boldsymbol{\mu}(\eta))$ goes to 0 if $\eta \to \xi$ (resp. is bounded by $C|\xi - \eta|$ for some $C < +\infty$).

The topologies on $L^2(\Omega, \mathcal{P}(D))$ are defined as follows.

Definition 3.2. *The strong topology on $L^2(\Omega, \mathcal{P}(D))$ is the one induced by the distance d_{L^2}, and the weak topology is the one induced on $\mathcal{P}_0(\Omega \times D)$ by the weak topology on $\mathcal{P}(\Omega \times D)$.*

Proposition 3.3. *W.r.t. the strong topology, $L^2(\Omega, \mathcal{P}(D))$ is a polish space. W.r.t. the weak topology, $L^2(\Omega, \mathcal{P}(D))$ is a separable compact space. Moreover, the strong topology is finer than the weak topology.*

Proof. The statement concerning the strong topology is a consequence of the fact that $\mathcal{P}(D)$ is itself a polish space, see for instance [19, Section 1.1]. As $\mathcal{P}_0(\Omega \times D)$ is closed in $\mathcal{P}(\Omega \times D)$, for the second statement we simply use the fact that $\mathcal{P}(\Omega \times D)$ is itself a separable compact space.

To compare the topologies we take a sequence $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ which converges strongly to some $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$. Up to extraction, we know that we can assume that $\boldsymbol{\mu}_n(\xi)$ converges in $\mathcal{P}(D)$ to $\boldsymbol{\mu}(\xi)$ for a.e. $\xi \in \Omega$. In particular, if $a \in C(\Omega \times D)$, we have that $\int_D a(\xi, \cdot)\mathrm{d}\boldsymbol{\mu}_n(\xi)$ converges to $\int_D a(\xi, \cdot)\mathrm{d}\boldsymbol{\mu}(\xi)$ for a.e. $\xi \in \Omega$. With the help of Lebesgue dominated convergence Theorem, we see that

$$\lim_{n\to+\infty} \int_{\Omega\times D} a\mathrm{d}\boldsymbol{\mu}_n = \lim_{n\to+\infty} \int_\Omega \left(\int_D a(\xi, \cdot)\mathrm{d}\boldsymbol{\mu}_n(\xi) \right) \mathrm{d}\xi = \int_\Omega \left(\int_D a(\xi, \cdot)\mathrm{d}\boldsymbol{\mu}(\xi) \right) \mathrm{d}\xi = \int_{\Omega\times D} a\mathrm{d}\boldsymbol{\mu}.$$

As a is arbitrary, this allows us to conclude that $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ converges to $\boldsymbol{\mu}$ for the weak topology. □

3.1. **A Benamou-Brenier type definition.** We are now ready to define the Dirichlet energy. The first step is to define the (generalized) continuity equation. Recall that $C^1_c(\mathring{\Omega} \times D, \mathbf{R}^p)$ is the set of C^1 functions defined on $\Omega \times D$ and valued in $\mathbf{R}^p$, whose support is compactly included in $\mathring{\Omega}$, but not necessarily in D, and $\mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ denotes the space of vector-valued measures on $\Omega \times D$ with finite mass.

Definition 3.4. *If $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ and if $E \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$, we say that the pair $(\boldsymbol{\mu}, E)$ satisfies the continuity equation if, for every $\varphi \in C^1_c(\mathring{\Omega} \times D, \mathbf{R}^p)$, one has*

$$\int_{\Omega\times D} \nabla_\Omega \cdot \varphi \mathrm{d}\boldsymbol{\mu} + \int_{\Omega\times D} \nabla_D \varphi \cdot \mathrm{d}E = 0.$$

In other words, the pair $(\boldsymbol{\mu}, E)$ satisfies the continuity equation if the equation

$$\nabla_\Omega \boldsymbol{\mu} + \nabla_D \cdot E = 0.$$

with no-flux boundary conditions on ∂D is satisfied in a weak sense. In we develop in coordinates, it means that for every $\alpha \in \{1, 2, \dots, p\}$, one has $\partial_\alpha \boldsymbol{\mu} + \sum_{i=1}^q \partial_i E^{i\alpha} = 0$. If the pair $(\boldsymbol{\mu}, E)$ satisfies the continuity equation, we want to define its Dirichlet energy by $\int_{\Omega\times D} \frac{|E|^2}{2\boldsymbol{\mu}}$. It is well known in optimal transport that this definition can be made by duality.

Definition 3.5. *If $(\boldsymbol{\mu}, E)$ satisfies the continuity equation, we define its Dirichlet energy $\mathrm{Dir}(\boldsymbol{\mu}, E)$ by*

$$\mathrm{Dir}(\boldsymbol{\mu}, E) := \sup \left\{ \int_{\Omega\times D} a\mathrm{d}\boldsymbol{\mu} + \int_{\Omega\times D} b \cdot \mathrm{d}E \ : \ (a, b) \in C(\Omega \times D, \mathcal{K}) \right\},$$

where $\mathcal{K} \subset \mathbf{R}^{1+pq}$ is the set of pair (x, y) with $x \in \mathbf{R}$ and $y \in \mathbf{R}^{pq}$ such that $x + \frac{1}{2}|y|^2 \leqslant 0$.

The following proposition is identical to the case of the Benamou-Brenier formula.

Proposition 3.6. *If $(\boldsymbol{\mu}, E)$ satisfies the continuity equation and $\mathrm{Dir}(\boldsymbol{\mu}, E) < +\infty$, then E is absolutely continuous w.r.t. $\boldsymbol{\mu}$, and if $\mathbf{v} : \Omega \times D \to \mathbf{R}^{pq}$ is the density of E w.r.t. $\boldsymbol{\mu}$, then one has*

$$\mathrm{Dir}(\boldsymbol{\mu}, E) = \mathrm{Dir}(\boldsymbol{\mu}, \mathbf{v}\boldsymbol{\mu}) = \int_{\Omega\times D} \frac{1}{2}|\mathbf{v}|^2 \mathrm{d}\boldsymbol{\mu}.$$

Proof. There is nothing to add to the proof of this when Ω is 1-dimensional, and such a proof can be found for instance in [27, Proposition 5.18]. □

Definition 3.7. *Let $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$. Its Dirichlet energy $\mathrm{Dir}(\boldsymbol{\mu})$ is defined by*

$$\mathrm{Dir}(\boldsymbol{\mu}) := \inf \left\{ \mathrm{Dir}(\boldsymbol{\mu}, E) \ : \ E \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq}) \text{ and } (\boldsymbol{\mu}, E) \text{ satisfies the continuity equation} \right\}.$$

Let us underline that if there exists no $E \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ such that $(\boldsymbol{\mu}, E)$ satisfies the continuity equation, then by convention $\mathrm{Dir}(\boldsymbol{\mu}) = +\infty$. To be sure that it is written somewhere, let us state the following proposition which identifies the Dirichlet energy if Ω is a segment of $\mathbf{R}$. It is a consequence of Theorem 2.2 and of the above definitions (see [27, Theorem 5.28]).

Proposition 3.8. *Assume that I is a segment of $\mathbf{R}$ and let $\boldsymbol{\mu} \in L^2(I, \mathcal{P}(D))$. Then $\mathrm{Dir}(\boldsymbol{\mu}) < +\infty$ if and only if $\boldsymbol{\mu}$ is absolutely continuous, and in this case*

$$\mathrm{Dir}(\boldsymbol{\mu}) = \int_I \frac{1}{2} |\dot{\boldsymbol{\mu}}|^2(t) \mathrm{d}t.$$

Now, let us show easy properties, namely that the Dirichlet energy behaves like the square of a H^1 norm: it is convex and l.s.c.

Proposition 3.9. *If $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ is such that $\mathrm{Dir}(\boldsymbol{\mu}) < +\infty$, then there exists a unique $E \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ such that $(\boldsymbol{\mu}, E)$ satisfies the continuity equation and $\mathrm{Dir}(\boldsymbol{\mu}) = \mathrm{Dir}(\boldsymbol{\mu}, E)$.*

Definition 3.10. *If $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ and if $E = \mathbf{v}\boldsymbol{\mu}$ is such that $(\boldsymbol{\mu}, E)$ satisfies the continuity equation and $\mathrm{Dir}(\boldsymbol{\mu}) = \mathrm{Dir}(\boldsymbol{\mu}, E) < +\infty$, then E and $\mathbf{v}$ are said* tangent *to $\boldsymbol{\mu}$.*

The terminology *tangent* comes from [4]. As in the case of absolutely continuous curves, there is a characterization of the tangent velocity field $\mathbf{v}$ which looks like [4, Proposition 8.4.5]: see Corollary 3.21 proved later.

Proposition 3.11. *The mapping $\mathrm{Dir} : L^2(\Omega, \mathcal{P}(D)) \to \mathbf{R} \cup \{+\infty\}$ is l.s.c. w.r.t. weak convergence. Moreover it is convex: for any $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ in $L^2(\Omega, \mathcal{P}(D))$ and any $t \in [0,1]$,*

$$\mathrm{Dir}((1-t)\boldsymbol{\mu} + t\boldsymbol{\nu}) \leqslant (1-t)\mathrm{Dir}(\boldsymbol{\mu}) + t\mathrm{Dir}(\boldsymbol{\nu}).$$

Proof of Proposition 3.9. To this extent, we can use the direct method of calculus of variations. Take $(E_n)_{n \in \mathbf{N}}$ a minimizing sequence, i.e. such that $(\boldsymbol{\mu}, E_n)$ satisfies the continuity equation for every $n \in \mathbf{N}$ and such that $\mathrm{Dir}(\boldsymbol{\mu}, E_n)$ converges to $\mathrm{Dir}(\boldsymbol{\mu})$ when $n \to +\infty$. For n large enough, $\mathrm{Dir}(\boldsymbol{\mu}, E_n)$ is finite, thus $E_n = \mathbf{v}_n \boldsymbol{\mu}$ and $\mathrm{Dir}(\boldsymbol{\mu}, E_n) = \frac{1}{2} \|\mathbf{v}_n\|^2_{L^2_{\boldsymbol{\mu}}(\Omega \times D, \mathbf{R}^{pq})}$. In particular, $\mathbf{v}_n$ is uniformly bounded in $L^2_{\boldsymbol{\mu}}(\Omega \times D, \mathbf{R}^{pq})$, up to extraction it converges weakly to some $\mathbf{v} \in L^2_{\boldsymbol{\mu}}(\Omega \times D, \mathbf{R}^{pq})$. By l.s.c. of the norm, we see that $\frac{1}{2} \|\mathbf{v}\|^2_{L^2_{\boldsymbol{\mu}}(\Omega \times D, \mathbf{R}^{pq})} \leqslant \mathrm{Dir}(\boldsymbol{\mu})$. Moreover, the weak convergence of $\mathbf{v}_n$ to $\mathbf{v}$ allows us to see that $(\boldsymbol{\mu}, \mathbf{v}\boldsymbol{\mu})$ satisfies the continuity equation.

To prove uniqueness, assume that E_1 and E_2 are such that $\mathrm{Dir}(\boldsymbol{\mu}) = \mathrm{Dir}(\boldsymbol{\mu}, E_1) = \mathrm{Dir}(\boldsymbol{\mu}, E_2)$. We write $E_1 = \mathbf{v}_1 \boldsymbol{\mu}$ and $E_2 = \mathbf{v}_2 \boldsymbol{\mu}$. Notice that $(\boldsymbol{\mu}, (E_1 + E_2)/2)$ satisfies the continuity equation, hence by convexity of the square,

$$\begin{aligned} \mathrm{Dir}(\boldsymbol{\mu}) \leqslant \mathrm{Dir}\left(\boldsymbol{\mu}, \frac{E_1 + E_2}{2}\right) &= \int_{\Omega \times D} \frac{1}{2} \left| \frac{\mathbf{v}_1 + \mathbf{v}_2}{2} \right|^2 \mathrm{d}\boldsymbol{\mu} \\ &\leqslant \frac{1}{2} \int_{\Omega \times D} \frac{1}{2} |\mathbf{v}_1|^2 \mathrm{d}\boldsymbol{\mu} + \frac{1}{2} \int_{\Omega \times D} \frac{1}{2} |\mathbf{v}_2|^2 \mathrm{d}\boldsymbol{\mu} \\ &= \frac{1}{2}(\mathrm{Dir}(\boldsymbol{\mu}, E_1) + \mathrm{Dir}(\boldsymbol{\mu}, E_2)) = \mathrm{Dir}(\boldsymbol{\mu}). \end{aligned}$$

In consequence all inequalities are equalities, thus $\mathbf{v}_1 = \mathbf{v}_2$ $\boldsymbol{\mu}$-a.e., which means $E_1 = E_2$. □

Proof of Proposition 3.11. Let us first prove the lower semi-continuity. Let $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ a sequence which converges weakly to some $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$. Without loss of generality, we can assume that $\liminf_{n\in\mathbf{N}} \mathrm{Dir}(\boldsymbol{\mu}_n) < +\infty$. Up to extraction, we can assume that $\mathrm{Dir}(\boldsymbol{\mu}_n)$ converges to $\liminf_n \mathrm{Dir}(\boldsymbol{\mu}_n)$. For each $n \in \mathbf{N}$, let $E_n = \mathbf{v}_n\boldsymbol{\mu}_n \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ be tangent to $\boldsymbol{\mu}_n$. As

$$|E_n|(\Omega \times D) = \int_{\Omega\times D} |\mathbf{v}_n| \mathrm{d}\boldsymbol{\mu}_n \leqslant \sqrt{\int |\mathbf{v}_n|^2 \mathrm{d}\boldsymbol{\mu}_n} = \sqrt{2\mathrm{Dir}(\boldsymbol{\mu}_n)},$$

we see that the mass of E_n is uniformly bounded. Up to extraction, we can assume that E_n weakly converges in $\mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ to some E. The weak convergence allows to pass to the limit in the continuity equation, thus $(\boldsymbol{\mu}, E)$ satisfies the continuity equation. Moreover, as $\mathrm{Dir} : L^2(\Omega, \mathcal{P}(D)) \times \mathcal{M}(\Omega \times D, \mathbf{R}^{pq}) \to \mathbf{R} \cup \{+\infty\}$ is a supremum of continuous (for the weak convergence) functionals, it is clear that $\mathrm{Dir}(\boldsymbol{\mu}, E) \leqslant \liminf_n \mathrm{Dir}(\boldsymbol{\mu}_n, E_n)$, which allows to conclude that $\mathrm{Dir}(\boldsymbol{\mu}) \leqslant \liminf_n \mathrm{Dir}(\boldsymbol{\mu}_n)$.

The convexity is simple: if E and F are tangent to respectively $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$, then one sees that $((1-t)\boldsymbol{\mu} + t\boldsymbol{\nu}, (1-t)E + tF)$ satisfies the continuity equation. Moreover, as $\mathrm{Dir} : L^2(\Omega, \mathcal{P}(D)) \times \mathcal{M}(\Omega \times D, \mathbf{R}^{pq}) \to \mathbf{R} \cup \{+\infty\}$ is a supremum of linear functionals, it is a convex function of its two arguments, thus

$$\mathrm{Dir}((1-t)\boldsymbol{\mu} + t\boldsymbol{\nu}) \leqslant \mathrm{Dir}((1-t)\boldsymbol{\mu} + t\boldsymbol{\nu}, (1-t)E + tF) \leqslant (1-t)\mathrm{Dir}(\boldsymbol{\mu}, E) + t\mathrm{Dir}(\boldsymbol{\nu}, F),$$

which is enough to conclude. □

Let us show a localization property: the tangent velocity field $\mathbf{v}$, depends only locally on the values of $\boldsymbol{\mu}$. In other words, if $\mathbf{v}$ is tangent on Ω, then it is tangent on every subset of Ω. In the next proposition, $\boldsymbol{\mu}|_{\tilde{\Omega}}$ and $\mathbf{v}|_{\tilde{\Omega}}$ will denote the restrictions of $\boldsymbol{\mu}$ and $\mathbf{v}$ to a subset $\tilde{\Omega}$ of Ω.

Proposition 3.12. *Let $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ such that $\mathrm{Dir}(\boldsymbol{\mu}) < +\infty$ and let $\mathbf{v} \in L^2_{\boldsymbol{\mu}}(\Omega\times D, \mathbf{R}^{pq})$ be tangent to $\boldsymbol{\mu}$. Then, if $\tilde{\Omega}$ is any subdomain compactly supported in $\mathring{\Omega}$, $\mathbf{v}|_{\tilde{\Omega}}$ is tangent to $\boldsymbol{\mu}|_{\tilde{\Omega}}$.*

Proof. Set $E = \mathbf{v}\boldsymbol{\mu}$ and let $\tilde{E} = \tilde{\mathbf{v}}\boldsymbol{\mu} \in \mathcal{P}(\tilde{\Omega} \times D)$ be the tangent momentum to $\boldsymbol{\mu}|_{\tilde{\Omega}}$. Take $\chi \in C^1(\tilde{\Omega})$ a smooth function compactly supported in $\tilde{\Omega}$. As χ does not depend on D, $(\boldsymbol{\mu}, (1-\chi)E + \chi\tilde{E})$ still satisfies the continuity equation. By convexity of the square and optimality of $\tilde{\mathbf{v}}$,

$$\begin{aligned}
\mathrm{Dir}(\boldsymbol{\mu}) \leqslant \mathrm{Dir}(\boldsymbol{\mu}, (1-\chi)E + \chi\tilde{E}) &= \int_{\Omega\times D} \frac{1}{2}|(1-\chi)\mathbf{v} + \chi\tilde{\mathbf{v}}|^2 \mathrm{d}\boldsymbol{\mu} \\
&\leqslant \int_{\Omega\times D} (1-\chi)\frac{1}{2}|\mathbf{v}|^2 \mathrm{d}\boldsymbol{\mu} + \int_{\tilde{\Omega}\times D} \frac{1}{2}\chi|\tilde{\mathbf{v}}|^2 \mathrm{d}\boldsymbol{\mu} \\
&\leqslant \int_{\Omega\times D} (1-\chi)\frac{1}{2}|\mathbf{v}|^2 \mathrm{d}\boldsymbol{\mu} + \int_{\tilde{\Omega}\times D} \chi\frac{1}{2}|\mathbf{v}|^2 \mathrm{d}\boldsymbol{\mu} = \mathrm{Dir}(\boldsymbol{\mu}).
\end{aligned}$$

Hence $(1-\chi)E + \chi\tilde{E}$ is also tangent to $\boldsymbol{\mu}$. By uniqueness of the tangent velocity field we deduce that $\mathbf{v}|_{\tilde{\Omega}} = \tilde{\mathbf{v}}$ at least on the support of χ. Arbitrariness of χ allows us to conclude. □

We will conclude this subsection by showing the following approximation result, which basically states that smooth functions are dense in $L^2(\Omega, \mathcal{P}(D))$.

Theorem 3.13. *Let $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$. Then there exists a sequence $\boldsymbol{\mu}_n \in L^2(\Omega, \mathcal{P}(D))$ with the following properties:*

- (i) *For any $n \in \mathbf{N}$, $\boldsymbol{\mu}_n(\mathrm{d}\xi, \mathrm{d}x) = \rho_n(\xi, x)\mathrm{d}\xi\mathrm{d}x$, where the density ρ_n of $\boldsymbol{\mu}_n$ w.r.t. to $\mathcal{L}_\Omega \otimes \mathcal{L}_D$ satisfies $\rho_n \in C^1(\mathring{\Omega} \times \mathring{D}) \cap L^\infty(\Omega \times D)$ and $\operatorname{ess\,inf}_{\Omega\times D} \rho_n > 0$.*
- (ii) *The sequence $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ converges strongly and a.e. to $\boldsymbol{\mu}$ in $L^2(\Omega, \mathcal{P}(D))$.*

(iii) *Moreover, if* $E \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ *is such that* $(\boldsymbol{\mu}, E)$ *satisfies the continuity equation and* $\mathrm{Dir}(\boldsymbol{\mu}, E) < +\infty$*, then we can choose* $\mathbf{v}_n \in C^1(\mathring{\Omega} \times \mathring{D}, \mathbf{R}^{pq}) \cap L^\infty(\Omega \times D, \mathbf{R}^{pq})$ *a smooth vector field such that* $(\boldsymbol{\mu}_n, \mathbf{v}_n\boldsymbol{\mu}_n)$ *satisfies the continuity equation and* $\mathrm{Dir}(\boldsymbol{\mu}_n, \mathbf{v}_n\boldsymbol{\mu}_n) \leqslant \mathrm{Dir}(\boldsymbol{\mu}, E)$ *for all* $n \in \mathbf{N}$*, and such that* $(\mathbf{v}_n\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ *converges weakly to* E*. In particular,* $\mathrm{Dir}(\boldsymbol{\mu}_n)$ *converges to* $\mathrm{Dir}(\boldsymbol{\mu})$ *as* $n \to +\infty$*.*

Proof. We will rely on the heat flow to regularize $\boldsymbol{\mu}$, and we will make it act both on Ω and D. Recall that Φ^Ω, Φ^D denote the heat flows, and K^Ω denotes the heat kernel. For any $t, s > 0$, we set $\boldsymbol{\mu}_{t,s} := \Phi^\Omega_t \Phi^D_s \boldsymbol{\mu}$, which is a shortcut to say that we define

$$\boldsymbol{\mu}_{t,s}(\xi) := \int_\Omega K^\Omega_t(\xi, \eta)[\Phi^D_s \boldsymbol{\mu}(\eta)]\mathrm{d}\eta.$$

(the integral is performed according to the affine structure on $\mathcal{P}(D)$, see (2.1)). By parabolic regularity, it is clear that for fixed $t, s > 0$, the measure $\boldsymbol{\mu}_{t,s}$ has a density w.r.t. $\mathcal{L}_\Omega \otimes \mathcal{L}_D$ which is smooth on $\mathring{\Omega} \times \mathring{D}$ and bounded from below and from above by strictly positive constants on $\Omega \times D$. Then we define $\boldsymbol{\mu}_n := \boldsymbol{\mu}_{t_n,s_n}$ for sequences $(t_n)_{n\in\mathbf{N}}$ and $(s_n)_{n\in\mathbf{N}}$ which goes to 0. In other words, point (i) is proved.

Let us prove (ii). It would be easy to get that $\boldsymbol{\mu}_{t_n,s_n}$ converges weakly to $\boldsymbol{\mu}$ as $n \to +\infty$, but we want strong convergence. Starting from the definition of $\boldsymbol{\mu}_{t,s}(\xi)$ and using Jensen's inequality and the triangle inequality, one sees that for a.e. $\xi \in \Omega$,

$$\begin{aligned} W_2^2(\boldsymbol{\mu}_{t,s}(\xi), \boldsymbol{\mu}(\xi)) &\leqslant \int_\Omega K^\Omega_t(\xi, \eta) W_2^2(\Phi^D_s \boldsymbol{\mu}(\eta), \boldsymbol{\mu}(\xi))\mathrm{d}\eta \\ &\leqslant 2\left[\int_\Omega K^\Omega_t(\xi, \eta) W_2^2(\Phi^D_s \boldsymbol{\mu}(\eta), \boldsymbol{\mu}(\eta))\mathrm{d}\eta + \int_\Omega K^\Omega_t(\xi, \eta) W_2^2(\boldsymbol{\mu}(\eta), \boldsymbol{\mu}(\xi))\mathrm{d}\eta\right] \\ &\leqslant 2\omega(s)^2 + 2\int_\Omega K^\Omega_t(\xi, \eta) W_2^2(\boldsymbol{\mu}(\eta), \boldsymbol{\mu}(\xi))\mathrm{d}\eta, \end{aligned}$$

where we recall that ω is defined in Proposition 2.5. Integrating w.r.t. ξ, we get

$$d^2_{L^2}(\boldsymbol{\mu}_{t_n,s_n}, \boldsymbol{\mu}) \leqslant 2\omega(s_n)^2 + 2\int_{\Omega\times\Omega} K^\Omega_{t_n}(\xi, \eta) W_2^2(\boldsymbol{\mu}(\eta), \boldsymbol{\mu}(\xi))\mathrm{d}\xi\mathrm{d}\eta.$$

We have to estimate the r.h.s. of the equation above. Recall that $\boldsymbol{\mu}$ is not assumed to have a finite Dirichlet energy, thus we have no regularity w.r.t. Ω. Fix $\delta > 0$. Thanks to Lusin's theorem (for Lusin's theorem in metric spaces, see [27, Box 1.6]), there exists a compact $X \subset \Omega$ such that $\mathcal{L}_\Omega(\Omega\backslash X) \leqslant \delta$ and the restriction of $\boldsymbol{\mu}$ to X is continuous. In particular, the function $a : X \times X \to \mathbf{R}$ defined by $a(\xi, \eta) := W_2^2(\boldsymbol{\mu}(\eta), \boldsymbol{\mu}(\xi))$ is continuous, and it stays continuous if we extend a by setting $a(\xi, \xi) = 0$ for every $\xi \in \Omega$. The function a being real-valued, it can be extended to the whole $\Omega \times \Omega$ in a positive continuous function still denoted by a. By construction, a vanishes on the diagonal of $\Omega \times \Omega$. If C is the diameter of $\mathcal{P}(D)$, then (to handle the first integral of the r.h.s. we use the fact that the integral of $K^\Omega_{t_n}$ w.r.t. its first variable is 1)

$$\begin{aligned} \int_{\Omega\times\Omega} K^\Omega_{t_n}(\xi, \eta) W_2^2(\boldsymbol{\mu}(\eta), \boldsymbol{\mu}(\xi))\mathrm{d}\xi\mathrm{d}\eta &\leqslant 2\int_{\Omega\times(\Omega\backslash X)} C^2 K^\Omega_{t_n}(\xi, \eta)\mathrm{d}\xi\mathrm{d}\eta + \int_{X\times X} K^\Omega_{t_n}(\xi, \eta) a(\xi, \eta)\mathrm{d}\xi\mathrm{d}\eta \\ &\leqslant 4C^2\delta + \int_{\Omega\times\Omega} K^\Omega_{t_n}(\xi, \eta) a(\xi, \eta)\mathrm{d}\xi\mathrm{d}\eta. \end{aligned}$$

Here we can use Proposition 2.6 to conclude that $\int_{\Omega\times\Omega} K^\Omega_{t_n}(\xi, \eta) a(\xi, \eta)\mathrm{d}\xi\mathrm{d}\eta$ goes to 0 as $n \to +\infty$. Putting all the pieces together, we conclude that

$$\limsup_{n\to+\infty} d^2_{L^2}(\boldsymbol{\mu}_{t_n,s_n}, \boldsymbol{\mu}) \leqslant 4C^2\delta.$$

As δ is arbitrary, $\boldsymbol{\mu}_{t_n,s_n}$ converges strongly to $\boldsymbol{\mu}$. Up to extraction, the convergence takes also place a.e., which means that for a.e. $\xi \in \Omega$, $\boldsymbol{\mu}_{t_n,s_n}(\xi)$ converges to $\boldsymbol{\mu}(\xi)$ in $\mathcal{P}(D)$. Thus (ii) is proved.

It remains to prove (iii), let us assume from now that $E \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ is such that $(\boldsymbol{\mu}, E)$ satisfies the continuity equation and $\mathrm{Dir}(\boldsymbol{\mu}, E) < +\infty$. Take $\mathbf{v} \in L^2_{\boldsymbol{\mu}}(\Omega \times D)$ such that $E = \mathbf{v}\boldsymbol{\mu}$. We define $E_n := \Phi^{\Omega}_{t_n}\Phi^{D}_{s_n}E$, which is a shortcut to say that

$$E_n(\xi, x) := \int_{\Omega\times D} K^{\Omega}_{t_n}(\xi,\eta)\left[\Phi^{D}_{s_n}\left(\mathbf{v}(\eta,\cdot)\boldsymbol{\mu}(\eta)\right)\right]\mathrm{d}\eta.$$

Call e_n the density of E_n w.r.t. $\mathcal{L}_\Omega \otimes \mathcal{L}_D$, we simply set $\mathbf{v}_n = E_n/\rho_n$ which is well defined as $\rho_n > 0$ a.e. on $\Omega \times D$. In other words, $E_n = \mathbf{v}_n\boldsymbol{\mu}_n$. By parabolic regularity, one has $\mathbf{v}_n \in C^1(\mathring{\Omega} \times \mathring{D}, \mathbf{R}^{pq}) \cap L^\infty(\Omega \times D, \mathbf{R}^{pq})$. It is clear that $(E_n)_{n\in\mathbf{N}}$ converges weakly to E. As the heat flow commutes with derivatives w.r.t. Ω or D, the pair $(\boldsymbol{\mu}_n, E_n)$ satisfies the continuity equation for any $n \in \mathbf{N}$. Moreover for every $(a, b) \in C(\Omega \times D, \mathcal{K})$ (cf. Definition 3.5), one has $(\Phi^{\Omega}_{t_n}\Phi^{D}_{s_n}a, \Phi^{\Omega}_{t_n}\Phi^{D}_{s_n}b) \in C(\Omega \times D, \mathcal{K})$: indeed, as $\mathcal{K}$ is convex, by the maximum principle $C(\Omega \times D, \mathcal{K})$ is stable by the heat flow. Hence, as the heat flow is self-adjoint and by definition of $\mathrm{Dir}(\boldsymbol{\mu}, E)$,

$$\int_{\Omega\times D} a\mathrm{d}\boldsymbol{\mu}_n + \int_{\Omega\times D} b\cdot \mathrm{d}E_n = \int_{\Omega\times D}(\Phi^{\Omega}_{t_n}\Phi^{D}_{s_n}a)\mathrm{d}\boldsymbol{\mu} + \int_{\Omega\times D}(\Phi^{\Omega}_{t_n}\Phi^{D}_{s_n}b)\cdot \mathrm{d}E \leqslant \mathrm{Dir}(\boldsymbol{\mu}, E).$$

As a and b are arbitrary, $\mathrm{Dir}(\boldsymbol{\mu}_n, E_n) \leqslant \mathrm{Dir}(\boldsymbol{\mu}, E)$ for all $n \in \mathbf{N}$. If we take for E the momentum tangent to $\boldsymbol{\mu}$ then

$$\limsup_{n\to+\infty}\mathrm{Dir}(\boldsymbol{\mu}_n) \leqslant \limsup_{n\to+\infty}\mathrm{Dir}(\boldsymbol{\mu}_n, \mathbf{v}_n\boldsymbol{\mu}_n) \leqslant \mathrm{Dir}(\boldsymbol{\mu}, E) = \mathrm{Dir}(\boldsymbol{\mu}) \leqslant \liminf_{n\to+\infty}\mathrm{Dir}(\boldsymbol{\mu}_n),$$

where the last inequality comes from the lower semi-continuity of Dir. Hence $\mathrm{Dir}(\boldsymbol{\mu}_n)$ converges to $\mathrm{Dir}(\boldsymbol{\mu})$ as $n \to +\infty$. □

3.2. **The smooth case.** In this subsection, we will briefly study the smooth case, i.e. the one where $\boldsymbol{\mu}$ has a smooth and strictly positive density w.r.t. $\mathcal{L}_\Omega \otimes \mathcal{L}_D$. It will help to understand the meaning of the continuity equation and we will use it in the sequel when reasoning by approximation.

We assume that $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ can be written as $\boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x) = \rho(\xi, x)\mathrm{d}\xi\mathrm{d}x$, where ρ is C^1 on $\mathring{\Omega}\times\mathring{D}$. Let $\mathbf{v} \in C^1(\mathring{\Omega} \times \mathring{D}, \mathbf{R}^{pq}) \cap L^\infty(\Omega \times D, \mathbf{R}^{pq})$ be a smooth vector field. If $(\boldsymbol{\mu}, \mathbf{v}\boldsymbol{\mu})$ satisfies the continuity equation, it means that they also satisfy in a strong sense:

$$\begin{cases} \nabla_\Omega\rho + \nabla_D\cdot(\mathbf{v}\rho) = 0 & \text{in } \mathring{\Omega}\times\mathring{D} \\ \mathbf{v}\cdot\mathbf{n_D} = 0 & \text{on } \mathring{\Omega}\times\partial D, \end{cases}$$

where $\mathbf{n}_D$ is the outward normal at D. In particular, if $\gamma : I \to \mathring{\Omega}$ is a smooth curve going from an interval of $\mathbf{R}$ to $\mathring{\Omega}$, then $\boldsymbol{\mu}^\gamma = \boldsymbol{\mu}\circ\gamma : I \to \mathcal{P}(D)$ an defines absolutely continuous curve valued in the Wasserstein space for which we know one velocity field. Indeed, we know that the (classical) continuity equation $\partial_t\boldsymbol{\mu}^\gamma + \nabla\cdot(\mathbf{v}^\gamma\boldsymbol{\mu}^\gamma)$ is satisfied in a strong sense (with Neumann boundary conditions on D) provided that we define $\mathbf{v}^\gamma := \mathbf{v}\cdot\dot\gamma : I \times D \to \mathbf{R}^q$. More precisely, if $i \in \{1, 2, \ldots, q\}$, the i-the component of $\mathbf{v}^\gamma$ at time $t \in I$ and at the point $x \in D$ is

$$(\mathbf{v}^\gamma(t, x))^i = \sum_{\alpha=1}^{p}\mathbf{v}(\gamma(t), x)^{i\alpha}\dot\gamma^\alpha(t).$$

In other words, the (generalized) continuity equation implies that we get (classical) continuity equation for every curve of Ω. In some sense, the (generalized) continuity equation is much stronger in higher dimensions. As we recalled previously, the velocity field $\mathbf{v}^\gamma$ is related to the metric derivative

of the curve $\boldsymbol{\mu}^\gamma$ in the Wasserstein space . Applying Theorem 2.2, we see that for any $s<t$ elements of I,

$$W_2(\boldsymbol{\mu}^\gamma(t),\boldsymbol{\mu}^\gamma(s)) \leqslant \int_s^t |\dot{\boldsymbol{\mu}}^\gamma(r)|\mathrm{d}r \leqslant \int_s^t \sqrt{\int_D |\mathbf{v}^\gamma(r,x)|^2\rho(\gamma(r),x)\mathrm{d}x}\mathrm{d}r. \tag{3.3}$$

With this key estimate, we can prove the following.

Proposition 3.14. *Let $\boldsymbol{\mu}\in L^2(\Omega,\mathcal{P}(D))$ be smooth: it admits a density ρ w.r.t. $\mathcal{L}_\Omega\otimes\mathcal{L}_D$ satisfying $\rho\in C^1(\mathring{\Omega}\times\mathring{D})\cap L^\infty(\Omega\times D)$. Suppose that $\mathbf{v}\in C^1(\mathring{\Omega}\times\mathring{D},\mathbf{R}^{pq})\cap L^\infty(\Omega\times D,\mathbf{R}^{pq})$ is such that $(\boldsymbol{\mu},\mathbf{v}\boldsymbol{\mu})$ satisfies the continuity equation.*

Then the function $\boldsymbol{\mu}$ is Lipschitz. Moreover, if $\xi\in\mathring{\Omega}$ and $e\in\mathbf{R}^p$,

$$\limsup_{\varepsilon\to 0}\frac{W_2(\boldsymbol{\mu}(\xi+\varepsilon e),\boldsymbol{\mu}(\xi))}{|\varepsilon|} \leqslant \sqrt{\int_D |\mathbf{v}(\xi,x)\cdot e|^2\boldsymbol{\mu}(\xi,\mathrm{d}x)}. \tag{3.4}$$

Proof. To prove that $\boldsymbol{\mu}$ is Lipschitz, just notice that $|\mathbf{v}^\gamma|\leqslant|\mathbf{v}||\dot\gamma|$ and that $\mathbf{v}$ and ρ are in $L^\infty(\Omega\times D)$, then use (3.3). Moreover, if $\xi\in\mathring{\Omega}$ and $e\in\mathbf{R}^p$ is a unit vector, take $\gamma:[0,\varepsilon_0]\to\Omega$ to be the straight line defined by $\gamma(t):=\xi+te$: it is well defined for if $\varepsilon_0>0$ is small enough. Then just use (3.3) and the continuity of $\mathbf{v}$ in $\mathring{\Omega}$. □

This result is all we need in the sequel but much more could be proved: for instance, in (3.4), the $\limsup$ is a $\lim$, and the inequality is an equality.

3.3. **Equivalence with Sobolev spaces valued in metric spaces.** Until now, we have not discussed the existence of solutions to the (generalized) continuity equation: this notion could be too strong or too loose. In this subsection, we will show that the set of $\boldsymbol{\mu}$ with finite Dirichlet energy coincides with an already known definition of Sobolev spaces valued in metric spaces given by Reshetnyak [25, 26]. This definition is restricted to the case where the starting space has a smooth structure (which is precisely our framework), but can be seen as particular case of a more general definition given by Hajłasz (a pedagogic and clear introduction to the latter can be found in [3, Chapter 5]).

Definition 3.15. *Let $\boldsymbol{\mu}\in L^2(\Omega,\mathcal{P}(D))$. For any $\nu\in\mathcal{P}(D)$, define $[\boldsymbol{\mu}]_\nu\in L^2(\Omega)$ by $[\boldsymbol{\mu}]_\nu(\xi):=W_2(\boldsymbol{\mu}(\xi),\nu)$. We say that $\boldsymbol{\mu}\in H^1(\Omega,\mathcal{P}(D))$ if there exists a countable family $(\nu_n)_{n\in\mathbf{N}}$ dense in $\mathcal{P}(D)$ such that $[\boldsymbol{\mu}]_{\nu_n}\in H^1(\Omega)$ for all $n\in\mathbf{N}$ and there exists a function $g\in L^2(\Omega)$ such that, for every $n\in\mathbf{N}$,*

$$|\nabla[\boldsymbol{\mu}]_{\nu_n}|\leqslant g \tag{3.5}$$

a.e. on Ω. The smallest g for which (3.5) *holds is called the metric gradient of $\boldsymbol{\mu}$ and is denoted by $g_{\boldsymbol{\mu}}$.*

Notice that $g_{\boldsymbol{\mu}}=\sup_n|\nabla[\boldsymbol{\mu}]_{\nu_n}|$. The definition looks slightly different than in [25]. However, it is equivalent because of the following result:

Proposition 3.16. *Let $\boldsymbol{\mu}\in H^1(\Omega,\mathcal{P}(D))$ and $g_{\boldsymbol{\mu}}\in L^2(\Omega)$ be its metric gradient. Then for all mappings $u:\mathcal{P}(D)\to\mathbf{R}$ which are C-Lipschitz, $u\circ\boldsymbol{\mu}\in H^1(\Omega)$ and $|\nabla(u\circ\boldsymbol{\mu})|\leqslant Cg_{\boldsymbol{\mu}}$ a.e. on Ω.*

Proof. Is is enough to copy the proof of [25, Theorem 5.1]. Indeed, in this proof, one only uses the functions $[\boldsymbol{\mu}]_\nu$ for measures ν belonging to a dense and countable subset of $\mathcal{P}(D)$. □

In particular, if $\boldsymbol{\mu}\in H^1(\Omega,\mathcal{P}(D))$, then $[\boldsymbol{\mu}]_\nu\in H^1(\Omega)$ with gradient bounded by $g_{\boldsymbol{\mu}}$ for all $\nu\in\mathcal{P}(D)$. Notice that the definition above can be stated for mappings valued in arbitrary metric spaces (separability of the target space is required). The main theorem of this subsection is the following, which states that the framework that we have developed coincides with the one of Reshetnyak.

Theorem 3.17. *Let $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$. Then $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ if and only if $\mathrm{Dir}(\boldsymbol{\mu}) < +\infty$. Moreover, if $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ and if $\mathbf{v}$ is tangent to $\boldsymbol{\mu}$, then for a.e. $\xi \in \Omega$,*

$$g_{\boldsymbol{\mu}}(\xi) = \sqrt{\int_D |\mathbf{v}(\xi, x)|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x)}.$$

As an immediate corollary, notice that if $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$, then

$$\mathrm{Dir}(\boldsymbol{\mu}) = \int_\Omega \frac{1}{2} |g_{\boldsymbol{\mu}}(\xi)|^2 \mathrm{d}\xi.$$

We will prove this theorem in three steps. The first one is to prove it if Ω is a segment of $\mathbf{R}$ (Proposition 3.18). It is just a rewriting of the definition of Reshetnyak and does not rely of the special structure of the Wasserstein space. The second step is to say that, roughly speaking, a function is in $H^1(\Omega)$ if it is in H^1 for a.e. lines, with some uniform control on the gradients. It enables us to get the result if Ω is a cube (Proposition 3.20). The third step is simply to write that every domain can be written as a (countable) union of cubes.

Proposition 3.18. *Theorem 3.17 holds if Ω is a segment of $\mathbf{R}$.*

Proof. Assume $\Omega = I$ is a segment of $\mathbf{R}$. The set of curves with finite Dirichlet energy coincides with the set of absolutely continuous curves, see Proposition 3.8. Given Theorem 2.2, we want to prove the equality $g_{\boldsymbol{\mu}} = |\dot{\boldsymbol{\mu}}|$ a.e. on I.

Assume that $\mathrm{Dir}(\boldsymbol{\mu}) < +\infty$ and take $\nu \in \mathcal{P}(D)$. Then, as $W_2(\cdot, \nu)$ is 1-Lipschitz, for all $s < t$ elements of I,

$$|[\boldsymbol{\mu}]_\nu(t) - [\boldsymbol{\mu}]_\nu(s)| \leqslant W_2(\boldsymbol{\mu}(t), \boldsymbol{\mu}(s)) \leqslant \int_s^t |\dot{\boldsymbol{\mu}}|(r)\mathrm{d}r.$$

It shows that the function $[\boldsymbol{\mu}]_\nu$ is in $H^1(I)$ and its gradient is smaller than $|\dot{\boldsymbol{\mu}}|$. Hence, as ν is arbitrary, $\boldsymbol{\mu} \in H^1(I, \mathcal{P}(D))$ and $g_{\boldsymbol{\mu}} \leqslant |\dot{\boldsymbol{\mu}}|$.

Reciprocally, assume $\boldsymbol{\mu} \in H^1(I, \mathcal{P}(D))$, take $(\nu_n)_{n \in \mathbf{N}}$ countable and dense in $\mathcal{P}(D)$ such that $[\boldsymbol{\mu}]_{\nu_n} \in H^1(I)$ for every $n \in \mathbf{N}$ with gradient bounded by $g_{\boldsymbol{\mu}}$. In particular, for any $n \in \mathbf{N}$ and any $s < t$ elements of I,

$$|[\boldsymbol{\mu}]_{\nu_n}(t) - [\boldsymbol{\mu}]_{\nu_n}(s)| \leqslant \int_s^t g_{\boldsymbol{\mu}}(r)\mathrm{d}r.$$

Then we choose ν_n arbitrary close to $\boldsymbol{\mu}(t)$: the r.h.s. is unchanged and the l.h.s. is arbitrary close to $W_2(\boldsymbol{\mu}(t), \boldsymbol{\mu}(s))$. Hence we conclude that

$$W_2(\boldsymbol{\mu}(s), \boldsymbol{\mu}(t)) \leqslant \int_t^s g_{\boldsymbol{\mu}}(r)\mathrm{d}r,$$

which is enough to say that $\boldsymbol{\mu}$ is an absolutely continuous curve and $|\dot{\boldsymbol{\mu}}| \leqslant g_{\boldsymbol{\mu}}$ a.e. on I by minimality of $|\dot{\boldsymbol{\mu}}|$. □

Now we will prove Theorem 3.17 at least locally, which means in the case where Ω is a cube. Up to an isometry and a dilatation, we can assume that Ω is the unit cube of $\mathbf{R}^p$. Recall that $(e_\alpha)_{1 \leqslant \alpha \leqslant p}$ is the canonical basis of $\mathbf{R}^p$. In the sequel, we will denote by $\Omega_\alpha \subset \mathbf{R}^p$ the α-face of the cube, which means the set of $(\xi^1, \dots, \xi^{\alpha-1}, 0, \xi^{\alpha+1}, \dots, \xi^p)$, with $0 \leqslant \xi^\beta \leqslant 1$ for all $\beta \neq \alpha$. The measure on Ω_α will be the $p-1$-dimensional Lebesgue measure. If $f : \Omega \to X$ is a given mapping (where X is any set) and $\xi \in \Omega_\alpha$ is fixed, then $f_\xi : [0,1] \to X$ is defined by $f_\xi(t) = f(\xi + te_\alpha)$: it is the restriction of f to a line directed by e_α and crossing Ω_α at ξ. Recall the following characterization for real-valued mappings:

Proposition 3.19. *Assume Ω is the unit cube of $\mathbf{R}^p$ and let $f \in L^2(\Omega)$ be a given function. The function f belongs to $H^1(\Omega)$ if and only if for any $\alpha \in \{1, 2, \dots, p\}$, for a.e. $\xi \in \Omega_\alpha$, the function f_ξ is in $H^1([0,1])$ and*

$$\int_{\Omega_\alpha} \left(\int_0^1 |\dot{f}_\xi(t)|^2 \mathrm{d}t \right) \mathrm{d}\xi < +\infty.$$

Moreover, for a.e. $\xi \in \Omega_\alpha$ and a.e. $t \in [0,1]$,

$$(\partial_\alpha f)(\xi + te_\alpha) = \dot{f}_\xi(t).$$

Proof. One can look at [12, Section 4.9]. □

Proposition 3.20. *Theorem 3.17 holds if Ω is the unit cube of $\mathbf{R}^p$.*

Proof. Implication $\mathrm{Dir}(\boldsymbol{\mu}) < +\infty \Rightarrow \boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$. Assume first that $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ is such that $\mathrm{Dir}(\boldsymbol{\mu}) < +\infty$ and take $\mathbf{v} \in L^2_{\boldsymbol{\mu}}(\Omega \times D, \mathbf{R}^{pq})$ the velocity field tangent to $\boldsymbol{\mu}$. Fix $\alpha \in \{1, 2, \dots, p\}$. Take two compactly supported test functions $\psi \in C^1_c(]0,1[\times D)$ and $a \in C^1_c(\Omega_\alpha)$. As a test function $\varphi \in C^1_c(\Omega \times D, \mathbf{R}^p)$ in the weak formulation of the continuity equation, choose $\varphi(\xi + te_\alpha, x) := (0, 0, \dots, 0, \psi(t,x)a(\xi), 0, \dots, 0)$ for $\xi \in \Omega_\alpha$ and $t \in [0,1]$ (only the α-th component of φ is not 0). If we expand we find that $\nabla_\Omega \cdot \varphi = a\partial_t \psi$ hence

$$0 = \int_{\Omega \times D} \nabla_\Omega \cdot \varphi \mathrm{d}\boldsymbol{\mu} + \int_{\Omega \times D} \nabla_D \varphi \cdot \mathbf{v} \mathrm{d}\boldsymbol{\mu} = \int_{\Omega_\alpha} \left(\int_{[0,1]\times D} \partial_t \psi(t,x) \mathrm{d}t \boldsymbol{\mu}(\xi + te_\alpha, \mathrm{d}x) \right) a(\xi) \mathrm{d}\xi$$
$$+ \int_{\Omega_\alpha} \left(\int_{[0,1]\times D} \nabla_D \psi(t,x) \cdot (\mathbf{v}(\xi + te_\alpha, x) \cdot e_\alpha) \mathrm{d}t \boldsymbol{\mu}(\xi + te_\alpha, \mathrm{d}x) \right) a(\xi) \mathrm{d}\xi.$$

Using the arbitrariness of a, we deduce that for a.e. $\xi \in \Omega_\alpha$, and for a fixed $\psi \in C^1_c(]0,1[\times D, \mathbf{R}^p)$,

$$\int_{[0,1]\times D} \partial_t \psi(t,x) \mathrm{d}t \boldsymbol{\mu}(\xi + te_\alpha, \mathrm{d}x) + \int_{[0,1]\times D} \nabla_D \psi(t,x) \cdot (\mathbf{v}(\xi + te_\alpha, x) \cdot e_\alpha) \mathrm{d}t \boldsymbol{\mu}(\xi + te_\alpha, \mathrm{d}x) = 0. \tag{3.6}$$

Now, taking a sequence $(\psi_n)_{n \in \mathbf{N}}$ which is dense in $C^1_c(]0,1[\times D, \mathbf{R}^p)$, we can say that for a.e. $\xi \in \Omega_\alpha$, for all $\psi \in C^1_c(]0,1[\times D, \mathbf{R}^p)$, (3.6) holds. For $\xi \in \Omega_\alpha$ define $\boldsymbol{\mu}_\xi : [0,1] \to \mathcal{P}(D)$ by $\boldsymbol{\mu}_\xi(t) = \boldsymbol{\mu}(\xi + te_\alpha)$ and $\mathbf{v}_\xi : [0,1] \times D \to \mathbf{R}^q$ by $\mathbf{v}_\xi(t,x) = \mathbf{v}(\xi + te_\alpha, x) \cdot e_\alpha$. By Fubini's theorem, for a.e. $\xi \in \Omega_\alpha$, $\mathbf{v}_\xi \in L^2_{\boldsymbol{\mu}_\xi}([0,1] \times D, \mathbf{R}^q)$. Hence (3.6) rewrites as: for a.e. $\xi \in \Omega_\alpha$, the curve $\boldsymbol{\mu}_\xi$ is an absolutely continuous curve in the Wasserstein space with a velocity field given by $\mathbf{v}_\xi$. By Proposition 3.18, if $\nu \in \mathcal{P}(D)$, then the function $[\boldsymbol{\mu}_\xi]_\nu$ is in $H^1([0,1])$ and

$$|\partial_t [\boldsymbol{\mu}_\xi]_\nu(t)| \leqslant \sqrt{\int_D |\mathbf{v}_\xi(t,x)|^2 \boldsymbol{\mu}_\xi(t, \mathrm{d}x)} = \sqrt{\int_D |\mathbf{v}(\xi + te_\alpha, x) \cdot e_\alpha|^2 \boldsymbol{\mu}(\xi + te_\alpha, \mathrm{d}x)}.$$

As the r.h.s. is integrable over $[0,1] \times \Omega_\alpha$ and α is arbitrary, we can use Proposition 3.19 to see that $[\boldsymbol{\mu}]_\nu \in H^1(\Omega)$. Moreover, taking the square of the previous equation and summing over $\alpha \in \{1, 2, \dots, p\}$, we see that for a.e. $\xi \in \Omega$

$$|\nabla [\boldsymbol{\mu}]_\nu(\xi)|^2 \leqslant \int_D |\mathbf{v}(\xi, x)|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x).$$

Thus, we conclude that $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ and for a.e. $\xi \in \Omega$,

$$g_{\boldsymbol{\mu}}(\xi) \leqslant \sqrt{\int_D |\mathbf{v}(\xi, x)|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x)}. \tag{3.7}$$

Implication $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D)) \Rightarrow \mathrm{Dir}(\boldsymbol{\mu}) < +\infty$. Let $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$. Take $(\nu_n)_{n\in\mathbf{N}}$ a sequence which is dense in $\mathcal{P}(D)$. For any $n \in \mathbf{N}$, the function $[\boldsymbol{\mu}]_{\nu_n}$ belongs to $H^1(\Omega)$. Fix $\alpha \in \{1, 2, \dots, p\}$. For any $n \in \mathbf{N}$, for a.e. $\xi \in \Omega_\alpha$, the function $[\boldsymbol{\mu}_\xi]_{\nu_n} : t \mapsto W_2(\boldsymbol{\mu}(\xi + te_\alpha), \nu_n)$ is in $H^1([0,1])$ with a gradient bounded by $g_{\boldsymbol{\mu}}(\xi + te_\alpha)$. As $\mathbf{N}$ is countable, we can exchange the "for a.e. $\xi \in \Omega_\alpha$" and the "for all $n \in \mathbf{N}$". Hence, for a.e. $\xi \in \Omega_\alpha$, the function $\boldsymbol{\mu}_\xi : [0,1] \to \mathcal{P}(D)$ belongs to $H^1([0,1], \mathcal{P}(D))$ with a gradient bounded by $g_{\boldsymbol{\mu}}(\xi + te_\alpha)$. For a given $\xi \in \Omega_\alpha$, we can use Proposition 3.18 and Theorem 2.2 to get the existence of a velocity field $\mathbf{w}_\xi^\alpha \in L^2_{\boldsymbol{\mu}_\xi}([0,1] \times D, \mathbf{R}^q)$ such that $(\boldsymbol{\mu}_\xi, \mathbf{w}_\xi^\alpha \boldsymbol{\mu}_\xi)$ satisfies the (1-dimensional) continuity equation and for a.e. $t \in [0,1]$,

$$\sqrt{\int_D |\mathbf{w}_\xi^\alpha(t,x)|^2 \boldsymbol{\mu}(\xi + te_\alpha, \mathrm{d}x)} \leqslant |\dot{\boldsymbol{\mu}}_\xi(t)| = |g_{\boldsymbol{\mu}_\xi}(t)| \leqslant g_{\boldsymbol{\mu}}(\xi + te_\alpha). \tag{3.8}$$

Now, do this for a.e. $\xi \in \Omega_\alpha$ and then for any $\alpha \in \{1, 2, \dots, p\}$. Define the velocity field $\mathbf{v} : \Omega \times D \to \mathbf{R}^{pq}$ component by component, the α-th component at the point $\xi + te_\alpha$ (with $\xi \in \Omega_\alpha$) being defined as $\mathbf{w}_\xi^\alpha(t)$. To justify that $\mathbf{v}$ is measurable, notice that $\mathbf{w}_\xi^\alpha$ is the solution of an optimization problem [4, Equation (8.3.11)] which depends in a measurable way of ξ, thus one can apply Proposition A.1. By the bound (3.8), it is clear that $\mathbf{v} \in L^2_{\boldsymbol{\mu}}(\Omega \times D, \mathbf{R}^{pq})$. Moreover, if $\varphi \in C^1_c(\Omega \times D, \mathbf{R}^p)$,

$$\begin{aligned}
\int_{\Omega\times D} \nabla_\Omega \cdot \varphi \mathrm{d}\boldsymbol{\mu} &= \sum_{\alpha=1}^p \int_{\Omega\times D} \partial_\alpha \varphi^\alpha(\xi, x) \boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x) \\
&= \sum_{\alpha=1}^p \int_{\Omega_\alpha} \left(\int_0^1 \partial_\alpha \varphi^\alpha(\xi + te_\alpha, x) \boldsymbol{\mu}(\xi + te_\alpha, \mathrm{d}x) \mathrm{d}t \right) \mathrm{d}\xi \\
&= -\sum_{\alpha=1}^p \int_{\Omega_\alpha} \left(\int_0^1 \nabla_D \varphi^\alpha(\xi + te_\alpha, x) \cdot \mathbf{w}_\xi^\alpha(t,x) \boldsymbol{\mu}(\xi + te_\alpha, \mathrm{d}x) \mathrm{d}t \right) \mathrm{d}\xi \\
&= -\sum_{\alpha=1}^p \int_{\Omega_\alpha} \left(\int_0^1 \nabla_D \varphi^\alpha(\xi + te_\alpha, x) (\mathbf{v}(\xi + te_\alpha, x) \cdot e_\alpha) \boldsymbol{\mu}(\xi + te_\alpha, \mathrm{d}x) \mathrm{d}t \right) \mathrm{d}\xi \\
&= -\int_{\Omega\times D} \nabla_D \varphi \cdot \mathbf{v} \mathrm{d}\boldsymbol{\mu}.
\end{aligned}$$

(The second and last inequalities are Fubini's theorem and the third one comes from the 1-dimensional continuity equations). Hence, we see that $(\boldsymbol{\mu}, \mathbf{v}\boldsymbol{\mu})$ satisfies the continuity equation.

To conclude, we need to show the opposite inequality in (3.7). If $\alpha \in \{1, 2, \dots, p\}$, for a.e. $\xi \in \Omega_\alpha$ and a.e. $t \in [0,1]$, one has, by definition of $g_{\boldsymbol{\mu}_\xi}$ and Proposition 3.18,

$$\sqrt{\int_D |\mathbf{w}_\xi^\alpha(t,x)|^2 \boldsymbol{\mu}(\xi + te_\alpha, \mathrm{d}x)} = g_{\boldsymbol{\mu}_\xi}(t) = \sup_{n\in\mathbf{N}} \left| \partial_\alpha [\boldsymbol{\mu}]_{\nu_n}(\xi + te_\alpha) \right|,$$

which can be rewritten as: for a.e. $\xi \in \Omega$, for all $\alpha \in \{1, 2, \dots, p\}$,

$$\sqrt{\int_D |\mathbf{v}(\xi, x) \cdot e_\alpha|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x)} = \sup_{n\in\mathbf{N}} \left| e_\alpha \cdot \nabla [\boldsymbol{\mu}]_{\nu_n}(\xi) \right|. \tag{3.9}$$

But this is not sufficient as it only holds in particular directions, namely the ones given by the canonical basis of $\mathbf{R}^p$. To go further, we will show that the velocity field $\mathbf{v}$ that we have constructed is tangent to $\boldsymbol{\mu}$, i.e. that $\mathrm{Dir}(\boldsymbol{\mu}) = \mathrm{Dir}(\boldsymbol{\mu}, \mathbf{v}\boldsymbol{\mu})$. Indeed, thanks to Theorem 2.2, as each $\mathbf{w}_\xi^\alpha$ is tangent, we can say that for a.e. $\xi \in \Omega$, there exists a sequence $(\psi_n)_{n\in\mathbf{N}}$ in $C^1(D, \mathbf{R}^p)$ such that $\mathbf{v}(\xi, \cdot)$ is the limit of the sequence $(\nabla\psi_n)_{n\in\mathbf{N}}$ in $L^2_{\boldsymbol{\mu}(\xi)}(D, \mathbf{R}^{pq})$. Now take, $\tilde{\mathbf{v}} \in L^2_{\boldsymbol{\mu}}(\Omega \times D, \mathbf{R}^{pq})$ such that $(\boldsymbol{\mu}, \tilde{\mathbf{v}}\boldsymbol{\mu})$

satisfies the continuity equation. In particular, for any $\varphi \in C^1_c(\Omega \times D, \mathbf{R}^p)$, one has

$$\int_{\Omega\times D} \nabla_D \varphi \cdot (\tilde{\mathbf{v}} - \mathbf{v}) \mathrm{d}\boldsymbol{\mu} = 0.$$

If we take $a \in C^1_c(\mathring{\Omega})$ and $\psi \in C^1(D, \mathbf{R}^p)$ and use $\varphi(\xi, x) := a(\xi)\psi(x)$ as a test function, we see that

$$\int_\Omega a(\xi) \left(\int_D \nabla\psi(x) \cdot (\tilde{\mathbf{v}}(\xi, x) - \mathbf{v}(\xi, x)) \boldsymbol{\mu}(\xi, \mathrm{d}x) \right) \mathrm{d}\xi = 0.$$

As a is arbitrary, we see that for all $\psi \in C^1(D, \mathbf{R}^p)$, for a.e. $\xi \in \Omega$,

$$\int_D \nabla\psi(x) \cdot (\tilde{\mathbf{v}}(\xi, x) - \mathbf{v}(\xi, x)) \boldsymbol{\mu}(\xi, \mathrm{d}x) = 0. \tag{3.10}$$

Taking a sequence $(\tilde{\psi}_n)_{n\in\mathbf{N}}$ dense in $C^1(D, \mathbf{R}^p)$, we conclude that for a.e. $\xi \in \Omega$, (3.10) holds for all $\psi \in C^1(D)$. Hence taking a sequence $(\psi_n)_{n\in\mathbf{N}}$ such that $(\nabla\psi_n)_{n\in\mathbf{N}}$ converges to $\mathbf{v}(\xi, \cdot)$ in $L^2_{\boldsymbol{\mu}(\xi)}(D, \mathbf{R}^{pq})$, one sees that for a.e. $\xi \in \Omega$,

$$\int_D \mathbf{v}(\xi, x) \cdot (\tilde{\mathbf{v}}(\xi, x) - \mathbf{v}(\xi, x)) \boldsymbol{\mu}(\xi, \mathrm{d}x) = 0.$$

In particular, integrating w.r.t. ξ, we see that $\mathbf{v}$ and $\tilde{\mathbf{v}} - \mathbf{v}$ are orthogonal in $L^2_{\boldsymbol{\mu}}(\Omega \times D, \mathbf{R}^{pq})$. Hence,

$$\mathrm{Dir}(\boldsymbol{\mu}, \tilde{\mathbf{v}}\boldsymbol{\mu}) = \int_{\Omega\times D} \frac{1}{2}|\tilde{\mathbf{v}}|^2 \mathrm{d}\boldsymbol{\mu} = \int_{\Omega\times D} \frac{1}{2}|\tilde{\mathbf{v}} - \mathbf{v} + \mathbf{v}|^2 \mathrm{d}\boldsymbol{\mu} \geqslant \int_{\Omega\times D} \frac{1}{2}|\mathbf{v}|^2 \mathrm{d}\boldsymbol{\mu} = \mathrm{Dir}(\boldsymbol{\mu}, \mathbf{v}\boldsymbol{\mu}).$$

As $\tilde{\mathbf{v}}$ is arbitrary, we conclude that $\mathrm{Dir}(\boldsymbol{\mu}) = \mathrm{Dir}(\boldsymbol{\mu}, \mathbf{v}\boldsymbol{\mu})$, which means that $\mathbf{v}$ is tangent to $\boldsymbol{\mu}$.

We know that there exists a unique ($\boldsymbol{\mu}$-a.e.) tangent velocity field, see Proposition 3.9. Hence, our construction to build $\mathbf{v}$, which was strongly dependent on the canonical basis of $\mathbf{R}^p$ gives in fact the tangent velocity field. In consequence, the same construction can be performed in any orthonormal basis and will lead to the same velocity field $\mathbf{v}$. In particular, for any unit vector $e \in \mathbf{R}^p$, for a.e. $\xi \in \Omega$, (3.9) holds provided that we replace e_α by e. Take $(\tilde{e}_m)_{m\in\mathbf{N}}$ a sequence of unit vector which is dense in the unit sphere of $\mathbf{R}^p$. Then, for a.e. $\xi \in \Omega$, one has

$$\sqrt{\int_D |\mathbf{v}(\xi, x) \cdot \tilde{e}_m|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x)} = \sup_{n\in\mathbf{N}} |\tilde{e}_m \cdot \nabla[\boldsymbol{\mu}]_{\nu_n}(\xi)|,$$

Taking the supremum in $m \in \mathbf{N}$ on both sides, and exchanging the supremum in $m \in \mathbf{N}$ and $n \in \mathbf{N}$ one sees that

$$\begin{aligned}\sqrt{\int_D |\mathbf{v}(\xi, x)|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x)} &= \sup_{m\in\mathbf{N}} \sqrt{\int_D |\mathbf{v}(\xi, x) \cdot \tilde{e}_m|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x)} \\ &= \sup_{m\in\mathbf{N}} \sup_{n\in\mathbf{N}} |\tilde{e}_m \cdot \nabla[\boldsymbol{\mu}]_{\nu_n}(\xi)| = \sup_{n\in\mathbf{N}} \sup_{m\in\mathbf{N}} |\tilde{e}_m \cdot \nabla[\boldsymbol{\mu}]_{\nu_n}(\xi)| = \sup_{n\in\mathbf{N}} |\nabla[\boldsymbol{\mu}]_{\nu_n}(\xi)| = g_{\boldsymbol{\mu}}(\xi),\end{aligned}$$

where the last equality comes from the definition of $g_{\boldsymbol{\mu}}$. □

To conclude the proof of the theorem, we just have to justify that we can put the pieces together.

Proof of Theorem 3.17. The domain $\mathring{\Omega}$ can be cut in a (countable) number of cubes $(\Omega_m)_{m\in\mathbf{N}}$. The boundary $\partial\Omega$ does not play any role as $\mathcal{L}_\Omega(\partial\Omega) = 0$.

Implication $\mathrm{Dir}(\boldsymbol{\mu}) < +\infty \Rightarrow \boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$. Assume first that $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ is such that $\mathrm{Dir}(\boldsymbol{\mu}) < +\infty$ and take $\mathbf{v} \in L^2_{\boldsymbol{\mu}}(\Omega \times D, \mathbf{R}^{pq})$ the velocity field tangent to $\boldsymbol{\mu}$. Fix $n \in \mathbf{N}$. On each cube Ω_m, we know that the function $[\boldsymbol{\mu}]_{\nu_n}$ is in $H^1(\Omega_m)$ with a gradient which is bounded by a function which does not depend on n and is in $L^2(\Omega)$, which is sufficient to say that $[\boldsymbol{\mu}]_{\nu_n} \in H^1(\Omega)$ with a gradient bounded by a function which does not depend on $n \in \mathbf{N}$.

Implication $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D)) \Rightarrow \mathrm{Dir}(\boldsymbol{\mu}) < +\infty$. Assume that $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$. For any $m \in \mathbf{N}$, one can construct a tangent velocity field $\mathbf{v} \in L^2_{\boldsymbol{\mu}}(\Omega_m \times D, \mathbf{R}^{pq})$. Combining Proposition 3.9 giving the uniqueness $\boldsymbol{\mu}$-a.e. of the tangent velocity field and Proposition 3.12 which enables to localize, one sees that if $\Omega_{m_1} \cap \Omega_{m_2} \neq \varnothing$, then the tangent velocity fields $\mathbf{v}_1 \in L^2_{\boldsymbol{\mu}}(\Omega_{m_1} \times D, \mathbf{R}^{pq})$ and $\mathbf{v} \in L^2_{\boldsymbol{\mu}}(\Omega_{m_2} \times D, \mathbf{R}^{pq})$ coincide $\boldsymbol{\mu}$-a.e. on $\Omega_{m_1} \cap \Omega_{m_2}$. Thus, one can define a velocity field $\mathbf{v}$ on the whole Ω, and it is straightforward to check that $\mathbf{v}$ is tangent to $\boldsymbol{\mu}$. □

Notice that we have proved an interesting corollary when proving the theorem in the case of a cube, which is a generalization of [4, Proposition 8.4.5].

Corollary 3.21. *Let $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ and $\mathbf{v} \in L^2_{\boldsymbol{\mu}}(\Omega \times D, \mathbf{R}^{pq})$ such that $(\boldsymbol{\mu}, \mathbf{v}\boldsymbol{\mu})$ satisfies the continuity equation. Then $\mathbf{v}$ is tangent to $\boldsymbol{\mu}$ if and only if, for a.e. $\xi \in \Omega$, there exists a sequence $(\psi_n)_{n\in\mathbf{N}}$ in $C^1(D, \mathbf{R}^p)$ such that $(\nabla\psi_n)_{n\in\mathbf{N}}$ converges to $\mathbf{v}(\xi, \cdot)$ in $L^2_{\boldsymbol{\mu}(\xi)}(D, \mathbf{R}^{pq})$.*

3.4. **Equivalence with Dirichlet energy in metric spaces.** In this subsection we will show that our definition coincides with the one of Koorevaar, Schoen, and Jost [19, 16]. As explained in the introduction, their formulation goes as follows.

Definition 3.22. *Let $\varepsilon > 0$ and $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$. We define the ε-Dirichlet energy of $\boldsymbol{\mu}$ by*

$$\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}) := C_p \int_{\Omega\times\Omega} \frac{W_2^2(\boldsymbol{\mu}(\xi), \boldsymbol{\mu}(\eta))}{2\varepsilon^{p+2}} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \mathrm{d}\xi\mathrm{d}\eta,$$

where the normalization constant C_p is defined as $C_p := |\eta|^2 \left(\int_{B(0,1)} |\xi \cdot \eta|^2 \mathrm{d}\xi\right)^{-1}$.

One can notice that the ε-Dirichlet energy is always finite as $\mathcal{P}(D)$ has a finite diameter, but it can blow up when $\varepsilon \to 0$. The goal is to prove that Dir_ε is a good approximation of Dir if ε is small enough. Before stating the main result, let us do the following observation, which will be useful in the sequel.

Proposition 3.23. *Let $\varepsilon > 0$ be fixed. Then the functional $\mathrm{Dir}_\varepsilon : L^2(\Omega, \mathcal{P}(D)) \to \mathbf{R}$ is continuous w.r.t. strong convergence and l.s.c. w.r.t. the weak convergence.*

Proof. The continuity w.r.t. strong convergence is simple: recall that $\mathcal{P}(D)$ has a finite diameter, thus Lebesgue dominated convergence theorem is enough. The lower semi-continuity relies on the fact that W_2^2 is a supremum of continuous linear functionals, thus is l.s.c. and convex.

More precisely, fix $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ and a sequence $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ which converges weakly to $\boldsymbol{\mu}$ in $L^2(\Omega, \mathcal{P}(D))$. If ξ and η are points of Ω, take $(\varphi(\xi, \eta, \cdot), \psi(\xi, \eta, \cdot))$ a pair of Kantorovitch potential between $\boldsymbol{\mu}(\xi)$ and $\boldsymbol{\mu}(\eta)$. In other words, $\varphi(\xi, \eta, \cdot)$ and $\psi(\xi, \eta, \cdot)$ are continuous functions (in fact uniformly Lipschitz), such that $\varphi(\xi, \eta, x) + \psi(\xi, \eta, y) \leqslant |x - y|^2$ for any $x, y \in D$, and such that

$$W_2^2(\boldsymbol{\mu}(\xi), \boldsymbol{\mu}(\eta)) = \int_D \varphi(\xi, \eta, x)\boldsymbol{\mu}(\xi, \mathrm{d}x) + \int_D \psi(\xi, \eta, y)\boldsymbol{\mu}(\eta, \mathrm{d}y). \tag{3.11}$$

One can do that in such a way that $\varphi : \Omega \times \Omega \to C(D)$ and $\psi : \Omega \times \Omega \to C(D)$ are measurable. Indeed, for fixed ξ and η, $(\varphi(\xi, \eta, \cdot), \psi(\xi, \eta, \cdot)) \in C(D) \times C(D)$ is a maximizer a functional which is continuous on $C(D) \times C(D)$ and which depends on ξ and η in a measurable way: hence we can apply Proposition A.1. Then, using the double convexification trick (see [31, Section 2.1]) which is a measurable operation, we can assume that (φ, ψ) are uniformly (w.r.t. ξ and η) Lipschitz and bounded as elements of $C(D)$. By the Kantorovitch duality, for every $n \in \mathbf{N}$,

$$\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_n) \geqslant \frac{C_p}{2\varepsilon^{p+2}} \int_{\Omega\times\Omega} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \left(\int_D \varphi(\xi, \eta, x)\boldsymbol{\mu}_n(\xi, \mathrm{d}x) + \int_D \psi(\xi, \eta, y)\boldsymbol{\mu}_n(\eta, \mathrm{d}y) \right) \mathrm{d}\xi\mathrm{d}\eta. \tag{3.12}$$

Now, apply Lusin's theorem to the mapping $\varphi : \Omega \times \Omega \to C(D)$ (for Lusin's theorem to other spaces than $\mathbf{R}$, see for instance [27, Box 1.6]). For any $\delta > 0$, we can find a compact $X \subset \Omega \times \Omega$ such that $\mathcal{L}_\Omega \otimes \mathcal{L}_\Omega([\Omega \times \Omega] \backslash X) \leqslant \delta$ and $\varphi : X \to C(D)$ is continuous on X. Now notice, as $|\varphi(\xi,\eta,x) - \varphi(\xi,\eta,y)| \leqslant C|x-y|$ uniformly in ξ and η, that $\varphi : X \times D \to \mathbf{R}$ is a continuous function for the product topology on $X \times D \subset \Omega \times \Omega \times D$. This function can be extended in a function $\tilde{\varphi} \in C(\Omega \times \Omega \times D)$. To sum up, there exists a continuous function $\tilde{\varphi}$, which coincides with φ on $X \times D$ (the important point is that there is coincidence on all D). Thus, denoting by C a uniform bound of φ and $\tilde{\varphi}$, one has that for every $\boldsymbol{\nu} \in L^2(\Omega, \mathcal{P}(D))$,

(3.13)
$$\left| \int_{\Omega\times\Omega} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \left(\int_D \varphi(\xi,\eta,x) \boldsymbol{\nu}(\xi, \mathrm{d}x) \right) \mathrm{d}\xi \mathrm{d}\eta - \int_{\Omega\times\Omega} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \left(\int_D \tilde{\varphi}(\xi,\eta,x) \boldsymbol{\nu}(\xi, \mathrm{d}x) \right) \mathrm{d}\xi \mathrm{d}\eta \right| \leqslant C\delta.$$

On the other hand, using Fubini's theorem one sees that

$$\int_{\Omega\times\Omega} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \left(\int_D \tilde{\varphi}(\xi,\eta,x) \boldsymbol{\mu}_n(\xi, \mathrm{d}x) \right) \mathrm{d}\xi \mathrm{d}\eta = \int_{\Omega\times D} \left(\int_{B(\xi,\varepsilon)\cap\Omega} \tilde{\varphi}(\xi,\eta,x) \mathrm{d}\eta \right) \boldsymbol{\mu}_n(\mathrm{d}\xi, \mathrm{d}x).$$

As $\tilde{\varphi}$ is continuous and bounded, it is not difficult to see that

$$(\xi, x) \in \Omega \times D \mapsto \int_{B(\xi,\varepsilon)\cap\Omega} \tilde{\varphi}(\xi,\eta,x) \mathrm{d}\eta \in \mathbf{R}$$

is continuous. Hence, using the weak convergence of $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$,

$$\lim_{n\to+\infty} \int_{\Omega\times\Omega} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \left(\int_D \tilde{\varphi}(\xi,\eta,x) \boldsymbol{\mu}_n(\xi, \mathrm{d}x) \right) \mathrm{d}\xi \mathrm{d}\eta = \int_{\Omega\times\Omega} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \left(\int_D \tilde{\varphi}(\xi,\eta,x) \boldsymbol{\mu}(\xi, \mathrm{d}x) \right) \mathrm{d}\xi \mathrm{d}\eta.$$

Using equation (3.13) with both $\boldsymbol{\mu}_n$ and $\boldsymbol{\mu}$ as $\boldsymbol{\nu}$, and using moreover the arbitrariness of δ, we conclude that we can replace $\tilde{\varphi}$ by φ in the equation above:

$$\lim_{n\to+\infty} \int_{\Omega\times\Omega} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \left(\int_D \varphi(\xi,\eta,x) \boldsymbol{\mu}_n(\xi, \mathrm{d}x) \right) \mathrm{d}\xi \mathrm{d}\eta = \int_{\Omega\times\Omega} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \left(\int_D \varphi(\xi,\eta,x) \boldsymbol{\mu}(\xi, \mathrm{d}x) \right) \mathrm{d}\xi \mathrm{d}\eta.$$

Of course there is exactly the same statement with ψ. With the help of this information, combining (3.12) and (3.11), we reach the conclusion that

$$\begin{aligned}
&\liminf_{n\to+\infty} \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_n) \\
&\quad \geqslant \liminf_{n\to+\infty} \frac{C_p}{2\varepsilon^{p+2}} \int_{\Omega\times\Omega} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \left(\int_D \varphi(\xi,\eta,x) \boldsymbol{\mu}_n(\xi, \mathrm{d}x) + \int_D \psi(\xi,\eta,y) \boldsymbol{\mu}_n(\eta, \mathrm{d}y) \right) \mathrm{d}\xi \mathrm{d}\eta \\
&\quad = \frac{C_p}{2\varepsilon^{p+2}} \int_{\Omega\times\Omega} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \left(\int_D \varphi(\xi,\eta,x) \boldsymbol{\mu}(\xi, \mathrm{d}x) + \int_D \psi(\xi,\eta,y) \boldsymbol{\mu}(\eta, \mathrm{d}y) \right) \mathrm{d}\xi \mathrm{d}\eta \\
&\quad = \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}).
\end{aligned}$$

□

We are now ready to state and prove the main theorem of this subsection.

Theorem 3.24. *Let $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$. Then $\mathrm{Dir}_\varepsilon(\boldsymbol{\mu})$ converges to $\mathrm{Dir}(\boldsymbol{\mu})$ as $\varepsilon \to 0$, and the sequence $(\mathrm{Dir}_{2^{-n}\varepsilon_0}(\boldsymbol{\mu}))_{n\in\mathbf{N}}$ is increasing for any $\varepsilon_0 > 0$.*

In addition for any $\varepsilon_0 > 0$, $\mathrm{Dir}_{2^{-n}\varepsilon_0}$ Γ-converges to Dir on the space $L^2(\Omega, \mathcal{P}(D))$ endowed with the weak topology as $n \to +\infty$.

Proof. For this proof, we are following the proof of [17, Theorem 8.3.1 and Lemma 8.3.4] (where the same result is shown for mapping from one Riemannian manifold to another) and also using a key result of [26].

Monotonicity of Dir_ε. If $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$, $\varepsilon > 0$ and $\lambda \in (0,1)$ then one has

$$\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}) \leqslant \lambda \mathrm{Dir}_{\lambda\varepsilon}(\boldsymbol{\mu}) + (1-\lambda)\mathrm{Dir}_{(1-\lambda)\varepsilon}(\boldsymbol{\mu}).$$

Indeed, this is a consequence of the triangle inequality and is valid for mappings valued in arbitrary metric spaces, see for instance [16, Example 1) (i)] or [17, Equation (8.3.4)] for a proof. In particular, by taking $\lambda = 1/2$, we see that the sequence $(\mathrm{Dir}_{2^{-n}\varepsilon_0}(\boldsymbol{\mu}))_{n\in\mathbf{N}}$ is increasing for any $\varepsilon_0 > 0$. Moreover, with well chosen λ, one sees that for a fixed $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ the function $\varepsilon \mapsto \varepsilon \mathrm{Dir}_\varepsilon(\boldsymbol{\mu})$ is subadditive, which is enough to ensure the convergence of $\mathrm{Dir}_\varepsilon(\boldsymbol{\mu})$ to some limit in $[0, +\infty]$ as $\varepsilon \to 0$.

The smooth case. Let $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ and $\mathbf{v} \in L^2_{\boldsymbol{\mu}}(\Omega \times D, \mathbf{R}^{pq})$ be smooth: they have the same regularity than in Proposition 3.14. Assume moreover that $(\boldsymbol{\mu}, \mathbf{v}\boldsymbol{\mu})$ satisfies the continuity equation. We will show that the limit of $\mathrm{Dir}_\varepsilon(\boldsymbol{\mu})$ is smaller than $\mathrm{Dir}(\boldsymbol{\mu}, \mathbf{v}\boldsymbol{\mu})$. Indeed, one can write

$$\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}) = \int_\Omega \mathrm{dir}_\varepsilon(\xi)\mathrm{d}\xi,$$

where

$$\mathrm{dir}_\varepsilon(\xi) := C_p \int_{\Omega \cap B(\xi,\varepsilon)} \frac{W_2^2(\boldsymbol{\mu}(\xi), \boldsymbol{\mu}(\eta))}{2\varepsilon^{p+2}} \mathrm{d}\eta.$$

If $\xi \notin \partial\Omega$ (it happens for a.e. ξ), for ε small enough, $B(\xi, \varepsilon) \subset \Omega$ and we can perform the following change of variables in spherical coordinates: denoting by $\mathbf{S}^{p-1}$ the unit sphere of $\mathbf{R}^p$ and σ its surface measure,

$$\mathrm{dir}_\varepsilon(\xi) = \frac{C_p}{2} \int_{\mathbf{S}^{d-1}} \left(\int_0^1 \frac{W_2^2(\boldsymbol{\mu}(\xi), \boldsymbol{\mu}(\xi + r\varepsilon e))}{\varepsilon^2} r^{p-1} \mathrm{d}r \right) \sigma(\mathrm{d}e).$$

Thanks to Proposition 3.14 we have an upper bound for the pointwise limit of the integrand, and we can pass to the limit as $\varepsilon \to 0$: we can use for instance Fatou's lemma as we have a pointewise $\limsup$ and all the quantity are bounded uniformly in $\varepsilon > 0$ from above by a constant (recall that $\boldsymbol{\mu}$ is Lipschitz). Hence, for a.e. $\xi \in \Omega$,

$$\begin{aligned}
\limsup_{\varepsilon\to 0} \mathrm{dir}_\varepsilon(\xi) &\leqslant \frac{C_p}{2} \int_{\mathbf{S}^{d-1}} \left[\int_0^1 \left(\int_D |\mathbf{v}(\xi, x) \cdot (re)|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x) \right) r^{p-1} \mathrm{d}r \right] \sigma(\mathrm{d}e) \\
&= \frac{C_p}{2} \int_D \left(\int_{B(0,1)} |\mathbf{v}(\xi, x) \cdot \theta|^2 \mathrm{d}\theta \right) \boldsymbol{\mu}(\xi, \mathrm{d}x) \\
&= \frac{1}{2} \int_D |\mathbf{v}(\xi, x)|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x),
\end{aligned}$$

where the last inequality comes from the definition of C_p. To integrate this inequality over Ω, we can use again Fatou's lemma as we still have a uniform bound on $\mathbf{v}$, hence

$$\limsup_{\varepsilon\to 0} \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}) \leqslant \int_\Omega \left(\limsup_{\varepsilon\to 0} \mathrm{dir}_\varepsilon(\xi) \right) \mathrm{d}\xi \leqslant \int_\Omega \left(\int_D \frac{1}{2} |\mathbf{v}(\xi, x)|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x) \right) \mathrm{d}\xi = \mathrm{Dir}(\boldsymbol{\mu}, \mathbf{v}\boldsymbol{\mu}).$$

General case: $\lim_\varepsilon \mathrm{Dir}_\varepsilon \leqslant \mathrm{Dir}$. Let $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ and $E \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ be tangent to $\boldsymbol{\mu}$, we know that $(\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}))_{\varepsilon>0}$ goes to some limit $l \in [0, +\infty]$ as $\varepsilon \to 0$. Let $\delta > 0$ be fixed. There exists $\varepsilon_0 > 0$ such that

$$l \leqslant \mathrm{Dir}_{\varepsilon_0}(\boldsymbol{\mu}) + \delta.$$

Let $(\boldsymbol{\mu})_{n\in\mathbf{N}}$ and $(\mathbf{v}_n)_{n\in\mathbf{N}}$ be the two sequences given by Theorem 3.13. We choose n large enough so that $\mathrm{Dir}_{\varepsilon_0}(\boldsymbol{\mu}_n) \geqslant \mathrm{Dir}_{\varepsilon_0}(\boldsymbol{\mu}) - \delta$: it is possible because $\mathrm{Dir}_{\varepsilon_0}$ is continuous w.r.t. strong convergence on $L^2(\Omega, \mathcal{P}(D))$. Hence,

$$l \leqslant \mathrm{Dir}_{\varepsilon_0}(\boldsymbol{\mu}) + \delta \leqslant \mathrm{Dir}_{\varepsilon_0}(\boldsymbol{\mu}_n) + 2\delta \leqslant \mathrm{Dir}(\boldsymbol{\mu}_n, \mathbf{v}_n\boldsymbol{\mu}_n) + 2\delta \leqslant \mathrm{Dir}(\boldsymbol{\mu}, E) + 2\delta = \mathrm{Dir}(\boldsymbol{\mu}) + 2\delta,$$

where the third inequality comes from the monotonicity and the smooth case which is treated above. As δ is arbitrary, we get that $l \leqslant \mathrm{Dir}(\boldsymbol{\mu})$, which means

$$\lim_{\varepsilon\to 0} \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}) \leqslant \mathrm{Dir}(\boldsymbol{\mu}).$$

This equation still holds if $\boldsymbol{\mu} \notin H^1(\Omega, \mathcal{P}(D))$ as the r.h.s. is infinite.

General case: $\liminf_\varepsilon \mathrm{Dir}_\varepsilon \geqslant \mathrm{Dir}$. For this part, we rely on [26, Theorem 1]. Let us remark that the constant c given in this theorem is precisely $\sqrt{2C_p}$ (one can look closely at equation (2.7) in [26]), thus this theorem states, with our notations, that if $\limsup_\varepsilon \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}) < +\infty$, then $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ and

$$\int_\Omega \frac{1}{2} g_{\boldsymbol{\mu}}(\xi)^2 \mathrm{d}\xi \leqslant \lim_{\varepsilon\to 0} \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}),$$

where $g_{\boldsymbol{\mu}}$ is defined in Definition 3.15. Taking in account Theorem 3.17, we end up with

$$\mathrm{Dir}(\boldsymbol{\mu}) \leqslant \lim_{\varepsilon\to 0} \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}).$$

If $\limsup_\varepsilon \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}) = +\infty$, then the equation above trivially holds.

The Γ-convergence. The statement of Γ-convergence is now easy. To summarize, until now we have proved the monotonicity and that

$$\mathrm{Dir}(\boldsymbol{\mu}) = \lim_{\varepsilon\to 0} \mathrm{Dir}_\varepsilon(\boldsymbol{\mu})$$

for every $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$. It is an exercise that we leave to the reader to check that any sequence of functionals which are l.s.c. (which is the case for the Dir_ε, see Proposition 3.23) and which converges in a increasing way in fact Γ-converges. □

3.5. **Boundary values.** It is well known that it is possible to make sense of the values of a H^1 real-valued function on hypersurfaces, in particular to give a meaning to the values of such a function on the boundary of a domain. As we want to define the Dirichlet problem, which consists in minimizing the Dirichlet energy with fixed values on the boundary $\partial\Omega$, we need to give a meaning to the boundary values of elements of $H^1(\Omega, \mathcal{P}(D))$. Koorevaar and Schoen have already developped a trace theory in a fairly general context [19, Section 1.12]. However, in our specific situation and in view of proving the dual formulation of the Dirichlet problem, we will define the boundary values by showing how one can extend the continuity equation for test functions $\varphi \in C^1(\Omega \times D, \mathbf{R}^p)$ which are no longer compactly supported in $\mathring{\Omega}$. Even if we do not prove it in this article, our definition of trace coincides with the one of [19]: to be convinced one can look at Proposition 6.6 and compare it to [19, Theorem 1.12.3]. Recall that $\mathbf{n}_\Omega$ denotes the outward normal to $\partial\Omega$.

Theorem 3.25. *Let $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$. Then there exists a vector-valued measure $\mathrm{BT}_{\boldsymbol{\mu}} \in \mathcal{M}(\Omega \times D, \mathbf{R}^p)$ supported on $\mathbf{n}_\Omega \partial\Omega \times D$ (which means that $\mathrm{BT}_{\boldsymbol{\mu}}(\varphi) = 0$ if $\varphi \cdot \mathbf{n}_\Omega = 0$ on $\partial\Omega \times D$) such that for any $\varphi \in C^1(\Omega \times D, \mathbf{R}^p)$ and for any $E \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ for which $(\boldsymbol{\mu}, E)$ satisfies the continuity equation and $\mathrm{Dir}(\boldsymbol{\mu}, E) < +\infty$,*

$$\int_{\Omega\times D} \nabla_\Omega \cdot \varphi \mathrm{d}\boldsymbol{\mu} + \int_{\Omega\times D} \nabla_D \varphi \cdot \mathrm{d}E = \mathrm{BT}_{\boldsymbol{\mu}}(\varphi). \tag{3.14}$$

Moreover if $\boldsymbol{\mu}$ is continuous then for any $\varphi \in C^1(\Omega \times D, \mathbf{R}^p)$,

$$\mathrm{BT}_{\boldsymbol{\mu}}(\varphi) = \int_{\partial\Omega} \left(\int_D \varphi(\xi, x) \cdot \mathbf{n}_\Omega(\xi) \boldsymbol{\mu}(\xi, \mathrm{d}x) \right) \sigma(\mathrm{d}\xi),$$

where σ is the surface measure on $\partial\Omega$.

$\mathrm{BT}_{\boldsymbol{\mu}}$ stands for "Boundary Term" of $\boldsymbol{\mu}$. It is not surprising that, if $\boldsymbol{\mu}$ is continuous, the value of $\mathrm{BT}_{\boldsymbol{\mu}}$ depends only on the values of $\boldsymbol{\mu}$ on the boundary.

Proof. Take $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ and $E = \mathbf{v}\boldsymbol{\mu} \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ such that $(\boldsymbol{\mu}, E)$ satisfies the continuity equation and $\mathrm{Dir}(\boldsymbol{\mu}, E) < +\infty$. The l.h.s. of (3.14) defines a vector-valued distribution on $\Omega \times D$ acting on φ. We need to show that it is of order 0 and that it does not depend on E.

We define $f : \Omega \to \mathbf{R}^p$ by, for a.e. $\xi \in \Omega$,

$$f(\xi) := \int_D \varphi(\xi, x)\boldsymbol{\mu}(\xi, \mathrm{d}x).$$

Using the continuity equation with test functions of the form $\chi\varphi^\alpha$, for $\chi \in C^1_c(\mathring{\Omega}, \mathbf{R}^p)$ and $\alpha \in \{1, 2, \dots, p\}$, one can see that $f \in H^1(\mathring{\Omega}, \mathbf{R}^p)$ and

$$\partial_\alpha f^\beta(\xi) = \int_D \partial_\alpha \varphi^\beta(\xi, x)\boldsymbol{\mu}(\xi, \mathrm{d}x) + \int_D \nabla_D \varphi^\beta(\xi, x) \cdot \mathbf{v}^\alpha(\xi, x)\boldsymbol{\mu}(\xi, \mathrm{d}x).$$

for all $\alpha, \beta \in \{1, 2, \dots, p\}$. In particular f admits on $\partial\Omega$ a trace $\bar{f} : \partial\Omega \to \mathbf{R}^p$. We apply the divergence theorem: one can find in [12, Section 4.3] a statement when $\partial\Omega$ is only Lipschitz and f has Sobolev regularity. In our case, given the expression of ∇f, it reads

$$\int_{\Omega\times D} \nabla_\Omega \cdot \varphi \mathrm{d}\boldsymbol{\mu} + \int_{\Omega\times D} \nabla_D \varphi \cdot \mathrm{d}E = \int_{\partial\Omega} \bar{f}(\xi) \cdot \mathbf{n}_\Omega(\xi)\sigma(\mathrm{d}\xi) \tag{3.15}$$

where $\mathbf{n}_\Omega$ is the outward normal to $\partial\Omega$ and σ its the surface measure. In particular we see that the r.h.s. of (3.14) does not depend on E. Moreover, as $\|f\|_\infty \leqslant \|\varphi\|_\infty$, the same L^∞ bounds holds for $\bar{f}$, thus

$$\left| \int_{\partial\Omega} \bar{f}(\xi) \cdot \mathbf{n}_\Omega(\xi)\sigma(\mathrm{d}\xi) \right| \leqslant \sigma(\partial\Omega)\|\varphi\|_\infty.$$

It allows to conclude that the r.h.s. of (3.14) is a distribution of order 0 acting on φ, hence it can be represented by a measure $\mathrm{BT}_{\boldsymbol{\mu}} \in \mathcal{M}(\Omega \times D, \mathbf{R}^p)$. From (3.15) it is clear that $\mathrm{BT}_{\boldsymbol{\mu}}$ is supported on $\mathbf{n}_\Omega\partial\Omega \times D$.

If we assume moreover that $\boldsymbol{\mu}$ is continuous, so is f. Indeed, for any $\xi, \eta \in \Omega$,

$$\begin{aligned}
|f(\xi) - f(\eta)| &= \left| \int_D \varphi(\xi, x)\boldsymbol{\mu}(\xi, \mathrm{d}x) - \int_D \varphi(\eta, x)\boldsymbol{\mu}(\xi, \mathrm{d}x \right| \\
&\leqslant \int_D |\varphi(\xi, x) - \varphi(\eta, x)|\boldsymbol{\mu}(\xi, \mathrm{d}x) + \left| \int_D \varphi(\eta, x)\boldsymbol{\mu}(\xi, \mathrm{d}x) - \int_D \varphi(\eta, x)\boldsymbol{\mu}(\eta, \mathrm{d}x) \right| \\
&\leqslant \|\nabla_\Omega \varphi\|_\infty |\xi - \eta| + \left| \int_D \varphi(\eta, x)\boldsymbol{\mu}(\xi, \mathrm{d}x) - \int_D \varphi(\eta, x)\boldsymbol{\mu}(\eta, \mathrm{d}x) \right|.
\end{aligned}$$

When $\xi \to \eta$, the first term obviously goes to 0, and the second one too by definition of the weak convergence (by assumption $\boldsymbol{\mu}(\xi) \to \boldsymbol{\mu}(\eta)$ in the weak sense). Thus $\bar{f}$ coincides with f, which gives the announced result. □

If $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$, using the disintegration theorem and testing against well chosen functions, one can show that there exists $\bar{\boldsymbol{\mu}} : \partial\Omega \to \mathcal{P}(D)$ defined σ-a.e. such that $\mathrm{BT}_{\boldsymbol{\mu}} = \mathbf{n}_\Omega\bar{\boldsymbol{\mu}} \otimes \sigma$. The mapping $\bar{\boldsymbol{\mu}}$ can be seen as a definition of the values of $\boldsymbol{\mu}$ on $\partial\Omega$.

Now we can define what it means to share the same boundary values and prove that the set of $\boldsymbol{\mu}$ with fixed boundary values is closed.

Definition 3.26. *Let $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ two elements of $H^1(\Omega, \mathcal{P}(D))$. We say that $\boldsymbol{\mu}|_{\partial\Omega} = \boldsymbol{\nu}|_{\partial\Omega}$ if $\mathrm{BT}_{\boldsymbol{\mu}} = \mathrm{BT}_{\boldsymbol{\nu}}$.*

Proposition 3.27. *Let $\boldsymbol{\mu}_b \in H^1(\Omega, \mathcal{P}(D))$ and $C \in \mathbf{R}$ be fixed. Then the set*

$$\{\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D)) \ : \ \boldsymbol{\mu}|_{\partial\Omega} = \boldsymbol{\mu}_b|_{\partial\Omega} \text{ and } \mathrm{Dir}(\boldsymbol{\mu}) \leqslant C\}$$

is closed for the weak topology on $L^2(\Omega, \mathcal{P}(D))$.

Proof. The proof is straightforward. Indeed, take a sequence $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ in $\in L^2(\Omega,\mathcal{P}(D))$ such that $\boldsymbol{\mu}_n|_{\partial\Omega} = \boldsymbol{\mu}_b|_{\partial\Omega}$ and $\mathrm{Dir}(\boldsymbol{\mu}_n) \leqslant C$ for any $n \in \mathbf{N}$, and assume it converges weakly to some $\boldsymbol{\mu} \in L^2(\Omega,\mathcal{P}(D))$. By lower semi-continuity of Dir, we know that $\mathrm{Dir}(\boldsymbol{\mu}) \leqslant C$. For any $n \in \mathbf{N}$ choose $E_n \in \mathcal{M}(\Omega\times D,\mathbf{R}^{pq})$ tangent to $\boldsymbol{\mu}_n$, similarly take E_b tangent to $\boldsymbol{\mu}_b$. The identity $\boldsymbol{\mu}_n|_{\partial\Omega} = \boldsymbol{\mu}_b|_{\partial\Omega}$ can be written: for every $\varphi \in C^1(\Omega\times D,\mathbf{R}^p)$

$$\int_{\Omega\times D} \nabla\cdot\varphi \mathrm{d}\boldsymbol{\mu}_n + \int_{\Omega\times D} \nabla_D\varphi\cdot \mathrm{d}E_n = \int_{\Omega\times D} \nabla\cdot\varphi \mathrm{d}\boldsymbol{\mu}_b + \int_{\Omega\times D} \nabla_D\varphi\cdot \mathrm{d}E_b. \tag{3.16}$$

As seen in the proof of Proposition 3.11, one can assume that, up to extraction, $(E_n)_{n\in\mathbf{N}}$ weakly converges to some E. It is easy to see that $(\boldsymbol{\mu},E)$ satisfies the continuity equation and that $\mathrm{Dir}(\boldsymbol{\mu},E)\leqslant C<+\infty$. Thus, we can pass to the limit in (3.16) and see that for any $\varphi\in C^1(\Omega\times D,\mathbf{R}^p)$,

$$\int_{\Omega\times D} \nabla\cdot\varphi \mathrm{d}\boldsymbol{\mu} + \int_{\Omega\times D} \nabla_D\varphi\cdot \mathrm{d}E = \int_{\Omega\times D} \nabla\cdot\varphi \mathrm{d}\boldsymbol{\mu}_b + \int_{\Omega\times D} \nabla_D\varphi\cdot \mathrm{d}E_b,$$

which exactly means that $\boldsymbol{\mu}|_{\partial\Omega} = \boldsymbol{\mu}_b|_{\partial\Omega}$. □

4. The Dirichlet problem and its dual

4.1. **Statement of the problem.** With all the tools at our disposal, we are ready to state the Dirichlet problem. It simply consists in minimizing the Dirichlet energy under the constraint that the values at the boundary are fixed.

Definition 4.1. *Let $\boldsymbol{\mu}_b \in H^1(\Omega,\mathcal{P}(D))$. Then the Dirichlet problem with boundary values $\boldsymbol{\mu}_b$ is defined as*

$$\min\left\{\mathrm{Dir}(\boldsymbol{\mu}) \ :\ \boldsymbol{\mu}\in H^1(\Omega,\mathcal{P}(D)) \text{ and } \boldsymbol{\mu}|_{\partial\Omega} = \boldsymbol{\mu}_b|_{\partial\Omega}\right\}.$$

A mapping $\boldsymbol{\mu}\in H^1(\Omega,\mathcal{P}(D))$ which realizes the minimum is called a solution of the Dirichlet problem.

Definition 4.2. *Let $\boldsymbol{\mu}\in H^1(\Omega,\mathcal{P}(D))$. We say that $\boldsymbol{\mu}$ is harmonic if it is a solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}$.*

With the work of the previous section, the existence of at least one solution is a straightforward application of the direct method of calculus of variations.

Theorem 4.3. *Let $\boldsymbol{\mu}_b\in H^1(\Omega,\mathcal{P}(D))$. Then there exists at least one solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_b$.*

Proof. There exists at least one $\boldsymbol{\mu}$ with finite Dirichlet energy which satisfies the boundary conditions, namely $\boldsymbol{\mu}_b$. Thus, one can consider a minimizing sequence $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$. By compactness of $L^2(\Omega,\mathcal{P}(D))$, we can assume, up to extraction, that this sequence converges weakly to some $\boldsymbol{\mu}\in L^2(\Omega,\mathcal{P}(D))$. By Proposition 3.27, we know that $\boldsymbol{\mu}$ also satisfies $\boldsymbol{\mu}|_{\partial\Omega} = \boldsymbol{\mu}_b|_{\partial\Omega}$. The lower semi-continuity of Dir allows to conclude that $\boldsymbol{\mu}$ is a minimizer of Dir. □

Let us spend a few words about the question of uniqueness. Of course, the proof above provides no information about it. By convexity of the Dirichlet energy (Proposition 3.11), we know that the set of solutions of the Dirichlet problem is convex. Recall that if $\Omega=[0,1]$ is a segment of $\mathbf{R}$, then the Dirichlet problem reduces to the problem of finding a geodesic between the two endpoints $\boldsymbol{\mu}_b(0)$ and $\boldsymbol{\mu}_b(1)$. It is well known that a sufficient condition for uniqueness is to impose that either $\boldsymbol{\mu}_b(0)$ or $\boldsymbol{\mu}_b(1)$ are absolutely continuous w.r.t. $\mathcal{L}_D$, and there can be non uniqueness when it is not the case (see for instance [27, Chapter 5]). Hence, it would natural, in order to investigate the question of uniqueness, to impose that for every $\xi\in\partial\Omega$, the measure $\boldsymbol{\mu}_b(\xi)$ is absolutely continuous w.r.t. $\mathcal{L}_D$. We do not know if uniqueness holds under this hypothesis: a difference with the case where Ω is a segment is the fact that we do not know a static or Lagrangian formulation. In other words,

we do not know the equivalent of transport plans, which in the case of a 1-dimensional Ω, allow to parametrize geodesics and to greatly simply the problem. However we are able to prove uniqueness in a non trivial case: the one of a family of elliptically contoured distributions treated in Subsection 7.2, see also the introduction where the strategy of the proof is discussed.

4.2. **Lipschitz extension.** To give ourselves the boundary conditions, we need a mapping $\boldsymbol{\mu}_b$ defined on the whole Ω, even though only its values near $\partial\Omega$ will play a role. Thus a natural question arises: if $\boldsymbol{\mu}_b$ is only defined on $\partial\Omega$, is it possible to extend it on Ω? The next theorem shows that the answer is positive in the case where $\boldsymbol{\mu}_b$ is Lipschitz on $\partial\Omega$. Indeed, in this case we can build an extension which is Lipschitz on Ω, thus in $H^1(\Omega, \mathcal{P}(D))$ thanks to Theorem 3.17.

Theorem 4.4. *Let $\boldsymbol{\mu}_l : \partial\Omega \to \mathcal{P}(D)$ a lipschitz mapping. Then there exists $\boldsymbol{\mu} : \Omega \to \mathcal{P}(D)$ Lipschitz such that $\boldsymbol{\mu}(\xi) = \boldsymbol{\mu}_l(\xi)$ for every $\xi \in \partial\Omega$.*

For a continuous $\boldsymbol{\mu}$ the boundary term $\mathrm{BT}_{\boldsymbol{\mu}}$ depends only on the values of $\boldsymbol{\mu}$ on $\partial\Omega$ (Theorem 3.25), hence the boundary term of the Lipschitz extension of $\boldsymbol{\mu}_l : \partial\Omega \to \mathcal{P}(D)$ does not depend on the extension. In other words, the following problem is well defined:

Definition 4.5. *Let $\boldsymbol{\mu}_l : \partial\Omega \to \mathcal{P}(D)$ a Lipschitz mapping. Then the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$ is defined as the Dirichlet problem with boundary values $\boldsymbol{\mu}_b$, where $\boldsymbol{\mu}_b$ is any Lipschitz extension of $\boldsymbol{\mu}_l$ on Ω.*

Now, let us prove the Lipschitz extension theorem. It relies on the following Lemma, which allows to treat the case where Ω is a ball.

Lemma 4.6. *Let $B(0,1)$ be the unit ball of $\mathbf{R}^p$ and $\mathbf{S}^{p-1} := \partial B(0,1)$ its boundary. Let $\boldsymbol{\mu}_l : \mathbf{S}^{d-1} \to \mathcal{P}(D)$ a Lipschitz mapping and take $x_0 \in D$. Define, for any $r \in [0,1]$ the map $T_r : D \to D$ by $T_r(x) = rx + (1-r)x_0$. Then the mapping $\boldsymbol{\mu} : B(0,1) \to \mathcal{P}(D)$ defined by*

$$\boldsymbol{\mu}(r\xi) := T_r \# [\boldsymbol{\mu}(\xi)]$$

for any $r \in [0,1]$ and any $\xi \in \mathbf{S}^{d-1}$ is Lipschitz.

Proof. If $\xi \in \mathbf{S}^{d-1}$ is fixed, then $r \in [0,1] \mapsto \boldsymbol{\mu}(r\xi)$ is the constant speed geodesic joining δ_{x_0} to $\boldsymbol{\mu}_l(\xi)$. Hence, we can write that $W_2(\boldsymbol{\mu}(r\xi), \boldsymbol{\mu}(s\xi)) \leqslant C|r-s|$, where C depends only on the diameter of $\mathcal{P}(D)$. On the other hand, as T_r is r-Lipschitz in D, then $\nu \mapsto T_r \# \nu$ is also r-Lipschitz in $\mathcal{P}(D)$. Hence, for any ξ and η in $\mathbf{S}^{d-1}$, one has $W_2(\boldsymbol{\mu}(r\xi), \boldsymbol{\mu}(r\eta)) \leqslant Cr|\xi - \eta|$, where C is the Lipschitz constant of $\boldsymbol{\mu}_l$. Putting the two estimates together, we deduce that for any $r, s \in [0,1]$ and any $\xi, \eta \in \mathbf{S}^{p-1}$,

$$W_2(\boldsymbol{\mu}(r\xi), \boldsymbol{\mu}(s\eta)) \leqslant C[|r-s| + \min(r,s)|\xi - \eta|],$$

which is enough to conclude that $\boldsymbol{\mu}$ is Lipschitz. □

Notice that the Lipschitz constant of the extension is not controlled by the Lipschitz constant of $\boldsymbol{\mu}_l$: the distance between δ_{x_0} and the range of $\boldsymbol{\mu}_l$ also plays a role as $\boldsymbol{\mu}(0) = \delta_{x_0}$. Hence, we cannot use a decomposition with Withney cubes to extend mappings defined on arbitrary closed subsets Ω, but only on the boundary of smooth sets: basically we need to use Lemma 4.6 only a finite number of times.

Proof of Theorem 4.4. We will use Lemma 4.6 in the following form: if Ω is a domain which is in a bilipschitz bijection with a ball, then Theorem 4.4 holds for this domain.

We reason by induction on $p \geqslant 1$ the dimension of Ω. In dimension 1, $\Omega = I$ is a segment. To extend a mapping defined only on the boundary of the segment I, we take the constant speed geodesic in $\mathcal{P}(D)$ between the values of $\boldsymbol{\mu}_l$ at the two endpoints of I.

Now assume that the result holds for some $p-1 \geqslant 1$ and let Ω be a compact domain with Lipschitz boundary in $\mathbf{R}^p$. The goal is to cut Ω in a finite number of pieces on which Lemma 4.6 apply. For

each $\xi \in \Omega$ we choose $r_\xi > 0$ such that $B(\xi, r_\xi) \cap \Omega$ is in a bilipschitz bijection with a ball. It is obvious that we can do that for $\xi \in \mathring{\Omega}$, and for points on $\partial\Omega$ we use the fact that Ω is locally the epigraph of a Lipschitz function. By compactness, we find balls $B_1, B_2, \dots, B_N$ covering Ω such that $B_n \cap \Omega$ is in a bilipschitz bijection with a ball for any $n \in \{1, 2, \dots, N\}$. We can of course assume that B_n is not included in B_m for any $n \neq m$. Then we define recursively $X_1 := B_1 \cap \Omega$ and $X_n = (B_n \cap \Omega) \backslash \mathring{X}_{n-1}$ for $n \in \{2, \dots, N\}$. For any $n \in \{1, 2, \dots, N\}$, X_n is still in a bilipschitz bijection with a ball. On $\bigcup_n \partial X_n$, which is made of $\partial\Omega$ and of pieces of spheres of $\mathbf{R}^p$, thus locally in bilipschitz bijection with Lipschitz domains of $\mathbf{R}^{p-1}$, we can use the induction assumption and extend $\boldsymbol{\mu}_l$. Then, we use Lemma 4.6 to extend $\boldsymbol{\mu}$ on $\mathring{X}_n$ for each $n \in \{1, 2, \dots, N\}$. We have obtained a function $\boldsymbol{\mu}$ which is continuous and Lipschitz on each $X_n, n \in \{1, 2, \dots, N\}$: it is globally Lipschitz on Ω. □

4.3. **The dual problem.** We will know show a rigorous proof of the absence of duality gap. The dual problem was already obtained, at least formally, in the introduction.

Theorem 4.7. *Let $\boldsymbol{\mu}_b \in H^1(\Omega, \mathcal{P}(D))$. Then one has*

$$\sup\left\{ \mathrm{BT}_{\boldsymbol{\mu}_b}(\varphi) \ : \ \varphi \in C^1(\Omega \times D, \mathbf{R}^p) \ \textit{and} \ \nabla_\Omega \cdot \varphi + \frac{|\nabla_D \varphi|^2}{2} \leqslant 0 \ \textit{on} \ \Omega \times D \right\}$$
$$= \min\left\{ \mathrm{Dir}(\boldsymbol{\mu}) \ : \ \boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D)) \ \textit{and} \ \boldsymbol{\mu}|_{\partial\Omega} = \boldsymbol{\mu}_b|_{\partial\Omega} \right\}.$$

Proof. We rely on the Fenchel-Rockafellar duality theorem which can be found in [31, Theorem 1.9]. Let $X := C(\Omega \times D, \mathbf{R}^{1+pq})$ the space of continuous functions defined on the compact space $\Omega \times D$ and valued in $\mathbf{R}^{1+pq}$ endowed with the norm of uniform convergence. An element of X will be written (a, b), where $a \in C(\Omega \times D)$ and $b \in C(\Omega \times D, \mathbf{R}^{pq})$. The dual space $X^\star$ is, by the Riesz theorem, $\mathcal{M}(\Omega \times D, \mathbf{R}^{1+pq})$. Again an element of $X^\star$ will be written $(\boldsymbol{\mu}, E)$ where $\boldsymbol{\mu} \in \mathcal{M}(\Omega \times D)$ is a signed measure and $E \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ is a vector-valued measure. We introduce the functionals $F : X \to \mathbf{R}$ and $G : X \to \mathbf{R}$ defined as, for any $(a, b) \in X$,

$$F(a,b) = \begin{cases} 0 & \text{if } a(\xi,x) + \dfrac{|b(\xi,x)|^2}{2} \leqslant 0 \text{ for every } (\xi,x) \in \Omega \times D \\ +\infty & \text{else,} \end{cases}$$

$$G(a,b) = \begin{cases} -\mathrm{BT}_{\boldsymbol{\mu}_b}(\varphi) & \text{if } (a,b) = (\nabla_\Omega \cdot \varphi, \nabla_D \varphi) \text{ for some } \varphi \in C^1(\Omega \times D, \mathbf{R}^p) \\ +\infty & \text{else.} \end{cases}$$

Notice that thanks to (3.14), G is well defined and does not depend on the choice of φ such that $(a, b) = (\nabla_\Omega \cdot \varphi, \nabla_D \varphi)$. Notice also that at the point $(-1, 0) \in X$, one has that F is finite and continuous and that G is finite (take $\varphi(\xi, x) := (-\xi^1, 0, 0, \dots, 0)$, where ξ^1 is the first component of ξ). As moreover F and G are convex, one can apply Fenchel-Rockafellar duality which means

$$-\min_{(\boldsymbol{\mu},E)\in X^\star} [F^\star(\boldsymbol{\mu}, E) + G^\star(-\boldsymbol{\mu}, -E)] = \inf_X (F + G)$$
$$= -\sup\left\{ \mathrm{BT}_{\boldsymbol{\mu}_b}(\varphi) \ : \ \varphi \in C^1(\Omega \times D, \mathbf{R}^p) \ \text{ and } \ \nabla_\Omega \cdot \varphi + \frac{|\nabla_D \varphi|^2}{2} \leqslant 0 \right\},$$

where the last inequality is just a rewriting of the definition of F and G. Let us compute $F^\star(\boldsymbol{\mu}, E)$. By definition,

$$F^\star(\boldsymbol{\mu}, E) = \sup\left\{ \int_{\Omega\times D} a \mathrm{d}\boldsymbol{\mu} + \int_{\Omega\times D} b \cdot \mathrm{d}E \ : (a,b) \in C(\Omega \times D, \mathcal{K}) \right\},$$

where $\mathcal{K}$ is defined in Definition 3.5. In particular, if $\boldsymbol{\mu}$ is not a positive measure, then choosing suitable negative a, one sees that $F^\star(\boldsymbol{\mu}, E) = +\infty$. Moreover, if $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$ and $(\boldsymbol{\mu}, E)$ satisfies

the continuity equation, then $F^\star(\boldsymbol{\mu}, E) = \mathrm{Dir}(\boldsymbol{\mu}, E)$: this is precisely Definition 3.5. On the other hand, we can compute $G^\star$: for any $(\boldsymbol{\mu}, E) \in X^\star$,

$$G^\star(-\boldsymbol{\mu}, -E) = \sup_{\varphi \in C^1(\Omega \times D, \mathbf{R}^p)} \left(\mathrm{BT}_{\boldsymbol{\mu}_b}(\varphi) - \int_{\Omega \times D} \nabla_\Omega \cdot \varphi \mathrm{d}\boldsymbol{\mu} - \int_{\Omega \times D} \nabla_D \varphi \cdot \mathrm{d}E \right).$$

By linearity of the expression inside the $\sup$ w.r.t. φ, we see that $G^\star(-\boldsymbol{\mu}, -E) < +\infty$ if and only if $G^\star(-\boldsymbol{\mu}, -E) = 0$, which translates in

$$\mathrm{BT}_{\boldsymbol{\mu}_b}(\varphi) = \int_{\Omega \times D} \nabla_\Omega \cdot \varphi \mathrm{d}\boldsymbol{\mu} + \int_{\Omega \times D} \nabla_D \varphi \cdot \mathrm{d}E$$

for every $\varphi \in C^1(\Omega \times D, \mathbf{R}^p)$. Let $a \in C(\Omega)$ a continuous function. It can always be written $a = \nabla_\Omega \cdot \varphi$, where $\varphi \in C^1(\Omega, \mathbf{R}^p)$ (take $\varphi = \nabla f$ where f solves $\Delta f = a$), thus using the fact that for such a φ,

$$\mathrm{BT}_{\boldsymbol{\mu}_b}(\varphi) = \int_{\Omega \times D} \nabla_\Omega \cdot \varphi \mathrm{d}\boldsymbol{\mu}_b = \int_{\Omega \times D} a \mathrm{d}\boldsymbol{\mu}_b = \int_\Omega a(\xi) \mathrm{d}\xi,$$

one sees that if $G^\star(-\boldsymbol{\mu}, -E) < +\infty$, then

$$\int_\Omega a(\xi) \mathrm{d}\xi = \int_{\Omega \times D} a \mathrm{d}\boldsymbol{\mu}.$$

Provided that $\boldsymbol{\mu}$ is a positive measure (recall that it happens if $F^\star(\boldsymbol{\mu}, E) < +\infty$) and by arbitrariness of a, it implies that the disintegration of $\boldsymbol{\mu}$ w.r.t. $\mathcal{L}_\Omega$ is made of probability measures on D, in other words that $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$. Once we have this information, testing with functions φ which are compactly supported on Ω, we see that if $G^\star(-\boldsymbol{\mu}, -E) < +\infty$ then $(\boldsymbol{\mu}, E)$ satisfies the continuity equation, and testing with arbitrary φ, we see that $\mathrm{BT}_{\boldsymbol{\mu}} = \mathrm{BT}_{\boldsymbol{\mu}_b}$. In the end, one concludes that

$$\min_{(\boldsymbol{\mu}, E) \in X^\star} [F^\star(\boldsymbol{\mu}, E) + G^\star(-\boldsymbol{\mu}, -E)] = \min \left\{ \mathrm{Dir}(\boldsymbol{\mu}) \ : \ \boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D)) \text{ and } \boldsymbol{\mu}|_{\partial\Omega} = \boldsymbol{\mu}_b|_{\partial\Omega} \right\}. \qquad \square$$

A natural question which arises is the existence of an optimal $\varphi \in C^1(\Omega \times D, \mathbf{R}^p)$ (or in a space of less regular functions). If Ω is a segment, the constraint on φ translates into the Hamilton-Jacobi equation

$$\partial_t \varphi + \frac{|\nabla \varphi|^2}{2} \leqslant 0,$$

whose explicit solutions are known. In particular, one can parametrize the function φ by its value at the initial time, the unknown becomes a scalar function defined on D. In the compact case, by the double convexification trick, one can get compactness in a maximizing sequence. In our case, the constraint reads in full coordinate

$$\sum_{\alpha=1}^p \partial_\alpha \varphi^\alpha + \frac{1}{2} \sum_{\alpha=1}^p \sum_{i=1}^q |\partial_i \varphi^\alpha|^2 \leqslant 0.$$

Now we do not know if one can parametrize a φ which satisfies the constraint by its values on the boundary of $\partial\Omega$ (and even if it were the case, on which part of the boundary?). Moreover, notice that the function φ is vector-valued, though the constraint involves only one scalar equation: to get compactness out of it seems more complicated. We have not investigated deeply the question of the existence of an optimal φ, but we believe that it can be substantially more complicated than in the case where Ω is a segment of $\mathbf{R}$.

5. Failure of the superposition principle

5.1. **The superposition principle.** In this section, we want to explain why a powerful tool to study curves valued in the Wasserstein space (i.e. the case where Ω is a segment of $\mathbf{R}$), namely the superposition principle, fails in higher dimensions. To say it briefly, there is no Lagrangian point of view for mappings into the Wasserstein space, one has to work only with the Eulerian one. Notice that the question of the existence of a superposition principle was already formulated by Brenier [9, Problem 3.4], but left unanswered. As we want to prove a negative result, we will not only provide a counterexample to the superposition principle, but also try to explain the obstruction and why this principle fails for all but few exceptional cases. Let us first recall the superposition principle for absolutely continuous curves.

The set Ω will be replaced by the unit segment $I = [0,1]$. As stated in Proposition 3.8, set $H^1(I,\mathcal{P}(D))$ coincides with the set of absolutely continuous curves. We denote by $\Gamma = C(I,D)$ the set of continuous curves valued in D endowed with the norm of uniform convergence, it is a polish space. If $f \in \Gamma$, then $\dot f$ denotes the derivative w.r.t. time of f provided that it exists. For any $t \in I$, $e_t : \Gamma \to D$ is the evaluation operator, which means $e_t(f) = f(t)$ for any $f \in \Gamma$. The following result can be found in [4, Section 8.2].

Theorem 5.1. *Let $\boldsymbol{\mu} \in H^1(I,\mathcal{P}(D))$. Then there exists a probability measure $Q \in \mathcal{P}(\Gamma)$ such that*

(i) *for any $t \in I$, $e_t \# Q = \boldsymbol{\mu}(t)$;*
(ii) *the following equality holds:*

$$\mathrm{Dir}(\boldsymbol{\mu}) = \int_\Gamma \left(\int_I \frac{1}{2} |\dot f(t)|^2 \mathrm{d}t \right) Q(\mathrm{d}\gamma).$$

The measure Q can be seen as a multimarginal transport plan coupling all the different instants, whose 2-marginals are almost optimal transport plans if they are taken between two very close instants. In other words, for any t and s in I, $(e_s,e_t)\#Q$ is a transport plan between $\boldsymbol{\mu}(s)$ and $\boldsymbol{\mu}(t)$ (by (i)), and it is almost an optimal transport plan if s is very close to t by (ii).

Another way to see it is the following: if $f \in \Gamma$, then we can also see it as an element $\boldsymbol{\mu}_f$ of $H^1(I,\mathcal{P}(D))$. Indeed, just set $\boldsymbol{\mu}(t) = \delta_{f(t)}$ for any $t \in I$, and one can define $E_f \in \mathcal{M}(I \times D, \mathbf{R}^q)$ by, for any $b \in C(I \times D, \mathbf{R}^q)$,

$$\int_{I\times D} b \cdot \mathrm{d}E_f := \int_I b(t,f(t)) \cdot \dot f(t) \mathrm{d}t.$$

With this choice, one can check that

$$\mathrm{Dir}(\boldsymbol{\mu}_f) = \mathrm{Dir}(\boldsymbol{\mu}_f, E_f) = \int_I \frac{1}{2} |\dot f(t)|^2 \mathrm{d}t.$$

Then, Theorem 5.1 is saying that there exists $Q \in \mathcal{P}(\Gamma)$ such that $\boldsymbol{\mu}$ is the mean w.r.t. Q of the $\boldsymbol{\mu}_f$'s (this is (i)), and such the E which is tangent to $\boldsymbol{\mu}$ is the mean w.r.t. Q of the E_f (indeed, thanks to Jensen's inequality, (ii) translates in the fact that $\mathrm{Dir}(\boldsymbol{\mu}) = \mathrm{Dir}(\boldsymbol{\mu}, \int_\Gamma E_f Q(\mathrm{d}f))$).

Let us try to see what a superposition principle would look like if the dimension of Ω is larger than 1. We denote by $\mathcal{F}$ the space $L^2(\Omega,D)$ which is a polish space. As it was already done in [9], if $f \in H^1(\Omega,D)$, then we can see it as an element $\boldsymbol{\mu}_f$ of $H^1(\Omega,\mathcal{P}(D))$ by setting $\boldsymbol{\mu}_f(\xi) := \delta_{f(\xi)}$. In other words, a classical function can be seen as a mapping valued in the Wasserstein space by identifying $f(\xi) \in D$ with $\delta_{f(\xi)} \in \mathcal{P}(D)$. More precisely, we define $\boldsymbol{\mu}_f \in L^2(\Omega,\mathcal{P}(D))$ and $E_f \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ by,

for any $a \in C(\Omega \times D)$ and $b \in C(\Omega \times D, \mathbf{R}^{pq})$,

$$\int_{\Omega\times D} a \mathrm{d}\boldsymbol{\mu}_f := \int_\Omega a(\xi, f(\xi))\mathrm{d}\xi,$$
$$\int_{\Omega\times D} b \cdot \mathrm{d}E_f := \int_\Omega b(\xi, f(\xi)) \cdot \nabla f(\xi)\mathrm{d}\xi.$$

Proposition 5.2. *If $f \in H^1(\Omega, D)$, and if $\boldsymbol{\mu}_f$ and E_f are defined as above, then E_f is tangent to $\boldsymbol{\mu}_f$ and*

$$\mathrm{Dir}(\boldsymbol{\mu}_f) = \mathrm{Dir}(\boldsymbol{\mu}_f, E_f) = \int_\Omega \frac{1}{2}|\nabla f(\xi)|^2 \mathrm{d}\xi.$$

Proof. To check the first part, take $\varphi \in C^1_c(\mathring{\Omega} \times D, \mathbf{R}^p)$. Defining $\tilde{\varphi} \in H^1(\Omega, \mathbf{R}^p)$ by $\tilde{\varphi}(\xi) = \varphi(\xi, f(\xi))$, we have that $\tilde{\varphi}$ is compactly supported in $\mathring{\Omega}$ and

$$\nabla \cdot \tilde{\varphi} = (\nabla_\Omega \cdot \varphi)(\xi, f(\xi)) + (\nabla_D \varphi)(\xi, f(\xi)) \cdot \nabla f(\xi).$$

Integrating this identity w.r.t. Ω, as the l.h.s. vanishes by compactness of the support of $\tilde{\varphi}$, we see that we can conclude that $(\boldsymbol{\mu}_f, E_f)$ satisfies the continuity equation.

Notice that E_f has a density $\mathbf{v}_f \in L^2_{\boldsymbol{\mu}_f}(\Omega \times D, \mathbf{R}^{pq})$ w.r.t. $\boldsymbol{\mu}$ given by $\mathbf{v}_f(\xi, x) = \nabla f(\xi)$. In particular, for a fixed ξ, $\mathbf{v}_f(\xi, \cdot)$ is constant hence the gradient of a function. Using Corollary 3.21, one sees that it is enough to conclude that E_f is tangent. Moreover, as $\mathbf{v}_f$ does not depend on x,

$$\mathrm{Dir}(\boldsymbol{\mu}_f) = \mathrm{Dir}(\boldsymbol{\mu}_f, E_f) = \int_{\Omega\times D} \frac{1}{2}|\mathbf{v}_f(\xi)|^2 \boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x) = \int_\Omega \frac{1}{2}|\mathbf{v}_f(\xi)|^2 \mathrm{d}\xi = \int_\Omega \frac{1}{2}|\nabla f(\xi)|^2 \mathrm{d}\xi. \qquad \square$$

We mention that [9, Theorem 3.1] states that if $f : \Omega \to D$ is a (classical) harmonic map, then $\boldsymbol{\mu}_f$ is also an harmonic mapping. To prove such a fact, Brenier showed how one can build a solution of the dual problem (with boundary values $\boldsymbol{\mu}_f$) from the function f.

By analogy, the superposition principle would read as follows: If $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ and $E \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ is tangent to $\boldsymbol{\mu}$, does there exist $Q \in \mathcal{P}(\mathcal{F})$ such that $\boldsymbol{\mu}$ is the mean of $\boldsymbol{\mu}_f$ w.r.t. Q and E is the mean of E_f w.r.t. Q? Thanks to Jensen's inequality and the uniqueness of the tangent momentum, the second condition can in fact be rewritten as

$$\mathrm{Dir}(\boldsymbol{\mu}) = \mathrm{Dir}(\boldsymbol{\mu}, E) = \int_{\mathcal{F}} \mathrm{Dir}(\boldsymbol{\mu}_f, E_f) Q(\mathrm{d}f) = \int_{\mathcal{F}} \left(\int_\Omega \frac{1}{2}|\nabla f(\xi)|^2 \mathrm{d}\xi \right) Q(\mathrm{d}f).$$

These considerations can be summarized by the following definition, which is the same as [9, Problem 3.4]. For $f \in \mathcal{F}$ we define its "classical" Dirichlet energy $\mathrm{Dir}_c(f)$ by

$$\mathrm{Dir}_c(f) = \begin{cases} \displaystyle\int_\Omega \frac{1}{2}|\nabla f(\xi)|^2 \mathrm{d}\xi & \text{if } f \in H^1(\Omega, D), \\ +\infty & \text{else.} \end{cases}$$

Definition 5.3. *Let $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$. We say that $\boldsymbol{\mu}$ admits a superposition principle if there exists $Q \in \mathcal{P}(\mathcal{F})$ such that*

(i) *for any $a \in C(\Omega \times D)$;*

$$\int_{\Omega\times D} a \mathrm{d}\boldsymbol{\mu} = \int_{\mathcal{F}} \left(\int_\Omega a(\xi, f(\xi))\mathrm{d}\xi \right) Q(\mathrm{d}f),$$

(ii) *the following identity holds:*

$$\int_{\mathcal{F}} \mathrm{Dir}_c(f) Q(\mathrm{d}f) \leqslant \mathrm{Dir}(\boldsymbol{\mu}).$$

In particular, with our definition, if Q represents $\boldsymbol{\mu} \in H^1(\Omega, D)$, then for Q-a.e. function f one has $\mathrm{Dir}_c(f) < +\infty$ hence f belongs to $H^1(\Omega, D)$. Let us underline that (i) is heuristically the same as (i) of Theorem 5.1, but in a form integrated over Ω because the evaluation operator does not make sense in higher dimensions: the elements of $\mathcal{F}$ are not necessarily continuous. In Definition 5.3, if (i) and (ii) holds, then the inequality in (ii) is in fact an equality because the reverse inequality always holds. Indeed, if $\boldsymbol{\mu}$ satisfies the superposition principle, we can say that $\boldsymbol{\mu} = \int_{\mathcal{F}} \boldsymbol{\mu}_f Q(\mathrm{d}f)$. By convexity of the Dirichlet energy (Proposition 3.11), we can apply Jensen's inequality, thus

$$\mathrm{Dir}(\boldsymbol{\mu}) \leqslant \int_{\mathcal{F}} \mathrm{Dir}(\boldsymbol{\mu}_f) Q(\mathrm{d}f) = \int_{\mathcal{F}} \mathrm{Dir}_c(f) Q(\mathrm{d}f).$$

5.2. **Counterexample.** We will first provide a counterexample which we will try to make as generic as possible. In what follows, we take $\Omega := \mathbf{U}$ to be the unit disk of $\mathbf{R}^2$ and $\mathbf{S}^1 = \partial\mathbf{U}$ its boundary. We also take $D = \mathbf{U}$. We view $\mathbf{U}$ as a subset of the complex plane $\mathbf{C}$: multiplication on $\mathbf{U}$ means complex multiplication.

Let $\boldsymbol{\mu}_s : \mathbf{S}^1 :\to \mathcal{P}(\mathbf{U})$ be the (complex) square root: it is the mapping defined by, for $\xi \in \mathbf{S}^1$,

$$\boldsymbol{\mu}_s(\xi) := \frac{1}{2} \sum_{z^2=\xi} \delta_z = \frac{1}{2}(\delta_{\sqrt{\xi}} + \delta_{-\sqrt{\xi}}),$$

where $\sqrt{\xi}$ is a (complex) square root of ξ. The function $\boldsymbol{\mu}_s$ is clearly Lipschitz (with Lipschitz constant equals to 2). In fact, if $\xi = e^{it}$ with $t \in \mathbf{R}$, one can write

$$\boldsymbol{\mu}_s(e^{it}) = \frac{1}{2}\left(\delta_{\exp(it/2)} + \delta_{\exp(it/2+i\pi)}\right).$$

The function $t \mapsto \boldsymbol{\mu}_s(e^{it})$ is 2π-periodic, but it cannot be written as a superposition of continuous 2π-periodic functions, only 4π-periodic ones. Hence, the superpositon principle with continuous functions fails for this mapping. This example is well known in the theory of Q-functions [10], we took it from there. To our purpose, we will need the fact that the superposition principle with $H^{1/2}$ functions fails for the mapping $\boldsymbol{\mu}_s$: roughly speaking, it holds because $H^{1/2}$ functions, in dimension 1, cannot have jumps.

Lemma 5.4. *There is no function $f \in H^{1/2}(\mathbf{S}^1, \mathbf{U})$ such that $f(\xi)^2 = \xi$ for a.e. $\xi \in \mathbf{S}^1$.*

As this lemma is not directly related to harmonic mappings, we postpone its proof to Appendix B. With the help of this lemma, we can prove that no mapping $\boldsymbol{\mu} \in H^1(\mathbf{U}, \mathcal{P}(\mathbf{U}))$ such that $\boldsymbol{\mu}|_{\partial\mathbf{U}} = \boldsymbol{\mu}_s$ can have a superposition principle: indeed, if it were the case, then we could restrict the superposition to $\partial\mathbf{U}$, and we would have a superposition principle for $\boldsymbol{\mu}_s$ with functions in $H^{1/2}$ which is a contradiction. To make this argument rigorous is a bit technical given the definition we chose for the boundary values of mappings in $H^1(\mathbf{U}), \mathcal{P}(\mathbf{U}))$: $\boldsymbol{\mu}$ is not necessarily continuous.

Proposition 5.5. *Let $\boldsymbol{\mu} \in H^1(\mathbf{U}, \mathcal{P}(\mathbf{U}))$ such that $\boldsymbol{\mu}|_{\partial\mathbf{U}} = \boldsymbol{\mu}_s$. Then $\boldsymbol{\mu}$ cannot admit a superposition principle.*

Proof. We will of course reason by contradiction. We assume that there exists $Q \in \mathcal{P}(\mathcal{F})$ which satisfies the points (i) and (ii) of Definition 5.3 (in fact only point (i) will be sufficient). Let $E = \mathbf{v}\boldsymbol{\mu}$ tangent to $\boldsymbol{\mu}$. Take $\delta > 0$ and $\varepsilon > 0$. We choose $\chi_\varepsilon \in C^1([0,1])$ an increasing function supported on $[1-\varepsilon, 1]$, such that $\chi_\varepsilon(1) = 1$. Define $a_\varepsilon \in C^1(\mathbf{U}, \mathbf{R}^2)$ and $b_\delta \in C^1(\mathbf{U} \times \mathbf{U})$ by, for any $\xi, x \in \mathbf{U}$,

$$a_\varepsilon(\xi) = \frac{\xi}{|\xi|}\chi_\varepsilon(|\xi|),$$

$$b_\delta(\xi, x) = \frac{|\xi - x^2|^2}{\delta^2}.$$

In words, a_ε is a vector-valued function, parallel to lines issued from the origin, and whose norm is increasing on the annulus of radii $1-\varepsilon$ and 1 from 0 to 1. Define $A_\varepsilon = \{\xi \in \mathbf{U} \,:\, 1-\varepsilon \leqslant |\xi| \leqslant 1\}$ the annulus outside which a_ε vanishes. A simple computation gives

$$|\nabla \cdot a_\varepsilon(\xi) - \dot\chi_\varepsilon(|\xi|)| \leqslant C \mathbb{1}_{A_\varepsilon}(\xi),$$

where C does not depend on ε. On the other hand, b_δ is a smooth scalar function, which vanishes if $x^2 = \xi$, which is larger than 1 if $|x^2 - \xi| \geqslant \delta$ and whose derivative is bounded by $C\delta^{-2}$. As a test function for the continuity equation, we take $\varphi(\xi, x) = a_\varepsilon(\xi) b_\delta(\xi, x)$. With this choice, for every $\xi \in \mathbf{S}^1$, one has

$$\int_{\mathbf{U}} \varphi(\xi, x) \boldsymbol{\mu}_s(\xi, \mathrm{d}x) = \frac{1}{2} \sum_{x^2=\xi} \varphi(\xi, x) = 0.$$

Thus, $\mathrm{BT}_{\boldsymbol{\mu}_s}(\varphi) = 0$ and the continuity equation tested against φ reads

$$\left| \int_{\mathbf{U}\times\mathbf{U}} \dot\chi_\varepsilon(|\xi|) b_\delta(\xi, x) \boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x) + \int_{\mathbf{U}\times\mathbf{U}} [a_\varepsilon(\xi) \cdot \nabla_\Omega b_\delta(\xi, x) + (a_\varepsilon(\xi) \otimes \nabla_D b_\delta(\xi, x)) \cdot \mathbf{v}(\xi, x)] \boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x) \right| \leqslant C\varepsilon.$$

Indeed, in the r.h.s, the reminder of $\nabla \cdot a_\varepsilon$ of order 1 has been integrated over A_ε whose area scales like ε. For the first integral, we use the assumption that $\boldsymbol{\mu}$ satisfies the superposition principle. For the second one, we bound ∇b_δ by $C\delta^{-2}$, notice that a_ε vanishes outside A_ε and use Cauchy-Schwarz:

$$\begin{aligned}
\int_{\mathcal{F}} \left(\int_{\mathbf{U}} \dot\chi_\varepsilon(|\xi|) b_\delta(\xi, f(\xi)) \mathrm{d}\xi \right) Q(\mathrm{d}f) &= \int_{\mathbf{U}\times\mathbf{U}} \dot\chi_\varepsilon(|\xi|) b_\delta(\xi, x) \boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x) \\
&\leqslant \frac{C}{\delta^2} \int_{A_\varepsilon \times \mathbf{U}} (1 + |\mathbf{v}(\xi, x)|) \boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x) + C\varepsilon \\
&\leqslant \frac{C}{\delta^2} \sqrt{\int_{\mathbf{U}\times\mathbf{U}} (1 + |\mathbf{v}(\xi, x)|^2) \boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x)} \sqrt{\int_{A_\varepsilon \times \mathbf{U}} \boldsymbol{\mu}(\mathrm{d}\xi, \mathrm{d}x)} + C\varepsilon \\
&\leqslant \frac{C}{\delta^2} \sqrt{1 + 2\mathrm{Dir}(\boldsymbol{\mu})} \sqrt{\varepsilon} + C\varepsilon \leqslant C \frac{\sqrt{\varepsilon}}{\delta^2},
\end{aligned}$$

where C denotes a generic constant which changes from one line to another and the inequality may hold only for small ε and δ. Let us call $\mathcal{F}_{\delta,\varepsilon} \subset \mathcal{F}$ the set of $f \in \mathcal{F}$ such that

$$\int_{\mathbf{U}} \dot\chi_\varepsilon(|\xi|) |f(\xi)^2 - \xi|^2 \mathrm{d}\xi \geqslant \delta^2.$$

By Markov's inequality, one can say that

$$Q(\mathcal{F}_{\delta,\varepsilon}) = Q\left(\left\{ f \in \mathcal{F} \,:\, \int_{\mathbf{U}} \dot\chi_\varepsilon(|\xi|) b_\delta(\xi, f(\xi)) \mathrm{d}\xi \geqslant 1 \right\} \right) \leqslant \int_{\mathcal{F}} \left(\int_{\mathbf{U}} \dot\chi_\varepsilon(|\xi|) b_\delta(\xi, f(\xi)) \mathrm{d}\xi \right) Q(\mathrm{d}f) \leqslant C \frac{\sqrt{\varepsilon}}{\delta^2}.$$

Now take the sequence $\varepsilon_n := 2^{-n}$. By the previous estimate, one sees that

$$\sum_{n=1}^{+\infty} Q(\mathcal{F}_{\delta,\varepsilon_n}) < +\infty.$$

By the Borel-Cantelli lemma, one has that $Q(\limsup_n \mathcal{F}_{\delta,\varepsilon_n}) = 0$ which means that for Q-a.e. $f \in \mathcal{F}$, there exists n_0 (which may depend on f) such that

$$\int_{\mathbf{U}} \dot\chi_{\varepsilon_n}(|\xi|) |f(\xi)^2 - \xi|^2 \mathrm{d}\xi \leqslant \delta^2$$

for all $n \geqslant n_0$. Recall also that Q-a.e. f belongs to $H^1(\Omega, D)$. For such an f, sending n to $+\infty$ and by definition of the trace of f,

$$\int_{\mathbf{S}^1} |\bar{f}(\xi)^2 - \xi|^2 \sigma(\mathrm{d}\xi) \leqslant \delta^2,$$

where in this formula $\bar{f}$ stands for the trace of f on $\mathbf{S}^1$ and σ the surface measure on $\partial\mathbf{U}$. Then using this estimate for smaller and smaller δ along a countable sequence, we conclude that Q-a.e. function f satisfies $\bar{f}(\xi)^2 = \xi$ a.e. on $\mathbf{S}^1$. But on the other hand the trace of Q-a.e. function f belongs to $H^{1/2}(\mathbf{S}^1, \mathbf{U})$, which is a clear contradiction with Lemma 5.4. □

From this Proposition, we deduce that there exists an harmonic and a Lipschitz mapping $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ for which the superposition principle fails: just take respectively a solution of the Dirichlet Problem with boundary values $\boldsymbol{\mu}_s$, or a Lipschitz extension of $\boldsymbol{\mu}_s$.

Though, these examples can seem too particular and rely too much on some singular boundary conditions. To produce a stronger example, we will use the fact that, roughly speaking, the set of $\boldsymbol{\mu}$ admitting a superposition principle is stable by approximation. Thus, by contraposition, any neighborhood of a $\boldsymbol{\mu}$ which does not admit a superposition principle will contain other measures not admitting a superposition principle.

Proposition 5.6. *Let $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ a sequence of elements of $H^1(\Omega, \mathcal{P}(D))$ such that, for every $n \in \mathbf{N}$, $\boldsymbol{\mu}_n$ admits a superposition principle. We assume that $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ converges weakly to $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ and that $\lim_n \mathrm{Dir}(\boldsymbol{\mu}_n) = \mathrm{Dir}(\boldsymbol{\mu})$. Then $\boldsymbol{\mu}$ admits a superposition principle.*

Proof. For any $n \in \mathbf{N}$, let $Q_n \in \mathcal{P}(\mathcal{F})$ such that (i) and (ii) of Definition 5.3 are satisfied. By Rellich's theorem (recall that D is compact), the functional $\mathrm{Dir}_c : \mathcal{F} \to \mathbf{R}$ has compact sublevel sets in the $L^2(\Omega, D)$-topology. As

$$\sup_{n\in\mathbf{N}} \int_{\mathcal{F}} \mathrm{Dir}_c(f) Q_n(\mathrm{d}f) = \sup_{n\in\mathbf{N}} \mathrm{Dir}(\boldsymbol{\mu}_n) < +\infty,$$

we can say [4, Remark 5.1.5] that $(Q_n)_{n\in\mathbf{N}}$ is tight, hence up to extraction it weakly converges in $\mathcal{P}(\mathcal{F})$ to some $Q \in \mathcal{P}(\mathcal{F})$. We will show that Q represents $\boldsymbol{\mu}$.

Let us take $a \in C(\Omega \times D)$ and define $A : \mathcal{F} \to \mathbf{R}$ by, for any $f \in \mathcal{F}$,

$$A(f) := \int_\Omega a(\xi, f(\xi))\mathrm{d}\xi.$$

The function A is continuous for the L^2 topology. Thus, starting from

$$\int_{\mathcal{F}} A(f) Q_n(\mathrm{d}f) = \int_{\Omega\times D} a \mathrm{d}\boldsymbol{\mu}_n,$$

which is valid by Definition 5.3, we can pass both terms to the limit (recall that $\boldsymbol{\mu}_n$ weakly converges to $\boldsymbol{\mu}$) and see that $(\boldsymbol{\mu}, Q)$ satisfies (i) of Definition 5.3.

Moreover, as Dir_c is l.s.c. (for the $L^2(\Omega, D)$ topology), we can say that

$$\int_{\mathcal{F}} \mathrm{Dir}_c(f) Q(\mathrm{d}f) \leqslant \liminf_{n\to+\infty} \int_{\mathcal{F}} \mathrm{Dir}_c(f) Q_n(\mathrm{d}f) = \liminf_{n\to+\infty} \mathrm{Dir}(\boldsymbol{\mu}_n) = \mathrm{Dir}(\boldsymbol{\mu}),$$

which gives point (ii) of Definition 5.3 and concludes the proof. □

Corollary 5.7. *There exists a smooth mapping in $H^1(\mathbf{U}, \mathcal{P}(\mathbf{U}))$ (it has the same regularity than in Proposition 3.14) for which the superposition principle fails.*

The ideal result would be the existence of a smooth and harmonic mapping for which the superposition fails. However, as we lack a theory of regularity for harmonic mappings valued in the Wasserstein space, we do not know how to produce one.

Proof. Let $\bar{\boldsymbol{\mu}} \in H^1(\mathbf{U}, \mathcal{P}(\mathbf{U}))$ be a solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_s$. By Proposition 5.5, we know that the superposition principle fails for $\bar{\boldsymbol{\mu}}$. Let $(\boldsymbol{\mu}_n, \mathbf{v}_n)_{n\in\mathbf{N}}$ the sequence of approximation of $\boldsymbol{\mu}$ given by Theorem 3.13. We notice that $\boldsymbol{\mu}_n$ converges strongly to $\bar{\boldsymbol{\mu}}$ and $\mathrm{Dir}(\boldsymbol{\mu}_n)$ converges to $\mathrm{Dir}(\bar{\boldsymbol{\mu}})$ when $n \to +\infty$. Thus, using Proposition 5.6, we see that for n large enough, $\boldsymbol{\mu}_n$ does not satisfy the superposition principle. □

5.3. **Local obstruction to the superposition principle.** The counterexample provided above shows a *global* obstruction. Indeed, the mapping $\boldsymbol{\mu}_s$ can be thought locally in Ω as a superposition of classical functions, but there is a contradiction if we try to make this superposition global. On the other hand, there is also (at least formally) *local obstructions* to the superposition principle. To describe them we will stay sloppy about the regularity issues and concentrate on heuristic explanations.

Indeed, if $\boldsymbol{\mu}$ admits a superposition principle given by $Q \in \mathcal{P}(\mathcal{F})$, and if $\mathbf{v}$ is the velocity field tangent to $\boldsymbol{\mu}$, then for Q-a.e. f, one has $\nabla f(\xi) = \mathbf{v}(\xi, f(\xi))$. To prove this fact, notice that the tangent momentum $E = \mathbf{v}\boldsymbol{\mu}$ is equal to $\int_{\mathcal{F}} E_f Q(\mathrm{d}f)$ (see the discussion preceding Definition 5.3), i.e. for any $b \in C(\Omega \times D, \mathbf{R}^{pq})$,

$$\int_{\Omega\times D} b \cdot \mathrm{d}E := \int_{\mathcal{F}} \left(\int_\Omega b(\xi, f(\xi)) \cdot \nabla f(\xi) \mathrm{d}\xi \right) Q(\mathrm{d}f).$$

Thus, one can say that

$$\begin{aligned}
\mathrm{Dir}(\boldsymbol{\mu}) = \int_{\Omega\times D} \frac{1}{2}|\mathbf{v}|^2 \mathrm{d}\boldsymbol{\mu} = \int_{\Omega\times D} \frac{1}{2}\mathbf{v}\cdot \mathrm{d}E &= \int_{\mathcal{F}} \left(\int_\Omega \frac{1}{2}\mathbf{v}(\xi, f(\xi)) \cdot \nabla f(\xi) \mathrm{d}\xi \right) Q(\mathrm{d}f) \\
&\leqslant \int_{\mathcal{F}} \left(\int_\Omega \frac{1}{4} \left[|\mathbf{v}(\xi, f(\xi))|^2 + |\nabla f(\xi)|^2 \right] \mathrm{d}\xi \right) Q(\mathrm{d}f) \\
&= \frac{1}{4} \int_{\Omega\times D} |\mathbf{v}|^2 \mathrm{d}\boldsymbol{\mu} + \frac{1}{2} \int_{\mathcal{F}} \left(\int_\Omega \frac{1}{2} |\nabla f(\xi)|^2 \mathrm{d}\xi \right) Q(\mathrm{d}f) \\
&= \mathrm{Dir}(\boldsymbol{\mu}).
\end{aligned}$$

In particular, the inequality is an equality: one sees that for Q-a.e. $f \in \mathcal{F}$, one has $\nabla f(\xi) = \mathbf{v}(\xi, f(\xi))$ for a.e. $\xi \in \Omega$.

The analogue if Ω is a segment is the fact that (using notations from Theorem 5.1) for Q-a.e. f, $\dot{f}(t) = \mathbf{v}(t, f(t))$: the measure Q is supported on the flow of the vector field $\mathbf{v}$ (see [4, Theorem 8.2.1]). In dimension larger than 1, the constraint $\nabla f = \mathbf{v}(\cdot, f)$ is much stronger. In particular, it implies that along every curve $\gamma : I \to \Omega$, the function $f \circ \gamma$ follows the flow of $\mathbf{v} \cdot \dot{\gamma}$. However, there are many different curves going from one point to another: if we want all the results to be coherent, some commutation properties of the flow of $\mathbf{v}$ along different directions are needed, which turns out to be a very strong constraint. Indeed, coordinatewise, the constraint reads for every $\alpha \in \{1, 2, \dots, p\}$ and $i \in \{1, 2, \dots, q\}$,

$$\partial_\alpha f^i(\xi) = \mathbf{v}^{\alpha i}(\xi, f(\xi)).$$

If we differentiate w.r.t. β, we find that

$$\partial_{\beta\alpha} f^i(\xi) = \partial_\beta \mathbf{v}^{\alpha i}(\xi, f(\xi)) + \sum_{j=1}^q \partial_\beta f^j(\xi) \partial_j \mathbf{v}^{\alpha i}(\xi, f(\xi)) = \left(\partial_\beta \mathbf{v}^{\alpha i} + \sum_{j=1}^q \mathbf{v}^{\beta j} \partial_j \mathbf{v}^{\alpha i} \right)(\xi, f(\xi)).$$

The l.h.s is clearly symmetric if we exchange the role of α and β, so must be the r.h.s. It implies that for all $\alpha, \beta \in \{1, 2, \dots, p\}$,

$$\partial_\alpha \mathbf{v}^{\beta i} + \sum_{j=1}^q \mathbf{v}^{\alpha j} \partial_j \mathbf{v}^{\beta i} = \partial_\beta \mathbf{v}^{\alpha i} + \sum_{j=1}^q \mathbf{v}^{\beta j} \partial_j \mathbf{v}^{\alpha i},$$

at least on the support of $\boldsymbol{\mu}$ in $\Omega \times D$. In other words, we see that $\mathbf{v}$ must satisfy a differential constraint for the superposition principle to hold, and there is no reason why this constraint would be satisfied for a generic $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$, even for a harmonic mapping.

An other way to understand the *local* failure of the superposition principle is the following. We will be sloppy and use the evaluation operators $e_\xi : \mathcal{F} \to D$ defined by $e_\xi(f) := f(\xi)$ (these operators are in principle not defined as elements of $\mathcal{F}$ are not continuous). If $\boldsymbol{\mu}$ admits a superposition principle, it would mean that for ξ and η very close, $(e_\xi, e_\eta)\#Q \in \mathcal{P}(D \times D)$ is a transport plan between $\boldsymbol{\mu}(\xi)$ and $\boldsymbol{\mu}(\eta)$ (because of point (i)) which is almost optimal (between of point (ii)). It also works with three measures: if ξ, η and θ are three points of Ω very close to each other (for instance located at the vertices of an equilateral triangle), then $(e_\xi, e_\eta, e_\theta)\#Q \in \mathcal{P}(D \times D \times D)$ is a coupling between $\boldsymbol{\mu}(\xi), \boldsymbol{\mu}(\eta)$ and $\boldsymbol{\mu}(\theta)$ whose 2-marginals are almost optimal transport plans. However, it is known that, if μ_1, μ_2 and $\mu_3 \in \mathcal{P}(D)$, then in general there exists no coupling between the three whose 2-marginals are optimal transport plans.

6. A Ishihara type property

As explained in the introduction, we want to show in this section that $F \circ \boldsymbol{\mu}$ is subharmonic (which means $\Delta(F \circ \boldsymbol{\mu}) \geqslant 0$) as soon as $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ is harmonic and $F : \mathcal{P}(D) \to \mathbf{R}$ is convex along generalized geodesics. As far as the regularity of F is concerned the simplest would be to assume that F is continuous on $\mathcal{P}(D)$. Nevertheless, this assumption is very strong and excludes natural functionals (like the internal energies). In the case where F is only l.s.c., we will need additional assumptions: it is the object of the following definition.

Definition 6.1. *We say that $F : \mathcal{P}(D) \to \mathbf{R} \cup \{+\infty\}$ is regular if it is l.s.c. on $\mathcal{P}(D)$, if*

$$\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D)) \mapsto \int_\Omega F(\boldsymbol{\mu}(\xi)) \mathrm{d}\xi$$

is l.s.c. for the weak convergence on $L^2(\Omega, \mathcal{P}(D))$, and if F is bounded on the bounded sets of $L^\infty(D) \cap \mathcal{P}(D)$.

Lower semi-continuity of F is a reasonable assumption. To impose that F is bounded on bounded sets of $L^\infty(D) \cap \mathcal{P}(D)$ is not a strong constraint as D is compact, we will need it to ensure that, by regularizing probability measures with the heat flow, we get measures for which F is finite.

Lower semi-continuity of $\mathfrak{F} : \boldsymbol{\mu} \mapsto \int_\Omega (F \circ \boldsymbol{\mu})$ is less usual: by a standard argument left to the reader, it implies that F is convex for the affine structure on $\mathcal{P}(D)$. However, we do not know in the general case if the fact that F is convex and l.s.c. on $\mathcal{P}(D)$ is enough to ensure lower semi-continuity of $\mathfrak{F}$. Indeed, to apply abstract functional analysis arguments, we would like to work in the space $\mathcal{M}(\Omega \times D)$ endowed with the total variation norm: it is the dual of the Banach space $(C(\Omega \times D), \|\cdot\|_\infty)$. If F is convex and l.s.c. on $\mathcal{P}(D)$, it can be shown easily that $\mathfrak{F}$ is convex and l.s.c. on $\mathcal{M}(\Omega \times D)$ endowed with the total variation norm. However, it only implies that $\mathfrak{F}$ is l.s.c. for the topology on $\mathcal{M}(\Omega \times D)$ defined by duality w.r.t. the dual of $\mathcal{M}(\Omega \times D)$, the latter being strictly larger than $C(\Omega \times D)$.

However, for the usual functionals on $\mathcal{P}(D)$ we can do an *ad hoc* analysis and we have the following results.

Proposition 6.2. *Let $V \in L^1(D)$ a l.s.c. function. Then the functional*

$$F : \mu \in \mathcal{P}(D) \mapsto \int_D V \mathrm{d}\mu$$

is regular.

Let $f : [0,+\infty) \to \mathbf{R}$ a proper and convex function such that $\lim_{t\to+\infty} f(t)/t = +\infty$. Then the functional defined by

$$F : \mu \in \mathcal{P}(D) \mapsto \begin{cases} \int_D f(\mu(x))\mathrm{d}x & \text{if } \mu \text{ is absolutely continuous w.r.t. } \mathcal{L}_D \\ +\infty & \text{else}, \end{cases}$$

is regular.

Proof. As V is l.s.c. on the compact D, it is bounded from below. As V is in $L^1(\Omega)$, the function F is clearly bounded on bounded sets of $L^\infty(D) \cap \mathcal{P}(D)$. Then, we can use [27, Proposition 7.1], seeing either V as a l.s.c. function on D, or as a l.s.c. on $\Omega \times D$ (constant w.r.t. its first variable) to get that both F and $\int_\Omega (F \circ \cdot)$ are l.s.c.

For the internal energy, to get lower semi-continuity of F we rely on [27, Proposition 7.7]. To get the lower semi-continuity of $\int_\Omega (F \circ \cdot)$, we can see that

$$\int_\Omega F(\boldsymbol{\mu}(\xi))\mathrm{d}\xi = \begin{cases} \int_{\Omega\times D} f(\boldsymbol{\mu}(\xi,x))\mathrm{d}\xi\mathrm{d}x & \text{if } \boldsymbol{\mu} \text{ is absolutely continuous w.r.t. } \mathcal{L}_\Omega \otimes \mathcal{L}_D \\ +\infty & \text{else}, \end{cases}$$

thus [27, Proposition 7.7] still applies. As f is bounded on bounded sets of $[0,+\infty)$, we see that F is bounded on bounded sets of $L^\infty(D) \cap \mathcal{P}(D)$. □

However, the interaction energy is not regular: it lacks convexity w.r.t. the affine structure on $\mathcal{P}(D)$ [27, Chapter 7]. For instance, take $\Omega = D = [0,1]$ and define $F : \mathcal{P}(D) \to \mathbf{R}$ by

$$F(\mu) := \int_{D\times D} |x-y|^2 \mu(\mathrm{d}x)\mu(\mathrm{d}y).$$

This functional is continuous on $\mathcal{P}(D)$ and bounded on bounded subsets of $L^\infty(D) \cap \mathcal{P}(D)$. However, if we define $\boldsymbol{\mu}_n(\xi) := \delta_{x_n(\xi)}$ with $x_n(\xi) = 1/2 + 1/2\cos(n\xi)$, one can see that $F(\boldsymbol{\mu}_n(\xi)) = 0$ for all $\xi \in \Omega$ and $n \in \mathbf{N}$, but $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ converges weakly on $\mathcal{P}(\Omega \times D)$ to $\boldsymbol{\mu} := \mathcal{L}_\Omega \otimes \mathcal{L}_D$, for which the value $\int_\Omega (F \circ \boldsymbol{\mu})$ is strictly positive. On the other hand, as soon as the interaction potential is continuous, the interaction energy is continuous on $\mathcal{P}(D)$.

Finally, let us recall that a function $f : \Omega \to \mathbf{R}$ is said subharmonic on $\mathring{\Omega}$ in the sense of distributions if $\Delta f \geqslant 0$ as a distribution in $\mathring{\Omega}$.

Theorem 6.3. *Let $F : \mathcal{P}(D) \to \mathbf{R} \cup \{+\infty\}$ a functional which is convex along generalized geodesics. Assume either that F is continuous (and everywhere finite) on $\mathcal{P}(D)$ or that F is regular. Let $\boldsymbol{\mu}_l : \partial\Omega \to \mathcal{P}(D)$ a Lipschitz mapping such that $\sup(F \circ \boldsymbol{\mu}_l) < +\infty$.*

Then there exists at least one solution $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ of the Dirichlet problem with boundary conditions $\boldsymbol{\mu}_l$ such that $(F \circ \boldsymbol{\mu}) : \Omega \to \mathbf{R}$ is subharmonic in $\mathring{\Omega}$ in the sense of distributions and

$$\operatorname*{ess\,sup}_\Omega (F \circ \boldsymbol{\mu}) \leqslant \sup_{\partial\Omega} (F \circ \boldsymbol{\mu}_l). \tag{6.1}$$

Moreover, if F is regular then $\boldsymbol{\mu}$ can be chosen in such a way that

$$\int_\Omega F(\boldsymbol{\mu}(\xi))\mathrm{d}\xi \leqslant \int_\Omega F(\boldsymbol{\nu}(\xi))\mathrm{d}\xi. \tag{6.2}$$

if $\boldsymbol{\nu}$ is an other solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$.

Let us make some comments. The first one is that (6.1) is nothing else than the maximum principle. It is not implied by the subharmonicity of $(F \circ \boldsymbol{\mu})$ as the latter holds only in $\mathring{\Omega}$ and we do not know

if $(F \circ \boldsymbol{\mu})$ is continuous. The second one is that (6.2) characterizes $\boldsymbol{\mu}$ if F is strictly convex. More generally, the subharmonicity of $F \circ \boldsymbol{\mu}$ would hold for $\boldsymbol{\mu}$ solution of the Dirichlet problem minimizing

$$\int_\Omega a(\xi) F(\boldsymbol{\mu}(\xi)) \mathrm{d}\xi,$$

where $a \in C(\Omega)$ is a continuous and strictly positive function (it comes from a slight modification of the proof which is left to the reader). The last comment is that this result is somehow disappointing because we cannot guarantee the subharmonicity to hold for all solutions. The main issue is that we reason by approximation, thus the solution $\boldsymbol{\mu}$ is constructed as the limit of some approximate mappings, the existence of the limit is coming from compactness. But as we have no uniqueness result for the Dirichlet problem, we can only identify the limit through (6.2) (which is a byproduct of the approximation process) but we cannot say much more.

The rest of this section is devoted to the proof of Theorem 6.3. In Subsection 6.1 we prove some preliminary results. The most difficult and interesting case is the one where F is not assumed to be continuous but only regular: it is the object of Subsections 6.2 and 6.3. To conclude, in Subsection 6.4, we briefly comment about the simplifications of the proof in the case of a continuous F.

6.1. **Preliminary results.** We prove first some technical results which would have overburden the previous sections. The first one deal with Rellich compactness theorem, as we will want some strong convergence of our solutions of the approximate problems.

Proposition 6.4. *Let $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ a sequence in $H^1(\Omega, \mathcal{P}(D))$ such that $\sup_n \mathrm{Dir}(\boldsymbol{\mu}_n) < +\infty$. Then, up to extraction, the sequence $(\boldsymbol{\mu}_n)_{n\in\mathbf{N}}$ converges strongly in $L^2(\Omega, \mathcal{P}(D))$ to some $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$.*

Proof. This is nothing else than the Rellich compactness theorem, but for mappings valued in metric spaces. Remark that $\mathcal{P}(D)$ has a finite diameter, thus in this result we only need a control on the Dirichlet energy of $\boldsymbol{\mu}_n$. We can find this result for instance in [19, Theorem 1.13] or in [3, Theorem 5.4.3]. In any way, this result is also a consequence of the next proposition. □

In fact, we will need a stronger result, as we want so show compactness if we only have a control of the approximate Dirichlet energies.

Proposition 6.5. *Let $(\boldsymbol{\mu}_\varepsilon)_{\varepsilon>0}$ a family in $L^2(\Omega, \mathcal{P}(D))$ such that $\liminf_\varepsilon \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_\varepsilon) < +\infty$. Then there exists a sequence $(\varepsilon_n)_{n\in\mathbf{N}}$ which goes to 0 such that $(\boldsymbol{\mu}_{\varepsilon_n})_{n\in\mathbf{N}}$ converges strongly in $L^2(\Omega, \mathcal{P}(D))$ to some $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$.*

There is a well known criterion for compactness in $L^2(\Omega)$: the Riesz-Fréchet-Kolmogorov theorem. It requires a uniform control of the L^2-norm of the difference between a function and its translated. Here, we have only a control of the distance between a function and its translated in average (thanks to Dir_ε), and our mappings take values in $\mathcal{P}(D)$ rather than $\mathbf{R}$. Nevertheless, the strategy of the proof of the Riesz-Fréchet-Kolmogorov theorem is rather straightforward to adapt. Recall that K^Ω denotes the heat kernel on Ω.

Proof. There exists a sequence $(\varepsilon_m)_{m\in\mathbf{N}}$, converging to 0, such that $\sup_m \mathrm{Dir}_{\varepsilon_m}(\boldsymbol{\mu}_{\varepsilon_m}) < +\infty$. For $t > 0$, define $\Phi^\Omega_t \boldsymbol{\mu}_{\varepsilon_m} \in L^2(\Omega, \mathcal{P}(D))$ by, for any $\xi \in \Omega$,

$$[\Phi^\Omega_t \boldsymbol{\mu}_{\varepsilon_m}](\xi) := \int_\Omega K^\Omega_t(\xi, \theta) \boldsymbol{\mu}_{\varepsilon_m}(\theta) \mathrm{d}\theta.$$

(We follow the heat flow but only in Ω). Let us estimate $d_{L^2}(\Phi^\Omega_t \boldsymbol{\mu}_{\varepsilon_m}, \boldsymbol{\mu}_{\varepsilon_m})$: we first use Jensen's inequality, then separate between values close or away from the diagonal of $\Omega \times \Omega$ at the level ε

(whose value will be fixed later), and we use the fact that $|K_t^\Omega(\xi,\theta)| \leqslant C_1 t^{-p/2}\exp(-|\xi-\theta|^2/(C_2 t))$ (see Proposition (ii) in 2.4).

$$\begin{aligned}
d^2_{L^2}(\Phi_t^\Omega \boldsymbol{\mu}_{\varepsilon_m}, \boldsymbol{\mu}_{\varepsilon_m}) &= \int_\Omega W_2^2\left(\int_\Omega K_t^\Omega(\xi,\theta)\boldsymbol{\mu}_{\varepsilon_m}(\theta)\mathrm{d}\theta, \boldsymbol{\mu}_{\varepsilon_m}(\xi)\right)\mathrm{d}\xi \\
&\leqslant \int_{\Omega\times\Omega} K_t^\Omega(\xi,\theta) W_2^2(\boldsymbol{\mu}_{\varepsilon_m}(\xi), \boldsymbol{\mu}_{\varepsilon_m}(\theta))\mathrm{d}\xi\mathrm{d}\theta \\
&= \int_{\Omega\times\Omega} K_t^\Omega(\xi,\theta) W_2^2(\boldsymbol{\mu}_{\varepsilon_m}(\xi), \boldsymbol{\mu}_{\varepsilon_m}(\theta))\mathbb{1}_{|\xi-\eta|\leqslant\varepsilon}\mathrm{d}\xi\mathrm{d}\theta + \int_{\Omega\times\Omega} K_t^\Omega(\xi,\theta) W_2^2(\boldsymbol{\mu}_{\varepsilon_m}(\xi), \boldsymbol{\mu}_{\varepsilon_m}(\theta))\mathbb{1}_{|\xi-\eta|\geqslant\varepsilon}\mathrm{d}\xi\mathrm{d}\theta \\
&\leqslant \frac{C_1}{t^{p/2}}\int_{\Omega\times\Omega} W_2^2(\boldsymbol{\mu}_{\varepsilon_m}(\xi), \boldsymbol{\mu}_{\varepsilon_m}(\theta))\mathbb{1}_{|\xi-\theta|\leqslant\varepsilon}\mathrm{d}\xi\mathrm{d}\theta + C_3\int_{\Omega\times\Omega} K_t^\Omega(\xi,\theta)\mathbb{1}_{|\xi-\theta|>\varepsilon}\mathrm{d}\xi\mathrm{d}\theta \\
&\leqslant \frac{2C_1\varepsilon^{p+2}}{C_p t^{p/2}}\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon_m}) + \frac{C_4}{t^{p/2}}\exp\left(-\frac{\varepsilon^2}{C_2 t}\right).
\end{aligned}$$

Choose the free parameters as follows: we fix $t>0$ and we use the scaling $\varepsilon = t^\beta$ with $p/(2(p+2)) < \beta < 1/2$. Thanks to this scaling, for t small enough one has that $\varepsilon^{p+2}t^{-p/2}$ is small and and $\varepsilon^2 t^{-1}$ is large. Moreover, by monotonicity of the Dirichlet energy, for m large enough one has $\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon_m}) \leqslant \mathrm{Dir}_{\varepsilon_m}(\boldsymbol{\mu}_{\varepsilon_m}) \leqslant C < +\infty$ (more precisely for each m large enough we choose ε of the form $2^N\varepsilon_m$ with $N\in\mathbf{N}$ and $\varepsilon \simeq t^\beta$). In consequence, for any $\delta>0$, there exists $t>0$ and $m_0\in\mathbf{N}$, such that for any $m\geqslant m_0$,

$$d_{L^2}(\Phi_t^\Omega\boldsymbol{\mu}_{\varepsilon_m}, \boldsymbol{\mu}_{\varepsilon_m}) \leqslant \delta.$$

On the other hand, for a fixed $t>0$, we want to show compactness of the family $\left(\Phi_t^\Omega\boldsymbol{\mu}_\varepsilon\right)$. Because of boundary effects we restrict ourselves to $\tilde\Omega \subset \mathring\Omega$ compactly supported in $\mathring\Omega$. We will show that the family $\left(\Phi_t^\Omega\boldsymbol{\mu}_\varepsilon\big|_{\tilde\Omega}\right)_{\varepsilon>0}$ is uniformly equi-Hölder, where $\Phi_t^\Omega\boldsymbol{\mu}_\varepsilon\big|_{\tilde\Omega}$ denotes the restriction of $\Phi_t^\Omega\boldsymbol{\mu}_\varepsilon$ to $\tilde\Omega$. Rather than to estimate $W_2^2(\Phi_t^\Omega\boldsymbol{\mu}_\varepsilon(\xi), \Phi_t^\Omega\boldsymbol{\mu}_\varepsilon(\eta))$, we prefer to work on the 1-Wasserstein distance $W_1(\Phi_t^\Omega\boldsymbol{\mu}_\varepsilon(\xi), \Phi_t^\Omega\boldsymbol{\mu}_\varepsilon(\eta))$ whose definition is recalled in Section 2. Take $\varphi\in C(D)$ a 1-Lipschitz function, up to translation by a constant we can assume that $\|\varphi\|_\infty \leqslant C$ with C independent of φ. Then for any $\xi,\eta\in\tilde\Omega$,

$$\begin{aligned}
\int_D \varphi(x)[\Phi_t^\Omega\boldsymbol{\mu}_\varepsilon(\xi)](\mathrm{d}x) - \int_D \varphi(x)[\Phi_t^\Omega\boldsymbol{\mu}_\varepsilon(\eta)](\mathrm{d}x) &= \int_{\Omega\times D}\varphi(x)[K_t^\Omega(\xi,\theta)-K_t^\Omega(\eta,\theta)]\boldsymbol{\mu}_\varepsilon(\mathrm{d}\theta,\mathrm{d}x) \\
&\leqslant C\|\varphi\|_\infty|\xi-\eta| \leqslant C|\xi-\eta|,
\end{aligned}$$

where we have used the fact that $K_t^\Omega(\cdot,\theta)$ is Lipschitz on $\tilde\Omega$ uniformly in $\theta\in\Omega$: by parabolic regularity the Lipschitz constant of a solution of the heat equation depends only on t, the distance to the boundary and the L^1 norm of the initial data. As the bound is independent on φ, we deduce that $W_1(\Phi_t^\Omega\boldsymbol{\mu}_\varepsilon(\xi), \Phi_t^\Omega\boldsymbol{\mu}_\varepsilon(\eta)) \leqslant C|\xi-\eta|$ for all ξ and η in $\tilde\Omega$. Using $W_2\leqslant C\sqrt{W_1}$ [27, Equation (5.1)], we see that, for a fixed t, the family $(\Phi_t^\Omega\boldsymbol{\mu}_{\varepsilon_m}\big|_{\tilde\Omega})_{m\in\mathbf{N}}$ is uniformly equi-continuous (more precisely $1/2$-Hölder).

Now we put the pieces together. For each $n\in\mathbf{N}$, take t_n such that $d_{L^2}(\Phi_{t_n}^\Omega\boldsymbol{\mu}_{\varepsilon_m}, \boldsymbol{\mu}_{\varepsilon_m}) \leqslant 1/n$ holds for m large enough. Take also $\Omega_n\subset\mathring\Omega$ compactly supported in $\mathring\Omega$ such that $\mathcal{L}_\Omega(\Omega\backslash\Omega_n)\leqslant 1/n$. Then, using Ascoli-Arzelà theorem, up to a subsequence, we know that $(\Phi_{t_n}^\Omega\boldsymbol{\mu}_{\varepsilon_m}\big|_{\Omega_n})_{m\in\mathbf{N}}$ converges strongly in $L^2(\Omega_n,\mathcal{P}(D))$, in particular it is a Cauchy sequence. Up to a diagonal extraction in $(\varepsilon_m)_{m\in\mathbf{N}}$ (we do not relabel the sequence), we can assume that $(\Phi_{t_n}^\Omega\boldsymbol{\mu}_{\varepsilon_m}\big|_{\Omega_n})_{m\in\mathbf{N}}$ is a Cauchy sequence for all $n\in\mathbf{N}$. Notice, as $\mathcal{P}(D)$ has a finite diameter, that $|d_{L^2}(\boldsymbol{\mu},\boldsymbol{\nu}) - d_{L^2}(\boldsymbol{\mu}|_{\Omega_n}, \boldsymbol{\nu}|_{\Omega_n})| \leqslant C/n$ for all

$\boldsymbol{\mu}, \boldsymbol{\nu} \in L^2(\Omega, \mathcal{P}(D))$. Hence, for any $n \in \mathbf{N}$, one has for m and m' large enough,

$$\begin{aligned} d_{L^2}(\boldsymbol{\mu}_{\varepsilon_m}, \boldsymbol{\mu}_{\varepsilon_{m'}}) &\leqslant d_{L^2}(\boldsymbol{\mu}_{\varepsilon_m}, \Phi^\Omega_{t_n}\boldsymbol{\mu}_{\varepsilon_m}) + d_{L^2}(\Phi^\Omega_{t_n}\boldsymbol{\mu}_{\varepsilon_m}\big|_{\Omega_n}, \Phi^\Omega_{t_n}\boldsymbol{\mu}_{\varepsilon_{m'}}\big|_{\Omega_n}) + d_{L^2}(\Phi^\Omega_{t_n}\boldsymbol{\mu}_{\varepsilon_{m'}}, \boldsymbol{\mu}_{\varepsilon_{m'}}) + \frac{2C}{n} \\ &\leqslant \frac{2+2C}{n} + d_{L^2}(\Phi^\Omega_{t_n}\boldsymbol{\mu}_{\varepsilon_m}\big|_{\Omega_n}, \Phi^\Omega_{t_n}\boldsymbol{\mu}_{\varepsilon_{m'}}\big|_{\Omega_n}), \end{aligned}$$

and $d_{L^2}(\Phi^\Omega_{t_n}\boldsymbol{\mu}_{\varepsilon_m}\big|_{\Omega_n}, \Phi^\Omega_{t_n}\boldsymbol{\mu}_{\varepsilon_{m'}}\big|_{\Omega_n})$ can be made arbitrary small for m and m' large enough. In other words, $(\boldsymbol{\mu}_{\varepsilon_m})_{m\in\mathbf{N}}$ is a Cauchy sequence in $L^2(\Omega, \mathcal{P}(D))$, thus it converges strongly. □

We will also need a result about the boundary conditions. Indeed, as the minimizers of Dir_ε will only live in $L^2(\Omega, \mathcal{P}(D))$, we cannot define and impose boundary values. To bypass this difficulty, we extend slightly our domain into a larger domain $\Omega_e \supset \Omega$ and impose the values of the mappings everywhere on $\Omega_e \backslash \mathring{\Omega}$.

Proposition 6.6. *Let $\boldsymbol{\mu}_l : \partial\Omega \to \mathcal{P}(D)$ a Lipschitz mapping. There exists a compact Ω_e such that $\Omega \subset \mathring{\Omega}_e$, and a Lipschitz mapping $\boldsymbol{\mu}_e \in L^2(\Omega_e \backslash \mathring{\Omega}, \mathcal{P}(D))$ such that $\boldsymbol{\mu}_e = \boldsymbol{\mu}_l$ on $\partial\Omega$ and*

$$\{\boldsymbol{\mu}_e(\xi) \ : \ \xi \in \Omega_e \backslash \mathring{\Omega}\} = \{\boldsymbol{\mu}_l(\xi) \ : \ \xi \in \partial\Omega\}. \tag{6.3}$$

Moreover, a mapping $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ satisfies $\boldsymbol{\mu}|_{\partial\Omega} = \boldsymbol{\mu}_l$ if and only if the mapping $\tilde{\boldsymbol{\mu}}$ defined on Ω_e by

$$\tilde{\boldsymbol{\mu}}(\xi) = \begin{cases} \boldsymbol{\mu}(\xi) & \text{if } \xi \in \mathring{\Omega} \\ \boldsymbol{\mu}_e(\xi) & \text{if } \xi \in \Omega_e \backslash \mathring{\Omega}, \end{cases}$$

belongs to $H^1(\Omega_e, \mathcal{P}(D))$.

Proof. As Ω has a Lipschitz boundary, one can say [19, Section 1.12] that there exists a compact Ω_e such that $\Omega \subset \mathring{\Omega}_e$, and $\Psi : [0,1] \times \partial\Omega \to \Omega_e \backslash \mathring{\Omega}$ a bilipschitz mapping such that $\Psi(0, \cdot)$ is the identity on $\partial\Omega$. Roughly speaking (for instance if $\partial\Omega$ is C^1), $\Psi(t, \xi) = \xi + t\mathbf{n}_\Omega(\xi)$ where $\mathbf{n}_\Omega$ is the outward normal to $\partial\Omega$. Then, one can define

$$\boldsymbol{\mu}_e(\Psi(t, \xi)) := \boldsymbol{\mu}_l(\xi)$$

for every $t \in [0,1]$ and $\xi \in \partial\Omega$: we extend $\boldsymbol{\mu}_l$ by keeping it constant along the normal to $\partial\Omega$. Because Ψ is bilipschitz and $\boldsymbol{\mu}_l$ is Lipschitz, it is clear that $\boldsymbol{\mu}_e$ is a Lipschitz mapping. Moreover, by construction, (6.3) obviously holds.

Let us prove the second point. Take $E \in \mathcal{M}(\Omega \times D, \mathbf{R}^{pq})$ and $E_e \in \mathcal{M}((\Omega_e \backslash \mathring{\Omega}) \times D, \mathbf{R}^{pq})$ the momenta tangent to respectively $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_e$. The tangent momentum of $\tilde{\mu}$, if it were to exist, must coincide with E on $\Omega \times D$ and with E_e on $(\Omega_e \backslash \mathring{\Omega}) \times D$ because of Proposition 3.12. Hence, if must be $\tilde{E} \in \mathcal{M}(\Omega_e \times D, \mathbf{R}^{pq})$ defined by

$$\int_{\Omega_e \times D} b \cdot \mathrm{d}\tilde{E} = \int_{\Omega \times D} b \cdot \mathrm{d}E + \int_{(\Omega_e \backslash \mathring{\Omega}) \times D} b \cdot \mathrm{d}E_e.$$

As we already have $\mathrm{Dir}(\tilde{\boldsymbol{\mu}}, \tilde{E}) < +\infty$, we see that $\tilde{\mu} \in H^1(\Omega_e, \mathcal{P}(D))$ if and only if $(\tilde{\boldsymbol{\mu}}, \tilde{E})$ satisfies the continuity equation. If $\varphi \in C^1_c(\Omega_e, \mathbf{R}^p)$,

$$\begin{aligned} &\int_{\Omega_e \times D} \nabla_\Omega \cdot \varphi \mathrm{d}\tilde{\boldsymbol{\mu}} + \int_{\Omega_e \times D} \nabla_D \varphi \cdot \mathrm{d}\tilde{E} \\ &\quad = \int_{\Omega \times D} \nabla_\Omega \cdot \varphi \mathrm{d}\boldsymbol{\mu} + \int_{\Omega \times D} \nabla_D \varphi \cdot \mathrm{d}E + \int_{(\Omega_e \backslash \mathring{\Omega}) \times D} \nabla_\Omega \cdot \varphi \mathrm{d}\boldsymbol{\mu}_e + \int_{(\Omega_e \backslash \mathring{\Omega}) \times D} \nabla_D \varphi \cdot \mathrm{d}E_e \\ &\quad = \mathrm{BT}_{\boldsymbol{\mu}}(\varphi) + \mathrm{BT}_{\boldsymbol{\mu}_e}(\varphi). \end{aligned}$$

By Whitney's theorem, the restriction of functions in $C^1_c(\mathring{\Omega}_e, \mathbf{R}^p)$ to Ω coincide with $C^1(\Omega, \mathbf{R}^p)$, thus we see that $\tilde{\boldsymbol{\mu}} \in H^1(\Omega_e, \mathcal{P}(D))$ if and only if $\mathrm{BT}_{\boldsymbol{\mu}} = -\mathrm{BT}_{\boldsymbol{\mu}_e}$. Considering the fact that the outward

normal to $\Omega_e\backslash\mathring{\Omega}$ is $-\mathbf{n}_\Omega$, and that $\boldsymbol{\mu}_e$ is continuous with values on $\partial\Omega$ given by $\boldsymbol{\mu}_l$, the proposition is proved. □

6.2. **The approximate problems and their optimality conditions.** In all this subsection, we assume that F is regular. As explained before, we use Dir_ε to approximate Dir, as the optimality conditions of Dir_ε imply that for each $\xi\in\Omega$, $\boldsymbol{\mu}(\xi)$ is a barycenter of all $\boldsymbol{\mu}(\eta)$ for η in the ball of center ξ and radius ε.

Let us introduce some notations that we will keep during the rest of the proof. We denote by $\Omega_e\supset\Omega$ and $\boldsymbol{\mu}_e\in H^1(\Omega_e\backslash\mathring{\Omega},\mathcal{P}(D))$ the objects given by Proposition 6.6. Take $\varepsilon_0>0$ such that $B(\xi,\varepsilon_0)\subset\Omega_e$ for all $\xi\in\partial\Omega$. We denote by

$$L^2_e(\Omega_e,\mathcal{P}(D)):=\{\boldsymbol{\mu}\in L^2(\Omega_e,\mathcal{P}(D))\ :\ \boldsymbol{\mu}|_{\Omega_e\backslash\mathring{\Omega}}=\boldsymbol{\mu}_e\}$$

the set of L^2 mappings which coincide with $\boldsymbol{\mu}_e$ on $\Omega_e\backslash\mathring{\Omega}$. This set $L^2_e(\Omega_e,\mathcal{P}(D))$ is clearly closed for the weak convergence on $L^2(\Omega_e,\mathcal{P}(D))$, in particular it is compact for the weak convergence. We also define $H^1_e(\Omega_e,\mathcal{P}(D)):=H^1(\Omega_e,\mathcal{P}(D))\cap L^2_e(\Omega_e,\mathcal{P}(D))$. In the rest of the proof, we extend the definitions of Dir_ε and Dir on $L^2_e(\Omega_e,\mathcal{P}(D))$. More precisely, if $\boldsymbol{\mu}\in L^2_e(\Omega_e,\mathcal{P}(D))$,

$$\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}):=C_p\int_{\Omega_e\times\Omega_e}\frac{W_2^2(\boldsymbol{\mu}(\xi),\boldsymbol{\mu}(\eta))}{2\varepsilon^{p+2}}\mathbb{1}_{|\xi-\eta|\leqslant\varepsilon}\mathrm{d}\xi\mathrm{d}\eta,$$

and

$$\mathrm{Dir}(\boldsymbol{\mu})$$
$$:=\inf\left\{\mathrm{Dir}(\boldsymbol{\mu},E)\ :\ E\in\mathcal{M}(\Omega_e\times D,\mathbf{R}^{pq})\text{ and }(\boldsymbol{\mu},E)\text{ satisfies the continuity equation on }\Omega_e\times D\right\}.$$

(we integrate over Ω_e and not only on Ω). We also use the notation

$$M:=\sup_{\partial\Omega}(F\circ\boldsymbol{\mu}_l),$$

by assumption M is finite. Remark that by construction, if $\boldsymbol{\mu}\in L^2_e(\Omega_e,\mathcal{P}(D))$, then for all $\xi\in\Omega_e\backslash\mathring{\Omega}$ one has $F(\boldsymbol{\mu}(\xi))\leqslant M$.

As F is l.s.c. on the compact set $\mathcal{P}(D)$, it is bounded from below. Hence, we can translate it by a constant and assume that $F\geqslant 0$ on $\mathcal{P}(D)$.

Let $\varepsilon>0$ and $\lambda>0$ be fixed. The approximate problem is defined as

$$\min\left\{\mathrm{Dir}_\varepsilon(\boldsymbol{\mu})+\lambda\int_{\Omega_e}F(\boldsymbol{\mu}(\xi))\mathrm{d}\xi\ :\ \boldsymbol{\mu}\in L^2_e(\Omega_e,\mathcal{P}(D))\right\}.$$

To add the term $\lambda\int_{\Omega_e}F\circ\boldsymbol{\mu}$ has two purposes: on the one hand, it ensures that $F\circ\boldsymbol{\mu}$ will be regular enough (namely in $L^1(\Omega_e)$) to extract information from the optimality conditions; on the other hand by taking the limit $\varepsilon\to 0$ and then $\lambda\to 0$, we will be able to say that $F\circ\boldsymbol{\mu}_{\varepsilon,\lambda}$ (where $\boldsymbol{\mu}_{\varepsilon,\lambda}$ is a minimizer of the approximate problem) converges pointewisely, and it is necessary to pass to the limit the (approximate) subharmonicity that we will get from the optimality conditions of the approximate problem.

The following result is easy with all the tools developed above.

Proposition 6.7. *For any $\varepsilon>0$ and $\lambda>0$, there exists a solution to the approximate problem.*

Proof. Let $\nu\in\mathcal{P}(D)$ any measure such that $F(\nu)<+\infty$ (it exists as F is regular). If we define $\boldsymbol{\mu}\in L^2_e(\Omega_e,\mathcal{P}(D))$ by $\boldsymbol{\mu}|_{\mathring{\Omega}}:=\nu$ and $\boldsymbol{\mu}|_{\Omega_e\backslash\mathring{\Omega}}:=\boldsymbol{\mu}_e$, one can see that $\int_{\Omega_e}F(\boldsymbol{\mu}(\xi))\mathrm{d}\xi<+\infty$, moreover as $\mathcal{P}(D)$ has a finite diameter $\mathrm{Dir}_\varepsilon(\boldsymbol{\mu})<+\infty$. Hence, the minimization problem is non empty. In consequence, we are minimizing over the set $L^2_e(\Omega_e,\mathcal{P}(D))$, which is compact for the weak convergence, a functional which is l.s.c. (see Proposition 3.23 and the regularity assumption on F): we can use the direct method of calculus of variations. □

Starting from now, for any $\varepsilon > 0$ and $\lambda > 0$, we denote by $\boldsymbol{\mu}_{\varepsilon,\lambda}$ a solution of the approximate problem.

Proposition 6.8. *Let $0 < \varepsilon \leqslant \varepsilon_0$ and $\lambda > 0$ be fixed. Then for a.e. $\xi \in \Omega$, $\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)$ is a minimizer over $\mathcal{P}(D)$ of*

$$\nu \mapsto \frac{C_p}{\varepsilon^{p+2}} \int_{B(\xi,\varepsilon)} W_2^2(\nu, \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))\mathrm{d}\eta + \lambda F(\nu).$$

Proof. We reason by contradiction. If the property does not hold, there exists $c > 0$ and a set $X \subset \mathring{\Omega}$ of strictly positive measure such that for all $\xi \in X$,

(6.4)
$$\frac{C_p}{\varepsilon^{p+2}} \int_{B(\xi,\varepsilon)} W_2^2(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi), \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))\mathrm{d}\eta + \lambda F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)) \geqslant c + \min_{\nu\in\mathcal{P}(D)} \left(\frac{C_p}{\varepsilon^{p+2}} \int_{B(\xi,\varepsilon)} W_2^2(\nu, \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))\mathrm{d}\eta + \lambda F(\nu) \right).$$

Now, consider $\delta > 0$ small and $Y \subset X$ such that $\mathcal{L}_\Omega(Y) = \delta$. On every point of $\xi \in Y$, we want to select a minimizer ν (which depends on ξ) of the r.h.s. of (6.4), and we want to dot it in a measurable way. Notice that

$$\nu \mapsto \frac{C_p}{\varepsilon^{p+2}} \int_{B(\xi,\varepsilon)} W_2^2(\nu, \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))\mathrm{d}\eta + \lambda F(\nu)$$

is the sum of a functional continuous w.r.t. ν and measurable w.r.t. ξ, and the functional λF which is l.s.c. w.r.t. ν but which does not depend on ξ. The fact that F is only l.s.c. prevents us from using directly Proposition A.1, though by some *ad hoc* measurable selection result which is stated and proved in the appendix (Proposition A.3), one can still choose $\nu(\xi)$ a minimizer in such a way that it is measurable in ξ. In other words, we construct $\tilde{\boldsymbol{\mu}} \in L^2_e(\Omega_e, \mathcal{P}(D))$ such that $\tilde{\boldsymbol{\mu}} = \boldsymbol{\mu}_{\varepsilon,\lambda}$ on $\Omega_e \backslash Y$ and

$$\frac{C_p}{\varepsilon^{p+2}} \int_{B(\xi,\varepsilon)} W_2^2(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi), \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))\mathrm{d}\eta + \lambda F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)) \geqslant c + \left(\frac{C_p}{\varepsilon^{p+2}} \int_{B(\xi,\varepsilon)} W_2^2(\tilde{\boldsymbol{\mu}}(\xi), \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))\mathrm{d}\eta + \lambda F(\tilde{\boldsymbol{\mu}}(\xi)) \right)$$

for all $\xi \in Y$. Now we evaluate:

$$\begin{aligned}
&\left(\mathrm{Dir}_\varepsilon(\tilde{\boldsymbol{\mu}}) + \lambda \int_{\Omega_e} F(\tilde{\boldsymbol{\mu}}(\xi))\mathrm{d}\xi \right) - \left(\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon,\lambda}) + \lambda \int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi \right) \\
&\quad = \frac{C_p}{2\varepsilon^{p+2}} \int_{\Omega_e\times\Omega_e} \left[W_2^2(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi), \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - W_2^2(\tilde{\boldsymbol{\mu}}(\xi), \tilde{\boldsymbol{\mu}}(\eta)) \right] \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \mathrm{d}\xi\mathrm{d}\eta + \lambda \int_Y [F(\tilde{\boldsymbol{\mu}}(\xi)) - F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))]\mathrm{d}\xi
\end{aligned}$$

The integral over $\Omega_e \times \Omega_e$ can be split over four parts: the one over $(\Omega_e\backslash Y) \times (\Omega_e\backslash Y)$, which vanishes because $\boldsymbol{\mu}_{\varepsilon,\lambda} = \tilde{\boldsymbol{\mu}}$ on this set; the one over $Y \times Y$, which can be bounded by $C\delta^2$, where C depends on the diameter of $\mathcal{P}(D)$ and on ε; and the ones over $(\Omega_e\backslash Y) \times Y$ and $Y \times (\Omega_e\backslash Y)$ which are equal by symmetry. Moreover, one has

$$\begin{aligned}
&\frac{C_p}{2\varepsilon^{p+2}} \int_{Y\times(\Omega_e\backslash Y)} \left[W_2^2(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi), \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - W_2^2(\tilde{\boldsymbol{\mu}}(\xi), \tilde{\boldsymbol{\mu}}(\eta)) \right] \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \mathrm{d}\xi\mathrm{d}\eta \\
&\qquad = \frac{C_p}{2\varepsilon^{p+2}} \int_{Y\times(\Omega_e\backslash Y)} \left[W_2^2(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi), \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - W_2^2(\tilde{\boldsymbol{\mu}}(\xi), \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) \right] \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \mathrm{d}\xi\mathrm{d}\eta \\
&\qquad \leqslant C\delta^2 + \frac{C_p}{2\varepsilon^{p+2}} \int_{Y\times\Omega_e} \left[W_2^2(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi), \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - W_2^2(\tilde{\boldsymbol{\mu}}(\xi), \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) \right] \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \mathrm{d}\xi\mathrm{d}\eta \\
&\qquad = C\delta^2 + \frac{C_p}{2\varepsilon^{p+2}} \int_Y \left(\int_{B(\xi,\varepsilon)} \left[W_2^2(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi), \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - W_2^2(\tilde{\boldsymbol{\mu}}(\xi), \boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) \right] \mathrm{d}\eta \right) \mathrm{d}\xi,
\end{aligned}$$

where the inequality comes from the fact that we have add the piece $Y \times Y$ which is of size δ^2 and one which we integrate a function which is bounded. Notice that we have used that $B(\xi,\varepsilon) \subset \Omega_e$ for $\xi \in \Omega$ as $\varepsilon < \varepsilon_0$. The part on $(\Omega_e \backslash Y) \times Y$ gives exactly the same amount, thus

$$\begin{aligned}
&\left(\mathrm{Dir}_\varepsilon(\tilde{\boldsymbol{\mu}}) + \lambda \int_{\Omega_e} F(\tilde{\boldsymbol{\mu}}(\xi))\mathrm{d}\xi\right) - \left(\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon,\lambda}) + \lambda \int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi\right)\\
&\leqslant C\delta^2 + \int_Y \left(\frac{C_p}{\varepsilon^{p+2}}\left[\int_{B(\xi,\varepsilon)} [W_2^2(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi),\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - W_2^2(\tilde{\boldsymbol{\mu}}(\xi),\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))]\mathrm{d}\eta\right] + \lambda\left[F(\tilde{\boldsymbol{\mu}}(\xi)) - F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\right]\right)\mathrm{d}\xi\\
&\leqslant C\delta^2 - c\delta,
\end{aligned}$$

where the last inequality comes precisely form the way we chose $\tilde{\boldsymbol{\mu}}$ on Y and of $\mathcal{L}_\Omega(Y) = \delta$. Hence, taking δ small enough, the r.h.s. is strictly negative, which is a contradiction with the optimality of $\boldsymbol{\mu}_{\varepsilon,\lambda}$. □

Remark that if $\lambda = 0$, our proof still works, and it precisely shows that $\boldsymbol{\mu}_{\varepsilon,0}(\xi)$ is a barycenter of the $\boldsymbol{\mu}_{\varepsilon,0}(\eta)$ for η running over the ball of center ξ and radius ε, a fact which was already stated by Jost [16]. The crucial result which allows us to get subharmonicity is the following, namely Jensen's inequality for functionals convex along generalized geodesics. Notice that $F \circ \boldsymbol{\mu}_{\varepsilon,\lambda}$ is integrable on Ω_e.

Proposition 6.9. *Let $0 < \varepsilon \leqslant \varepsilon_0$ and $\lambda > 0$ be fixed. Then, for a.e. $\xi \in \Omega$,*

$$\int_{B(\xi,\varepsilon)} [F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))]\mathrm{d}\eta \geqslant 0.$$

Proof. Let us take a point $\xi \in \Omega$ for which the conclusion of Proposition 6.8 holds and such that $F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)) < +\infty$: it is the case for a.e. points of Ω. As a competitor, we use $S_t^F[\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)]$ for small $t > 0$, which means that we let $\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)$ follow the gradient flow of F, see Theorem 2.3. By Proposition 6.8,

$$\begin{aligned}
&\frac{C_p}{\varepsilon^{p+2}} \int_{B(\xi,\varepsilon)} W_2^2(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi),\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))\mathrm{d}\eta + \lambda F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\\
&\qquad\leqslant \frac{C_p}{\varepsilon^{p+2}} \int_{B(\xi,\varepsilon)} W_2^2(S_t^F[\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)],\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))\mathrm{d}\eta + \lambda F(S_t^F[\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)]).
\end{aligned}$$

By the very definition of gradient flows, $F(S_t^F[\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)]) \leqslant F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))$. Thus, rearranging the terms and dividing by $2t > 0$,

$$\int_{B(\xi,\varepsilon)} \frac{W_2^2(S_t^F[\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)],\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - W_2^2(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi),\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))}{2t}\mathrm{d}\eta \geqslant 0.$$

For a.e. $\eta \in B(\xi,\varepsilon)$, one has that $F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) < +\infty$. Hence, using Theorem 2.3, we see that for a.e. $\eta \in B(\xi,\varepsilon)$, the quantity

$$\frac{W_2^2(S_t^F[\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)],\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - W_2^2(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi),\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))}{2t}$$

has a $\limsup$ bounded by $F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))$ and is uniformly bounded in t by $F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta))$ (by Theorem 2.3 and positivity of F), the latter being integrable on $B(\xi,\varepsilon)$. Hence, by Fatou's lemma, we can pass to the limit $t \to 0$ and conclude that

$$\int_{B(\xi,\varepsilon)} [F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))]\mathrm{d}\eta \geqslant 0.$$ □

Let us conclude this subsection by proving a maximum principle, but for mappings which are ε-subharmonic. Recall that M is the supremum of $F \circ \boldsymbol{\mu}$ on $\Omega_e \backslash \mathring{\Omega}$ for any $\boldsymbol{\mu} \in L^2_e(\Omega_e, \mathcal{P}(D))$.

Proposition 6.10. *Let $0 < \varepsilon \leqslant \varepsilon_0$ and $\lambda > 0$ be fixed. Then, for a.e. $\xi \in \Omega_e$, one has $F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)) \leqslant M$.*

Proof. Let $\delta > 0$ be fixed and consider $f_\delta : \Omega_e \to \mathbf{R} \cup \{+\infty\}$ defined by $f_\delta(\xi) = F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)) + \delta|\xi - \xi_0|^2$, where ξ_0 is any point of Ω. By strict convexity of the square function and thanks to Proposition 6.9, for a.e. $\xi \in \Omega$,

$$\int_{B(\xi,\varepsilon)} [f_\delta(\eta) - f_\delta(\xi)] \mathrm{d}\eta > 0.$$

In particular, the essential supremum of f_δ cannot be reached on $\mathring{\Omega}$, it must be reached on $\Omega_e \backslash \mathring{\Omega}$. On $\Omega_e \backslash \mathring{\Omega}$ we control the values of $F \circ \boldsymbol{\mu}_{\varepsilon,\lambda}$ by M, in consequence $\operatorname{ess\,sup}_{\Omega_e} f_\delta \leqslant M + C\delta$, where C depends on the diameter of Ω. Sending δ to 0 (along a sequence), we get the result. □

6.3. **Limit to the Dirichlet problem.** In all this subsection, we still assume that F is regular.

The goal is now to pass to the limit and to show that $\boldsymbol{\mu}_{\varepsilon,\lambda}$ converges to $\boldsymbol{\mu}$ a solution of the Dirichlet problem such that $F \circ \boldsymbol{\mu}$ is subharmonic. Recall that Dir_ε Γ-converges to Dir when $\varepsilon \to 0$, see Theorem 3.24. To get subharmonicity, we will need strong convergence, it implies to take first the limit $\varepsilon \to 0$ and then $\lambda \to 0$. But on the other hand, we need a uniform bound on the minimal values of the approximate problems to pass to the limit. To get them implies that we need to produce at least one mapping $\boldsymbol{\mu}$ in $H^1_e(\Omega_e, \mathcal{P}(D))$ such that $\int_{\Omega_e}(F \circ \boldsymbol{\mu}) < +\infty$. To do this, we cannot rely on the Lipschitz extension (Theorem 4.4): there is no way to guarantee that $\int_\Omega (F \circ \boldsymbol{\mu}) < +\infty$. To get this uniform bound, we will take first the limit $\lambda \to 0$ and then $\varepsilon \to 0$ (relying only on weak convergence). It will produce a solution $\tilde{\boldsymbol{\mu}} \in H^1_e(\Omega_e, \mathcal{P}(D))$ of the Dirichlet problem with $\int_{\Omega_e}(F \circ \tilde{\boldsymbol{\mu}}) < +\infty$ but we cannot guarantee subharmonicity of $F \circ \tilde{\boldsymbol{\mu}}$. However it brings uniform bounds and enables us to take the limit $\varepsilon \to 0$, $\lambda \to 0$ and get a solution $\bar{\boldsymbol{\mu}}$ of the Dirichlet problem for which $F \circ \bar{\boldsymbol{\mu}}$ is subharmonic.

We take two sequences $(\varepsilon_n)_{n\in\mathbf{N}}$, $(\lambda_m)_{m\in\mathbf{N}}$ that both converge to 0 while being strictly positive. More precisely we take $\varepsilon_n := \varepsilon_0 2^{-n}$ for any $n \in \mathbf{N}$, thus we always have $\varepsilon_n \leqslant \varepsilon_0$ and $\mathrm{Dir}_{\varepsilon_n}$ converges in an increasing way and Γ-converges to Dir. We will not relabel the sequences when extracting subsequences. Moreover, to avoid heavy notations, we will drop the indexes n and m; and $\lim_{n\to+\infty}$, $\lim_{m\to+\infty}$ will be denoted respectively by $\lim_{\varepsilon\to 0}$ and $\lim_{\lambda\to 0}$.

Proposition 6.11. *Up to extraction, there exists $\tilde{\boldsymbol{\mu}} \in H^1_e(\Omega_e, \mathcal{P}(D))$ such that*

$$\tilde{\boldsymbol{\mu}} := \lim_{\varepsilon\to 0} \left(\lim_{\lambda \to 0} \boldsymbol{\mu}_{\varepsilon,\lambda} \right),$$

where the limits are taken weakly in $L^2_e(\Omega_e, \mathcal{P}(D))$. Moreover, $\tilde{\boldsymbol{\mu}}$ is a minimizer of Dir in the space $H^1_e(\Omega_e, \mathcal{P}(D))$ and

$$\int_{\Omega_e} F(\tilde{\boldsymbol{\mu}}(\xi)) \mathrm{d}\xi < +\infty. \tag{6.5}$$

Proof. The existence of $\tilde{\boldsymbol{\mu}} \in L^2_e(\Omega_e, \mathcal{P}(D))$ is trivial: recall that $L^2_e(\Omega_e, \mathcal{P}(D))$ is compact for the weak convergence. Moreover, using Proposition 6.10, we have that for $\varepsilon \leqslant \varepsilon_0$ and $\lambda > 0$,

$$\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)) \mathrm{d}\xi \leqslant M|\Omega_e|.$$

By the regularity assumption on F, we can pass this inequality to the weak limit and get (6.5).

The minimizing property of $\tilde{\boldsymbol{\mu}}$ is more involved. Assume by contradiction that there exists $\boldsymbol{\nu} \in H^1_e(\Omega, \mathcal{P}(D))$ such that $\mathrm{Dir}(\boldsymbol{\nu}) < \mathrm{Dir}(\tilde{\boldsymbol{\mu}})$. By the Γ-convergence of Dir_ε to Dir and the positivity of F,

one has

$$\mathrm{Dir}(\boldsymbol{\nu}) < \mathrm{Dir}(\tilde{\boldsymbol{\mu}}) \leqslant \liminf_{\varepsilon\to 0}\left(\liminf_{\lambda\to 0}\left(\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon,\lambda}) + \lambda\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi\right)\right).$$

In particular, we can choose $\varepsilon > 0$ small enough such that (by monotonicity of Dir_ε)

$$\mathrm{Dir}_\varepsilon(\boldsymbol{\nu}) \leqslant \mathrm{Dir}(\boldsymbol{\nu}) < \liminf_{\lambda\to 0}\left(\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon,\lambda}) + \lambda\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi\right).$$

We regularize $\boldsymbol{\nu}$ in the following way: for $t > 0$, we denote by $\boldsymbol{\nu}_t := (\mathbb{1}_{\mathring{\Omega}}\Phi_t^D)\boldsymbol{\nu}$ the element of $L^2_e(\Omega_e, \mathcal{P}(D))$ for which the heat flow on D has been followed only in $\mathring{\Omega}$: in other words, for any $t > 0$,

$$\boldsymbol{\nu}_t(\xi) := \begin{cases} (\Phi_t^D)[\boldsymbol{\nu}(\xi)] & \text{if } \xi \in \mathring{\Omega}, \\ \boldsymbol{\nu}(\xi) = \boldsymbol{\mu}_e(\xi) & \text{if } \xi \in \Omega_e\backslash\mathring{\Omega}. \end{cases}$$

Clearly, $\boldsymbol{\nu}_t \in L^2_e(\Omega_e, \mathcal{P}(D))$. Moreover, as $W_2(\boldsymbol{\nu}_t(\xi), \boldsymbol{\nu}(\xi)) \leqslant \omega(t)$ with $\omega(t) \to 0$ as $t \to 0$ (see Proposition 2.5), we see that $\boldsymbol{\nu}_t$ converges strongly in $L^2_e(\Omega_e, \mathcal{P}(D))$ to $\boldsymbol{\nu}$. In particular, thanks to the continuity of Dir_ε, there exists t small enough such that

$$\mathrm{Dir}_\varepsilon(\boldsymbol{\nu}_t) < \liminf_{\lambda\to 0}\left(\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon,\lambda}) + \lambda\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi\right).$$

Because of the standard $L^\infty - L^1$ estimate for the heat flow [6, Section 7], one has that $\{\boldsymbol{\nu}_t(\xi) \,:\, \xi \in \mathring{\Omega}\}$ is included in a bounded set of $L^\infty(D) \cap \mathcal{P}(D)$. In particular, $F \circ \boldsymbol{\nu}_t$ is bounded on $\mathring{\Omega}$. As it is also bounded on $\Omega_e\backslash\mathring{\Omega}$ by M, we see that $\int_{\Omega_e} F \circ \boldsymbol{\nu}_t < +\infty$. Hence, for some λ small enough,

$$\mathrm{Dir}_\varepsilon(\boldsymbol{\nu}_t) + \lambda\int_{\Omega_e} F(\boldsymbol{\nu}_t(\xi))\mathrm{d}\xi < \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon,\lambda}) + \lambda\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi,$$

which is a contradiction with the optimality of $\boldsymbol{\mu}_{\varepsilon,\lambda}$. □

Proposition 6.12. *Up to extraction, there exists $\bar{\boldsymbol{\mu}} \in H^1_e(\Omega_e, \mathcal{P}(D))$ such that*

$$\bar{\boldsymbol{\mu}} := \lim_{\lambda\to 0}\left(\lim_{\varepsilon\to 0}\boldsymbol{\mu}_{\varepsilon,\lambda}\right),$$

where the limits are taken strongly in $L^2_e(\Omega_e, \mathcal{P}(D))$. Moreover, $\bar{\boldsymbol{\mu}}$ is a minimizer of Dir *in the space $H^1_e(\Omega_e, \mathcal{P}(D))$ and for any other minimizer $\boldsymbol{\nu}$ of* Dir *in $H^1_e(\Omega_e, \mathcal{P}(D))$,*

$$\int_{\Omega_e} F(\bar{\boldsymbol{\mu}}(\xi))\mathrm{d}\xi \leqslant \liminf_{\lambda\to 0}\left(\liminf_{\varepsilon\to 0}\left(\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi\right)\right) \leqslant \int_{\Omega_e} F(\boldsymbol{\nu}(\xi))\mathrm{d}\xi. \tag{6.6}$$

Proof. Using $\tilde{\boldsymbol{\mu}}$ as a competitor in the approximate problem, given the monotonicity of Dir_ε, one has that

$$\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon,\lambda}) + \lambda\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi \leqslant \mathrm{Dir}(\tilde{\boldsymbol{\mu}}) + \lambda\int_{\Omega_e} F(\tilde{\boldsymbol{\mu}}(\xi))\mathrm{d}\xi \leqslant C,$$

where the constant C is uniform in $\varepsilon > 0$ and $0 < \lambda \leqslant 1$. In particular, using the Rellich-like theorem (Proposition 6.5), we see that, up to extraction, $\boldsymbol{\mu}_{\varepsilon,\lambda}$ converges strongly in $L^2_e(\Omega_e, \mathcal{P}(D))$ to some $\bar{\boldsymbol{\mu}}_\lambda$ when $\varepsilon \to 0$. Moreover, by Γ-convergence of Dir_ε and the regularity of F,

$$\mathrm{Dir}(\bar{\boldsymbol{\mu}}_\lambda) + \lambda\int_{\Omega_e} F(\bar{\boldsymbol{\mu}}_\lambda(\xi))\mathrm{d}\xi \leqslant \liminf_{\varepsilon\to 0}\left(\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon,\lambda}) + \lambda\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi\right) \leqslant C. \tag{6.7}$$

Hence, we have a uniform bound on $\mathrm{Dir}(\bar{\boldsymbol{\mu}}_\lambda)$, and we can apply Rellich theorem (Proposition 6.4) to see that $\bar{\boldsymbol{\mu}}_\lambda$ converges strongly in $L^2(\Omega_e, \mathcal{P}(D))$ to some $\bar{\boldsymbol{\mu}} \in H^1_e(\Omega_e, \mathcal{P}(D))$ when $\lambda \to 0$. Moreover,

using the lower semi-continuity of Dir and positivity of F,

$$\mathrm{Dir}(\bar{\boldsymbol{\mu}}) \leqslant \liminf_{\lambda\to 0}\left(\mathrm{Dir}(\bar{\boldsymbol{\mu}}_\lambda) + \lambda\int_{\Omega_e} F(\bar{\boldsymbol{\mu}}_\lambda(\xi))\mathrm{d}\xi\right). \tag{6.8}$$

Let us assume by contradiction that $\bar{\boldsymbol{\mu}}$ is not a minimizer of Dir. Thanks to Proposition 6.11, it boils down to assume that $\mathrm{Dir}(\tilde{\boldsymbol{\mu}}) < \mathrm{Dir}(\bar{\boldsymbol{\mu}})$. In particular, as $F\circ\tilde{\boldsymbol{\mu}}$ is integrable on Ω_e and with the help of (6.8), it means that there exists λ small enough such that

$$\mathrm{Dir}(\tilde{\boldsymbol{\mu}}) + \lambda\int_{\Omega_e} F(\tilde{\boldsymbol{\mu}}(\xi))\mathrm{d}\xi < \mathrm{Dir}(\bar{\boldsymbol{\mu}}_\lambda) + \lambda\int_{\Omega_e} F(\bar{\boldsymbol{\mu}}_\lambda(\xi))\mathrm{d}\xi.$$

Using the fact that $\mathrm{Dir}_\varepsilon(\tilde{\boldsymbol{\mu}}) \to \mathrm{Dir}(\tilde{\boldsymbol{\mu}})$ to handle the l.h.s. and (6.7) to deal with the r.h.s., we see that for $\varepsilon > 0$ small enough,

$$\mathrm{Dir}_\varepsilon(\tilde{\boldsymbol{\mu}}) + \lambda\int_{\Omega_e} F(\tilde{\boldsymbol{\mu}}(\xi))\mathrm{d}\xi < \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon,\lambda}) + \lambda\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi,$$

which is a contradiction with the optimality of $\boldsymbol{\mu}_{\varepsilon,\lambda}$. Hence, $\bar{\boldsymbol{\mu}}$ is a minimizer of Dir over $H^1_e(\Omega_e,\mathcal{P}(D))$.

Remark that in (6.6) the first inequality is a consequence of the fact that F is regular. Assume by contradiction that there exists $\boldsymbol{\nu}\in H^1_e(\Omega_e,\mathcal{P}(D))$ a minimizer of Dir such that (6.6) does not hold. In particular as $\mathrm{Dir}(\bar{\boldsymbol{\mu}}) = \mathrm{Dir}(\boldsymbol{\nu})$, and by Γ-convergence of Dir_ε and lower semi-continuity of Dir,

$$\mathrm{Dir}(\boldsymbol{\nu}) = \mathrm{Dir}(\bar{\boldsymbol{\mu}}) \leqslant \liminf_{\lambda\to 0}\left(\liminf_{\varepsilon\to 0}\left(\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon,\lambda})\right)\right),$$

thus one can write that for some λ small enough,

$$\mathrm{Dir}(\boldsymbol{\nu}) + \lambda\int_{\Omega_e} F(\boldsymbol{\nu}(\xi))\mathrm{d}\xi < \liminf_{\varepsilon\to 0}\left(\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_{\varepsilon,\lambda}) + \lambda\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi\right):$$

it leads to the same contradiction than before by taking $\varepsilon > 0$ small enough. □

Now, the key result to get subharmonicity of $F\circ\bar{\boldsymbol{\mu}}$ is that we can pass at the pointwise limit the quantity $F\circ\boldsymbol{\mu}_{\varepsilon,\lambda}$.

Proposition 6.13. *For a.e. $\xi\in\Omega$,*

$$F(\bar{\boldsymbol{\mu}}(\xi)) = \lim_{\lambda\to 0}\left(\lim_{\varepsilon\to 0}\left(F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\right)\right).$$

Proof. As the convergence of $\boldsymbol{\mu}_{\varepsilon,\lambda}$ to $\bar{\boldsymbol{\mu}}$ holds strongly in $L^2_e(\Omega,\mathcal{P}(D))$, we can, up to extraction, assume that it holds a.e. In other words, for a.e. $\xi\in\Omega$,

$$\bar{\boldsymbol{\mu}}(\xi) = \lim_{\lambda\to 0}\left(\lim_{\varepsilon\to 0}\left(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)\right)\right)$$

in $\mathcal{P}(D)$. By lower semi-continuity of F on $\mathcal{P}(D)$, the inequality

$$F(\bar{\boldsymbol{\mu}}(\xi)) \leqslant \liminf_{\lambda\to 0}\left(\liminf_{\varepsilon\to 0}\left(F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\right)\right)$$

holds for a.e. $\xi\in\Omega$. On the other hand, use (6.6) with $\boldsymbol{\nu} = \bar{\boldsymbol{\mu}}$: up to extraction one has

$$\int_{\Omega_e} F(\bar{\boldsymbol{\mu}}(\xi))\mathrm{d}\xi = \lim_{\lambda\to 0}\left(\lim_{\varepsilon\to 0}\left(\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))\mathrm{d}\xi\right)\right).$$

By combining the two equations above (recall that all the functions $F\circ\boldsymbol{\mu}_{\varepsilon,\lambda}$ and $F\circ\bar{\boldsymbol{\mu}}$ are positive and bounded above by M thanks to Proposition 6.10), we reach the desired conclusion (this is just an adaptation of the proof of Scheffé's lemma). □

Proposition 6.14. *The function $F \circ \bar{\boldsymbol{\mu}}$ is subharmonic on $\mathring{\Omega}$. Moreover,*

$$\operatorname*{ess\,sup}_{\Omega}(F \circ \bar{\boldsymbol{\mu}}) \leqslant M.$$

Proof. The fact that the essential supremum of $F \circ \bar{\boldsymbol{\mu}}$ is bounded by M is a simple combination of Propositions 6.10 and 6.13. For the subharmonicity, take $\psi \in C_c^\infty(\mathring{\Omega})$ a smooth and positive function compactly supported in $\mathring{\Omega}$. For $0 < \varepsilon \leqslant \varepsilon_0$ small enough, one has, thanks to Proposition 6.9,

$$\int_{\Omega_e} \psi(\xi) \left(\frac{1}{\varepsilon^{d+2}} \int_{B(\xi,\varepsilon)} [F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\eta)) - F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))] \mathrm{d}\eta \right) \mathrm{d}\xi \geqslant 0.$$

Performing a discrete integration by parts (which is possible if ε is smaller than the distance between $\partial\Omega$ and the support of ψ), one sees that

$$\int_{\Omega_e} F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi)) \left(\frac{1}{\varepsilon^{d+2}} \int_{B(\xi,\varepsilon)} [\psi(\eta) - \psi(\xi)] \mathrm{d}\eta \right) \mathrm{d}\xi \geqslant 0.$$

Now send $\varepsilon \to 0$ and then $\lambda \to 0$. By smoothness of ψ, the quantity $\varepsilon^{-(d+2)} \int_{B(\xi,\varepsilon)} [\psi(\eta) - \psi(\xi)] \mathrm{d}\eta$ converges to $\Delta\psi(\xi)$ (up to a multiplicative constant). On the other hand, $F(\boldsymbol{\mu}_{\varepsilon,\lambda}(\xi))$ converges pointwisely to $F(\bar{\boldsymbol{\mu}})$ (see Proposition 6.13) while being bounded by M. By Lebesgue dominated convergence theorem,

$$\int_{\Omega_e} F(\boldsymbol{\mu}(\xi)) \Delta\psi(\xi) \mathrm{d}\xi \geqslant 0,$$

which exactly means that $F \circ \boldsymbol{\mu}$ is subharmonic in the sense of distributions as ψ is an arbitrary smooth and positive function. □

Now we can conclude:

Proof of Theorem 6.3 if F is regular. We take $\boldsymbol{\mu}$ the restriction of $\bar{\boldsymbol{\mu}}$ to Ω. Thanks to Proposition 6.6, the fact that $\bar{\boldsymbol{\mu}}$ is a minimizer of Dir among $H^1_e(\Omega_e, \mathcal{P}(D))$ is translated into the fact that $\boldsymbol{\mu}$ is a solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$. The subharmonicity and the upper bound of $F \circ \bar{\boldsymbol{\mu}}$ are preserved by restriction. To get the minimality of $\int_\Omega (F \circ \bar{\boldsymbol{\mu}})$, we just use (6.6). □

6.4. **Simplifications in the continuous case.** In this subsection, we assume that F is continuous. In particular, as $\mathcal{P}(D)$ is compact, it implies that F is bounded. The proof is simpler because we do not need to add the term $\lambda \int F \circ \boldsymbol{\mu}$ in the approximate problem. Indeed, strong convergence in $L^2(\Omega, \mathcal{P}(D))$ of a sequence $\boldsymbol{\mu}_n$ to $\boldsymbol{\mu}$ implies, up to extraction, the convergence a.e. of $(F \circ \boldsymbol{\mu}_n)$ to $(F \circ \boldsymbol{\mu})$.

We define $\Omega_e, \boldsymbol{\mu}_e$ and the functional spaces $L^2_e(\Omega_e, \mathcal{P}(D)), H^1_e(\Omega_e, \mathcal{P}(D))$ as in the beginning of Subsection 6.2.

Proof of Theorem 6.3 if F is continuous. For any $\varepsilon > 0$, we take $\boldsymbol{\mu}_\varepsilon \in L^2_e(\Omega_e, \mathcal{P}(D))$ a minimizer of Dir_ε over $L^2_e(\Omega_e, \mathcal{P}(D))$.

We can still apply Proposition 6.8 and conclude that for a.e. $\xi \in \Omega$, $\boldsymbol{\mu}_\varepsilon(\xi)$ is a barycenter of the $\boldsymbol{\mu}_\varepsilon(\eta)$ for $\eta \in B(\xi, \varepsilon)$. The proof of Jensen's inequality (Proposition 6.9) works in the same way as F is bounded on $\mathcal{P}(D)$. Hence, the maximum principle given by Proposition 6.10 is still true as it is only implied by Proposition 6.9.

To pass to the limit $\varepsilon \to 0$, we use the fact that (along an appropriate sequence) Dir_ε Γ-converges to Dir. Hence, up to extraction, $\boldsymbol{\mu}_\varepsilon$ converges to $\bar{\boldsymbol{\mu}}$ which is a minimizer of Dir over $L^2_e(\Omega_e, \mathcal{P}(D))$. Thanks to Proposition 6.5, the convergence takes place strongly in $L^2_e(\Omega_e, \mathcal{P}(D))$ and a.e. By continuity of F, we deduce that the conclusion of Proposition 6.13 still holds: $F \circ \boldsymbol{\mu}_\varepsilon$ converges a.e. to

$F \circ \bar{\boldsymbol{\mu}}$ as $\varepsilon \to 0$. Thus the proof of Proposition 6.14 works exactly in the same way and it is enough to take for $\boldsymbol{\mu}$ the restriction of $\bar{\boldsymbol{\mu}}$ to Ω. □

7. Examples

To conclude, we give examples of situations where the computation of harmonic mappings can be done explicitly. The first one is rather simple: when D is a segment of $\mathbf{R}$ the space $\mathcal{P}(D)$ has a structure of Hilbert space, hence the study is considerably simpler and all the machinery developed above is too heavy. The second one is trickier: we restrict ourselves to a family of elliptically contoured distributions, which is a geodesically convex subset of finite dimension. Thus we end up with mappings valued in a finite-dimensional Riemannian manifold, on which we can write explicit Euler-Lagrange equations.

Before studying these examples, let us say that the case where the measures on $\partial\Omega$ are Dirac masses has already been treated. Indeed, we have already mentioned [9, Theorem 3.1] which states that if $f : \Omega \to D$ is harmonic, then the measure $\boldsymbol{\mu}_f$ (defined by $\boldsymbol{\mu}_f(\xi) = \delta_{f(\xi)}$ for all $\xi \in \Omega$) is a measured-valued harmonic mapping. This result has been extended in [21, Theorem 3.3] to the case where space D is a simply connected manifold with negative curvature.

7.1. **One dimensional target.** In this subsection, we assume that $D = I = [0,1]$ is the unit interval. The important point is that the space $\mathcal{P}(I)$ has a very simple structure: the right object to characterize an element $\mu \in \mathcal{P}(D)$ is its inverse distribution function $F_\mu^{[-1]} : [0,1] \to [0,1]$ defined by

$$F_\mu^{[-1]}(t) := \inf\{x \in [0,1] \ : \ \mu([0,x]) \geqslant t\}.$$

It is well known that $F_\mu^{[-1]}$ is increasing, right continuous, and that there is a bijection between the set of increasing and right continuous mappings $[0,1] \to [0,1]$ and $\mathcal{P}(I)$. Moreover, for any $\mu, \nu \in \mathcal{P}(I)$, one has (see for instance [27, Proposition 2.17])

$$W_2^2(\mu,\nu) = \int_0^1 |F_\mu^{[-1]}(t) - F_\nu^{[-1]}(t)|^2 \mathrm{d}t. \tag{7.1}$$

Introduce the Hilbert space $\mathcal{H} := L^2([0,1])$ with its usual norm (denoted by $|\cdot|_{\mathcal{H}}$) and the subspace $\mathcal{H}_{\mathrm{i}}$ of increasing functions: if $f \in \mathcal{H}$, then we say that $f \in \mathcal{H}_{\mathrm{i}}$ if $f(t) \in [0,1]$ for a.e. $t \in [0,1]$ and if for any $0 \leqslant t_1 < t_2 \leqslant t_3 < t_4 \leqslant 1$, one has

$$\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} f(t)\mathrm{d}t \leqslant \frac{1}{t_4 - t_3} \int_{t_3}^{t_4} f(t)\mathrm{d}t$$

Notice that $\mathcal{H}_{\mathrm{i}}$ is clearly a convex and closed subset of $\mathcal{H}$. Any $f \in \mathcal{H}_{\mathrm{i}}$ has a unique increasing and right continuous representative. Indeed, take the representative given by the Lebesgue differentiation theorem: except on a subset N which is negligible, it is increasing. Then, on N and on any point of discontinuity, choose the right limit. Uniqueness is easy as any increasing and right continuous representative is continuous except at a countable number of points. This discussion can be summarized in the following proposition.

Proposition 7.1. *If we define $\Psi(\mu) := F_\mu^{[-1]}$, then Ψ is a one-to-one isometry between $\mathcal{P}(I)$ and $\mathcal{H}_{\mathrm{i}}$.*

Now we need to make the bridge between the space $H^1(\Omega, \mathcal{P}(I))$ and the space $H^1(\Omega, \mathcal{H})$. To this extent, we rely on the characterization by Reshetnyak.

Proposition 7.2. *Let $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(I))$. Then $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(I))$ if and only if $(\Psi \circ \boldsymbol{\mu}) \in H^1(\Omega, \mathcal{H})$ and in this case*

$$g_{\boldsymbol{\mu}} = |\nabla(\Psi \circ \boldsymbol{\mu})|_{\mathcal{H}}.$$

The proof is left to the reader: one can check that the definition of Reshetnyak coincides with the usual definition of $H^1(\Omega, \mathcal{H})$. Thus, we can say the following:

Theorem 7.3. *Let $\boldsymbol{\mu}_l : \partial\Omega \to \mathcal{P}(I)$ a given Lipschitz mapping. Then there exists a unique $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(I))$ solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$. Moreover, $(\Psi \circ \boldsymbol{\mu})$ is the solution of the minimization problem*

$$\min\left\{\int_\Omega |\nabla f(\xi)|^2_{\mathcal{H}} \mathrm{d}\xi \ : \ f \in H^1(\Omega, \mathcal{H}) \ \textit{and} \ f|_{\partial\Omega} = \Psi \circ \boldsymbol{\mu}_l\right\}. \tag{7.2}$$

Proof. Everything relies on this simple observation: combing Proposition 7.2 and Theorem 3.17 (which gives the expression of $g_{\boldsymbol{\mu}}$), we see that

$$\mathrm{Dir}(\boldsymbol{\mu}) := \int_\Omega |\nabla(\Psi \circ \boldsymbol{\mu})(\xi)|^2_{\mathcal{H}} \mathrm{d}\xi.$$

With the help of Proposition 6.6, one can be convinced that to impose $\mathrm{BT}_{\boldsymbol{\mu}} = \mathrm{BT}_{\boldsymbol{\mu}_b}$ is the same than to say that the the trace of $(\Psi \circ \boldsymbol{\mu})$ is $(\Psi \circ \boldsymbol{\mu}_l)$. Then, one takes f to be the unique harmonic extension of $(\Psi \circ \boldsymbol{\mu}_l)$ in $H^1(\Omega, \mathcal{H})$: it is the minimizer of (7.2). By the maximum principle, as $(\Psi \circ \boldsymbol{\mu}_l) \in \mathcal{H}_{\mathrm{i}}$ on $\partial\Omega$, it is clear that $f \in H^1(\Omega, \mathcal{H}_{\mathrm{i}})$. Thus, we can simply set $\boldsymbol{\mu} := \Psi^{-1} \circ f$. □

7.2. **Family of elliptically contoured distributions.** We finish by studying the case where the boundary values belong to a family of elliptically contoured distributions: they are parametrized by their covariance matrix. It can be seen as a generalization of the case where the measures are Gaussian. In this subsection, we would like to show that at least one solution of the harmonic problem is valued in the family of elliptically contoured distributions if it is the case for the boundary values, and to give a full solution (existence, uniqueness and Euler-Lagrange equation) under the additional assumption that the covariance matrices of the boundary values are non singular.

We will deal with centered measures (i.e. measures with zero mean) because the contribution of the mean to the Dirichlet energy can be handled independently. More precisely if $\mu \in \mathcal{P}(D)$ we denote by $m(\mu) := \int_D x\mu(\mathrm{d}x) \in D$ its mean and μ_0 the centered measured defined as the push forward of μ by $(x \mapsto x - m(\mu))$. It is well known [31, Problem 1] that if μ, ν are two probability measures then

$$W_2^2(\mu, \nu) = W_2^2(\mu_0, \nu_0) + |m(\mu) - m(\nu)|^2.$$

If $\boldsymbol{\mu} \in L^2(\Omega, \mathcal{P}(D))$, we use the formula above on $\mathrm{Dir}_\varepsilon(\boldsymbol{\mu})$:

$$\mathrm{Dir}_\varepsilon(\boldsymbol{\mu}) = \mathrm{Dir}_\varepsilon(\boldsymbol{\mu}_0) + C_p \int_{\Omega\times\Omega} \frac{|m(\boldsymbol{\mu}(\xi)) - m(\boldsymbol{\mu}(\eta))|^2}{2\varepsilon^{p+2}} \mathbb{1}_{|\xi-\eta|\leqslant\varepsilon} \mathrm{d}\xi\mathrm{d}\eta.$$

Then, sending ε to 0 and using [17, Theorem 8.3.1] to handle the part involving the Dirichlet energy of the means, one sees that

$$\mathrm{Dir}(\boldsymbol{\mu}) = \mathrm{Dir}(\boldsymbol{\mu}_0) + \frac{1}{2}\int_\Omega |\nabla[m(\boldsymbol{\mu})](\xi)|^2 \mathrm{d}\xi.$$

The term involving $m(\boldsymbol{\mu})$ is easy to minimize (because $m(\boldsymbol{\mu})$ is a vector-valued function, it boils down to take the harmonic extension) and it can be done independently from the term involving $\mathrm{Dir}(\boldsymbol{\mu}_0)$. In other words, it is not restrictive to work only with centered measures.

Let us go back to the family of elliptically contoured distributions. As we have assumed that D is compact, we cannot work with non compactly supported measures, in particular with Gaussian measures. For the rest of the subsection, we fix $\rho \in L^1(\mathbf{R}^q)$ a positive function compactly supported such that $\rho\mathcal{L}_D$ is a probability measure with zero mean and the identity matrix as a covariance

matrix. Recall that the covariance matrix $\mathrm{cov}(\mu)$ of a measure $\mu \in \mathcal{P}(\mathbf{R}^q)$ is defined as: for any $i, j \in \{1, 2, \dots, q\}$,

$$\mathrm{cov}(\mu)_{ij} := \int_{\mathbf{R}^q} x_i x_j \mu(\mathrm{d}x).$$

For technical reasons, we also assume that ρ is radial and that the Boltzmann entropy of $\rho\mathcal{L}_D$ (see (7.6) below) is finite. Let us denote by $S_q(\mathbf{R})$ the set of symmetric $q \times q$ matrices and $S_q^+(\mathbf{R}) \subset S_q(\mathbf{R})$ the set of symmetric and positive $q \times q$ matrices. The space $S_q(\mathbf{R})$ is equiped with its canonical scalar product $\langle \cdot, \cdot \rangle$ defined by $\langle A, B \rangle = \mathrm{Tr}(AB)$. The unique symmetric square root of a matrix $A \in S_q^+(\mathbf{R})$ is denoted by $A^{1/2}$. Instead of parametrizing measures by their covariance matrix we will do it by the square root of their covariance matrix, i.e. by their standard deviation: it is more natural for homogeneity reasons and the formulas are slightly simpler.

Definition 7.4. *For any $A \in S_q^+(\mathbf{R})$ we denote by $\rho_A \in \mathcal{P}(\mathbf{R}^q)$ the image measure of $\rho\mathcal{L}_D$ by the map $x \in \mathbf{R}^q \mapsto Ax \in \mathbf{R}^q$.*

The set of all ρ_A for $A \in S_q^+(\mathbf{R})$ is denoted by $\mathcal{P}_{\mathrm{ec}}(\mathbf{R}^q)$ and is called a family of elliptically contoured distributions (with reference measure $\rho\mathcal{L}_D$).

Thanks to the normalization of ρ, the measure ρ_A has zero mean and covariance matrix A^2. Notice that if A is invertible then

$$\rho_A(\mathrm{d}x) := \frac{1}{\det(A)} \rho\left(A^{-1}x\right) \mathrm{d}x.$$

We would recover the Gaussian case by taking $\rho(x) = (2\pi)^{-q/2} \exp(-|x|^2/2)$, but this function is not compactly supported.

The crucial tool to establish that a harmonic extension of a mapping valued in a family of elliptically contoured distributions stays in the same family is the existence of a retraction on the set $\mathcal{P}_{\mathrm{ec}}(\mathbf{R}^q)$. Let us call $\mathcal{P}_2(\mathbf{R}^q)$ the set of probability measures on $\mathbf{R}^q$ with finite second moment.

Definition 7.5. *Let $R : \mathcal{P}_2(\mathbf{R}^q) \to \mathcal{P}_{\mathrm{ec}}(\mathbf{R}^q)$ the application defined by $R(\mu) := \rho_A$, where $A := \mathrm{cov}(\mu)^{1/2}$ is the symmetric square root of the covariance matrix of μ.*

Proposition 7.6. *The application $R : \mathcal{P}_2(\mathbf{R}^q) \to \mathcal{P}_{\mathrm{ec}}(D)$ leaves $\mathcal{P}_{\mathrm{ec}}(\mathbf{R}^q)$ unchanged and is a contraction (i.e. is 1-Lipschitz) provided that $\mathcal{P}_2(\mathbf{R}^q)$ and $\mathcal{P}_{\mathrm{ec}}(\mathbf{R}^q)$ are endowed with the quadratic Wasserstein distance W_2.*

Proof. The first part is obvious by the way we normalize ρ. The second part is a reformulation of Theorem 2.1 and Theorem 2.4 of [13]. □

Let us prove state and prove here an easy technical lemma which will be crucial in the sequel.

Lemma 7.7. *Let $\boldsymbol{\mu}_l : \partial\Omega \to \mathcal{P}(D)$ a Lipschitz function and $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ such that $\boldsymbol{\mu}|_{\partial\Omega} = \boldsymbol{\mu}_l$. Let $M : \mathcal{P}(D) \to \mathcal{P}(D)$ a 1-Lipschitz mapping. Then $M \circ \boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ with $(M \circ \boldsymbol{\mu})|_{\partial\Omega} = (M \circ \boldsymbol{\mu}_l)$ and*

$$\mathrm{Dir}(M \circ \boldsymbol{\mu}) \leqslant \mathrm{Dir}(\boldsymbol{\mu}).$$

Proof. As M is a contraction and from the definition of Dir_ε it is obvious that

$$\mathrm{Dir}_\varepsilon(M \circ \boldsymbol{\mu}) \leqslant \mathrm{Dir}_\varepsilon(\boldsymbol{\mu})$$

holds for any $\varepsilon > 0$. Then it is sufficient to send ε to 0. To get the assertion involving the boundary conditions, one can use for instance Proposition 6.6. □

As we work in the compactly supported case, we add some assumption that D is large enough in order for the boundary of D to be invisible. More precisely, the following lemma will help us to handle the finiteness of D.

Lemma 7.8. *Let $\tilde{D} \subset D$ be a convex compact subset of D. Let $\boldsymbol{\mu}_l : \partial\Omega \to \mathcal{P}(\tilde{D})$ be a Lipschitz mapping. If $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(\tilde{D}))$ is a solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$, then, seen as an element of $H^1(\Omega, \mathcal{P}(D))$ (extending $\boldsymbol{\mu}$ by 0 on $D\backslash\tilde{D}$), $\boldsymbol{\mu}$ is also a solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$ (with $\boldsymbol{\mu}_l$ seen as a mapping valued in $\mathcal{P}(D)$).*

Proof. It relies on a simple observation. Let $P_{\tilde{D}} : D \to \tilde{D}$ be the Euclidean projection on $\tilde{D}$. One has that $\nu \mapsto P_{\tilde{D}}\#\nu$ is a 1-Lipschitz function from $(\mathcal{P}(D), W_2)$ to $(\mathcal{P}(\tilde{D}), W_2)$ which leaves the boundary values $\boldsymbol{\mu}_l$ unchanged. Thus we can apply Lemma 7.7 to see that $P_{\tilde{D}}$ maps any competitor from $H^1(\Omega, \mathcal{P}(D))$ into a competitor in $H^1(\Omega, \mathcal{P}(\tilde{D}))$. □

We will say that $\tilde{D} \subset D$ is compatible with ρ if it is a compact convex subset of D and for any $\mu \in \mathcal{P}(\tilde{D})$, one has $R(\mu) \in \mathcal{P}(D)$. It holds if D is large enough compared to $\tilde{D}$ and the diameter of the support of ρ. In the sequel, we will use the notations $\mathcal{P}_{\mathrm{ec}}(\tilde{D}) := \mathcal{P}(\tilde{D}) \cap \mathcal{P}_{\mathrm{ec}}(\mathbf{R}^q)$ and $\mathcal{P}_{\mathrm{ec}}(D) := \mathcal{P}(D) \cap \mathcal{P}_{\mathrm{ec}}(\mathbf{R}^q)$

Theorem 7.9. *Take $\tilde{D} \subset D$ compatible with ρ. Let $\boldsymbol{\mu}_l : \partial\Omega \to \mathcal{P}_{\mathrm{ec}}(\tilde{D})$ a Lipschitz mapping valued in the family of elliptically contoured distributions. Then there exists $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(D))$ a solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$ such that $\boldsymbol{\mu}(\xi) \in \mathcal{P}_{\mathrm{ec}}(D)$ for a.e. $\xi \in \Omega$.*

The assumption that $\tilde{D}$ is compatible with D can be translated in the fact that the supports of the $\boldsymbol{\mu}_l(\xi)$ for $\xi \in \partial\Omega$ are small compared to D.

Proof. Let $\tilde{\boldsymbol{\mu}}$ be a solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$, it exists thanks to Theorem 4.4 and Theorem 4.3. According to Lemma 7.8, we can choose $\tilde{\boldsymbol{\mu}}$ such that $\tilde{\boldsymbol{\mu}} \in \mathcal{P}(\tilde{D})$ a.e. As R is a contraction which leaves the boundary values unchanged, it is clear thanks to Lemma 7.7 that $\boldsymbol{\mu} := R \circ \tilde{\boldsymbol{\mu}}$ is a solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$. By construction, $\boldsymbol{\mu}$ is valued in $\mathcal{P}_{\mathrm{ec}}(\mathbf{R}^q)$ and also in $\mathcal{P}(D)$ as $\tilde{D}$ is compatible with ρ. □

We believe that, conducting a careful analysis, one can prove that all solutions of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$ are valued in $\mathcal{P}_{\mathrm{ec}}(D)$.

Now, we want to go further and give a more explicit description of the solution valued in the family of elliptically contoured distributions. To this extent, we rely on the fact that the manifold $S_q^+(\mathbf{R})$, when endowed with the distance induced by W_2 through the application $A \mapsto \rho_A$, has a structure of Riemannian manifold, at least when restricted to the set of non singular matrix.

More precisely, if A and B are in $S_q^+(\mathbf{R})$ it is known (see for instance [13]) that

$$W_2^2(\rho_A, \rho_B) = \mathrm{Tr}\left(A^2 + B^2 - 2(AB^2A)^{1/2}\right).$$

(When A and B are replaced by their square, the r.h.s. is sometimes called the Bures metric.) Notice that if A and B commute then $W_2^2(\rho_A, \rho_B) = \mathrm{Tr}((A-B)^2)$, which justifies that the right choice is to parametrize elements of the family of elliptically contoured distributions by the square root of their covariance matrix. Denote by $S_q^{++}(\mathbf{R})$ the set of $q \times q$ symmetric definite positive matrices. If $A \in S_q^+(\mathbf{R})$, we can define the linear map $L_A : S_q(\mathbf{R}) \to S_q(\mathbf{R})$ by $L_A := A \otimes \mathrm{Id} + \mathrm{Id} \otimes A$. More explicitly for any $H \in S_q(\mathbf{R})$

$$L_A(H) = AH + HA.$$

The map L_A is symmetric, and is moreover positive definite as soon as A is positive definite (in this case in particular it is invertible). If A is diagonal, then L_A is also diagonal in the canonical basis for matrices. In particular, if A and B commute, then L_A and L_B also commute. If $A \in S_q^{++}(\mathbf{R})$ and $B \in S_q(\mathbf{R})$, a lengthy but straightforward computation leads to

$$\lim_{t\to 0} \frac{W_2^2(\rho_A, \rho_{A+tB})}{t^2} = \langle B, \mathfrak{g}_A(B)\rangle \tag{7.3}$$

where $\mathfrak{g}_A : S_q(\mathbf{R}) \to S_q(\mathbf{R})$ is a linear map defined as

$$\mathfrak{g}_A := \frac{1}{2}(L_A)^2(L_{A^2})^{-1}.$$

More explicitly, if A is a diagonal matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_q$ and $B = (B_{ij})_{1\leqslant i,j\leqslant q}$ then

$$\langle B, \mathfrak{g}_A(B)\rangle = \frac{1}{2}\sum_{1\leqslant i,j\leqslant q}\frac{(\lambda_i+\lambda_j)^2}{\lambda_i^2+\lambda_j^2}B_{ij}^2.$$

Notice that $\mathfrak{g}_A$ always defines a scalar product on the space $S_q(\mathbf{R})$. As a consequence, we can define the Riemannian manifold $(S_q^{++}(\mathbf{R}), \mathfrak{g})$: at each point $A \in S_q^{++}(\mathbf{R})$ the tangent space, which is isomorphic to $S_q(\mathbf{R})$, is endowed with the scalar product $\mathfrak{g}_A$. If we do that, as we already know that $\mathcal{P}_{\mathrm{ec}}(\mathbf{R}^q)$ is a geodesic space and thanks to (7.3), we see that the Riemannian distance $d_{\mathfrak{g}}$ induced by $\mathfrak{g}$ satisfies $d_{\mathfrak{g}}(A,B) = W_2(\rho_A, \rho_B)$ for any $A, B \in S_q^{++}(\mathbf{R})$. From this identity we can derive the following consequence. Take $\mathbf{A} \in H^1(\Omega, (S_q^{++}(\mathbf{R}), \mathfrak{g}))$ a matrix-valued function and define $\rho_{\mathbf{A}} \in L^2(\Omega, \mathcal{P}(D))$ by $\rho_{\mathbf{A}}(\xi) = \rho_{\mathbf{A}(\xi)}$ for a.e. $\xi \in \Omega$. Then $\rho_{\mathbf{A}} \in H^1(\Omega, \mathcal{P}(D))$ and

$$\mathrm{Dir}(\rho_{\mathbf{A}}) = \int_\Omega \frac{1}{2}\sum_{\alpha=1}^p \langle \partial_\alpha \mathbf{A}(\xi), \mathfrak{g}_{\mathbf{A}(\xi)}(\partial_\alpha \mathbf{A}(\xi))\rangle \mathrm{d}\xi. \tag{7.4}$$

To justify this identity, one can use for instance the formulation with Dir_ε (Theorem 3.24), replace the Wasserstein distance W_2 by the Riemannian distance $d_{\mathfrak{g}}$, and use the already known equivalence between Dir and the limit of Dir_ε when $\varepsilon \to 0$ for mappings valued in a Riemannian manifold [17, Theorem 8.3.1].

Notice that the metric tensor $\mathfrak{g}_A$ diverges as A becomes singular. Thus, it is natural to assume that the boundary values have non singular covariance matrices. With this assumption we are able to provide the full solution of the Dirichlet problem.

Theorem 7.10. *Take $\tilde{D} \subset D$ compatible with ρ. Let $\boldsymbol{\mu}_l : \partial\Omega \to \mathcal{P}_{\mathrm{ec}}(\tilde{D})$ a Lipschitz mapping such that $\det(\mathrm{cov}(\boldsymbol{\mu}_l(\xi))) > 0$ for all $\xi \in \partial\Omega$ and define $\mathbf{A}_l(\xi) = \mathrm{cov}(\boldsymbol{\mu}_l(\xi))^{1/2}$ for all $\xi \in \partial\Omega$.*

Then there exists a unique solution $\bar{\boldsymbol{\mu}} \in H^1(\Omega, \mathcal{P}(D))$ of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$ and $\bar{\boldsymbol{\mu}}(\xi) \in \mathcal{P}_{\mathrm{ec}}(D)$ for a.e. $\xi \in \Omega$. Moreover, if $\bar{\mathbf{A}} \in H^1(\Omega, (S_q^{++}(\mathbf{R}), \mathfrak{g}))$ is defined by $\bar{\mathbf{A}}(\xi) := \mathrm{cov}(\bar{\boldsymbol{\mu}}(\xi))^{1/2}$ for a.e. $\xi \in \Omega$, then the following holds:

(i) $\operatorname{ess\,inf}_{\xi\in\Omega} \det(\bar{\mathbf{A}}(\xi)) > 0$;

(ii) $\bar{\mathbf{A}}$ *is a minimizer of*

$$\int_\Omega \frac{1}{2}\sum_{\alpha=1}^p \langle \partial_\alpha \mathbf{B}(\xi), \mathfrak{g}_{\mathbf{B}(\xi)}(\partial_\alpha \mathbf{B}(\xi))\rangle \mathrm{d}\xi.$$

among all $\mathbf{B} \in H^1(\Omega, (S_q^{++}(\mathbf{R}), \mathfrak{g}))$ which have boundary values $\mathbf{A}_l$;

(iii) $\bar{\mathbf{A}}$ *is a weak solution of*

$$\sum_{\alpha=1}^p \partial_\alpha\left(L_{\bar{\mathbf{A}}} L_{\bar{\mathbf{A}}^2}^{-1}(\partial_\alpha \bar{\mathbf{A}})\right) + \sum_{\alpha=1}^p \left(L_{\bar{\mathbf{A}}} L_{\bar{\mathbf{A}}^2}^{-1}(\partial_\alpha \bar{\mathbf{A}})\right)^2 = 0. \tag{7.5}$$

Notice that we are able to prove uniqueness among *all* mappings valued in the Wasserstein space and not only those valued in the family of elliptically contoured distributions: it is one of the only case where we can prove that uniqueness holds for the Dirichlet problem. Remark also that (7.5) is nothing else than the Euler-Lagrange equation associated to the problem of calculus of variations (ii). We will not discuss the question of the regularity of $\bar{\mathbf{A}}$. As $\bar{\mathbf{A}}$ is minimizing a Dirichlet energy its regularity is expected to be better than $H^1(\Omega, (S_q^{++}(\mathbf{R}), \mathfrak{g}))$. On the other hand, as $(S_q^{++}(\mathbf{R}), \mathfrak{g})$ is positively curved we cannot say automatically that $\bar{A}$ cannot blow up somewhere (except if Ω is

of dimension 2 where $\bar{\mathbf{A}}$ will be smooth, see [14, Section 4.2]). A more detailed analysis could the topic for an other study.

The rest of this subsection (and, incidentally, this article) is dedicated to the proof of Theorem 7.10 which is obtained by putting together Propositions 7.11, 7.12 and 7.15. More precisely, the first step is to show the existence of one solution $\bar{\boldsymbol{\mu}}$ of the Dirichlet problem taking values in the family of elliptically contoured distributions for which the covariance matrices stay non singular inside Ω (Proposition 7.11). Then, using the explicit expression (7.4), it is fairly easy to show that (ii) and (iii) are satisfied (Proposition 7.12). The hardest part is the question of uniqueness. As explained in the introduction, we will first show that any solution $\boldsymbol{\mu}$ of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$ must have $\bar{\mathbf{v}}$ as tangent velocity field, where $\bar{\mathbf{v}}$ is the tangent velocity field of $\bar{\boldsymbol{\mu}}$. Then, as $\bar{\mathbf{v}}$ will happen to be smooth enough (linear, hence Lipschitz w.r.t. D), we will use the results about uniqueness of the (1-dimensional) continuity equation for smooth velocity field (Proposition 7.15).

Let us begin by showing that for at least one solution of the Dirichlet problem the covariance matrices stay non singular inside Ω. As a tool to measure regularity of elliptically contoured distributions, we will use the Boltzmann entropy. We define $H : \mathcal{P}(D) \to [0, +\infty]$ by

$$H(\mu) := \begin{cases} \displaystyle\int_D \mu(x) \ln(\mu(x)) \mathrm{d}x & \text{if } \mu \text{ is absolutely continuous w.r.t. } \mathcal{L}_D, \\ +\infty & \text{else.} \end{cases} \tag{7.6}$$

It is known that H is convex along generalized geodesics [4, Theorem 9.4.10] and it is regular according to Proposition 6.2. Moreover, an explicit computation leads to $H(\rho_A) = -\ln(\det A) + H(\rho \mathcal{L}_D)$ (with the convention $\ln(0) = -\infty$). Also, using the fact that Gaussian measures are the ones which minimize H for a covariance matrix, we get that for any $\mu \in \mathcal{P}(D)$,

$$H(\mu) \geqslant -\frac{1}{2} \ln\left(\det\left(\mathrm{cov}(\mu)\right)\right) + C, \tag{7.7}$$

where the constant C is the entropy of a standard normal distribution.

Proposition 7.11. *Take $\tilde{D} \subset D$ compatible with ρ. Let $\boldsymbol{\mu}_l : \partial\Omega \to \mathcal{P}_{\mathrm{ec}}(\tilde{D})$ a Lipschitz mapping such that $\det\left(\mathrm{cov}(\boldsymbol{\mu}_l(\xi))\right) > 0$ for all $\xi \in \partial\Omega$. Then there exists $\bar{\boldsymbol{\mu}} \in H^1(\Omega, \mathcal{P}(D))$ a solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$ such that $\bar{\boldsymbol{\mu}}(\xi) \in \mathcal{P}_{\mathrm{ec}}(D)$ for a.e. $\xi \in \Omega$ and such that*

$$\operatorname*{ess\,inf}_{\xi \in \Omega} \left[\det\left(\mathrm{cov}(\bar{\boldsymbol{\mu}}(\xi))\right)\right] > 0.$$

Proof. Notice, thanks to the explicit formula for H on $\mathcal{P}_{\mathrm{ec}}(\mathbf{R}^q)$ and as $\boldsymbol{\mu}_l$ is continuous, that $\sup_{\partial\Omega}(H \circ \boldsymbol{\mu}_l) < +\infty$. Take $\boldsymbol{\mu} \in H^1(\Omega, \mathcal{P}(\tilde{D}))$ the solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$ given by Theorem 6.3 (with $F = H$). Set $\bar{\boldsymbol{\mu}} := R \circ \boldsymbol{\mu}$. By the same argument as in Theorem 7.9, $\bar{\boldsymbol{\mu}} \in H^1(\Omega, \mathcal{P}_{\mathrm{ec}}(D))$ is a solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$. Using first the estimate (7.7) and then the maximum principle (6.1),

$$\begin{aligned} \operatorname*{ess\,sup}_{\xi \in \Omega} \left[-\ln\left(\det\left(\mathrm{cov}(\bar{\boldsymbol{\mu}}(\xi))\right)\right)\right] &= \operatorname*{ess\,sup}_{\xi \in \Omega} \left[-\ln\left(\det\left(\mathrm{cov}(\boldsymbol{\mu}(\xi))\right)\right)\right] \\ &\leqslant -2C + 2 \operatorname*{ess\,sup}_{\xi \in \Omega} H(\boldsymbol{\mu}(\xi)) \\ &\leqslant -2C + 2 \sup_{\xi \in \partial\Omega} H(\boldsymbol{\mu}_l(\xi)) < +\infty. \end{aligned}$$

□

Until the end of the subsection, $\bar{\boldsymbol{\mu}} \in H^1(\Omega, \mathcal{P}_{\mathrm{ec}}(D))$ will denote the object defined in Proposition 7.11 and for a.e. $\xi \in \Omega$, one defines $\bar{\mathbf{A}}(\xi) = \mathrm{cov}(\bar{\boldsymbol{\mu}}(\xi))^{1/2}$. Notice that (i) of Theorem 7.10 is proved. Now let us derive the equation satisfied by $\bar{\mathbf{A}}$.

Proposition 7.12. *The mapping $\bar{\mathbf{A}} \in H^1(\Omega, (S_q^{++}(\mathbf{R}), \mathfrak{g}))$ is a weakly harmonic map, more precisely a minimizer of*

$$\mathbf{B} \in H^1(\Omega, (S_q^{++}(\mathbf{R}), \mathfrak{g})) \mapsto \int_\Omega \frac{1}{2} \sum_{\alpha=1}^p \langle \partial_\alpha \mathbf{B}(\xi), \mathfrak{g}_{\mathbf{B}(\xi)}(\partial_\alpha \mathbf{B}(\xi)) \rangle \mathrm{d}\xi.$$

among all $\mathbf{B}$ which have boundary values $\mathbf{A}_l$. In particular, $\bar{\mathbf{A}}$ satisfies the Euler-Lagrange equation (7.5)*.*

Proof. We need to prove that, for any $\mathbf{B} \in H^1(\Omega, (S_q^{++}(\mathbf{R}), \mathfrak{g}))$ with boundary values $\mathbf{A}_l$ one has

$$\int_\Omega \frac{1}{2} \sum_{\alpha=1}^p \langle \partial_\alpha \mathbf{B}(\xi), \mathfrak{g}_{\mathbf{B}(\xi)}(\partial_\alpha \mathbf{B}(\xi)) \rangle \mathrm{d}\xi \geqslant \int_\Omega \frac{1}{2} \sum_{\alpha=1}^p \langle \partial_\alpha \bar{\mathbf{A}}(\xi), \mathfrak{g}_{\bar{\mathbf{A}}(\xi)}(\partial_\alpha \bar{\mathbf{A}}(\xi)) \rangle \mathrm{d}\xi = \mathrm{Dir}(\rho_{\bar{\mathbf{A}}}) = \mathrm{Dir}(\bar{\boldsymbol{\mu}}).$$

To prove it, if we take any $\mathbf{B} \in H^1(\Omega, (S_q^{++}(\mathbf{R}), \mathfrak{g}))$ we can build $\boldsymbol{\mu} := \rho_{\mathbf{B}}$ and we have, thanks to (7.4), the identity

$$\mathrm{Dir}(\boldsymbol{\mu}) = \int_\Omega \frac{1}{2} \sum_{\alpha=1}^p \langle \partial_\alpha \mathbf{B}(\xi), \mathfrak{g}_{\mathbf{B}(\xi)}(\partial_\alpha \mathbf{B}(\xi)) \rangle \mathrm{d}\xi.$$

A priori, $\boldsymbol{\mu}$ is valued in $\mathcal{P}(\mathbf{R}^q)$. If we denote by $P_D : \mathbf{R}^q \to D$ the Euclidean projection on D, then

$$\mathrm{Dir}(\bar{\boldsymbol{\mu}}) \leqslant \mathrm{Dir}(P_D \# \boldsymbol{\mu}) \leqslant \mathrm{Dir}(\boldsymbol{\mu}),$$

where the first inequality comes from the optimality of $\bar{\boldsymbol{\mu}}$ (notice that $P_D \# \cdot$ leaves the boundary values unchanged) and the second one from the fact that $P_D \# \cdot$ is a contraction (Lemma 7.7).

To get the Euler-Lagrange equation it is actually easier if we take the covariance matrix and not its square root as the variable. In other words we define $\bar{\mathbf{C}} := \bar{\mathbf{A}}^2$. As $\bar{\mathbf{A}}$ is never singular, this change of variables is smooth. We have $\partial_\alpha \bar{\mathbf{C}} = L_{\bar{\mathbf{A}}}(\partial_\alpha \bar{\mathbf{A}})$ and in particular

$$\langle \partial_\alpha \bar{\mathbf{A}}, \mathfrak{g}_{\bar{\mathbf{A}}}(\partial_\alpha \bar{\mathbf{A}}) \rangle = \langle \partial_\alpha \bar{\mathbf{C}}, L_{\bar{\mathbf{C}}}^{-1}(\partial_\alpha \bar{\mathbf{C}}) \rangle.$$

If we take $\mathbf{D} : \Omega \to S_q(\mathbf{R})$ smooth and compactly supported on Ω and that we consider $\mathbf{B} := \bar{\mathbf{C}} + t\mathbf{D}$ as a competitor for small t, we reach the conclusion that

$$\sum_{\alpha=1}^p \langle \partial_\alpha \mathbf{D}, L_{\bar{\mathbf{C}}}^{-1}(\partial_\alpha \bar{\mathbf{C}}) \rangle + \frac{1}{2} \sum_{\alpha=1}^p \left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t=0} \langle \partial_\alpha \bar{\mathbf{C}}, L_{\bar{\mathbf{C}}+t\mathbf{D}}^{-1}(\partial_\alpha \bar{\mathbf{C}}) \rangle = 0.$$

A simple computation leads to

$$L_{\bar{\mathbf{C}}+t\mathbf{D}}^{-1}(\partial_\alpha \bar{\mathbf{C}}) = L_{\bar{\mathbf{C}}}^{-1}(\partial_\alpha \bar{\mathbf{C}}) - t L_{\bar{\mathbf{C}}}^{-1} \left[\mathbf{D}(L_{\bar{\mathbf{C}}}^{-1}(\partial_\alpha \bar{\mathbf{C}})) + (L_{\bar{\mathbf{C}}}^{-1}(\partial_\alpha \bar{\mathbf{C}}))\mathbf{D} \right] + o(t^2).$$

Using the properties of the Trace and the symmetry of $L_{\bar{\mathbf{C}}}^{-1}$, we conclude that the Euler-Lagrange equation reads

$$\sum_{\alpha=1}^p \langle \partial_\alpha \mathbf{D}, L_{\bar{\mathbf{C}}}^{-1}(\partial_\alpha \bar{\mathbf{C}}) \rangle - \sum_{\alpha=1}^p \langle \mathbf{D}, (L_{\bar{\mathbf{C}}}^{-1}(\partial_\alpha \bar{\mathbf{C}}))^2 \rangle = 0.$$

Coming back to $\bar{\mathbf{C}} = \bar{\mathbf{A}}^2$ and $\partial_\alpha \bar{\mathbf{C}} = L_{\bar{\mathbf{A}}}(\partial_\alpha \bar{\mathbf{A}})$, as $\mathbf{D}$ is arbitrary we see that we get the weak formulation of (7.5). □

As far as the regularity issues are concerned, notice that $\bar{\mathbf{A}}$ is uniformly bounded from below as a symmetric matrix (this is (i) of Theorem 7.10) and also bounded from above as a symmetric matrix (as $\rho_{\bar{\mathbf{A}}} \in \mathcal{P}(D)$ and D is compact), hence the operators $L_{\bar{\mathbf{A}}(\xi)} : S_q(\mathbf{R}) \to S_q(\mathbf{R})$ are bounded with a bounded inverse uniformly in $\xi \in \Omega$. In other words, the metric tensor $\mathfrak{g}_{\bar{\mathbf{A}}(\xi)}$ is equivalent to the canonical scalar product uniformly in $\xi \in \Omega$. In particular, the regularity $\bar{\boldsymbol{\mu}} \in H^1(\Omega, \mathcal{P}(D))$ translates in $\bar{\mathbf{A}} \in H^1(\Omega, S_q(\mathbf{R}))$ where $S_q(\mathbf{R})$ is endowed with its usual euclidean norm $|\cdot|$.

To conclude it remains to prove uniqueness. The first step is to identify the tangent velocity field to $\bar{\boldsymbol{\mu}}$ and a (at least formal) solution of the dual problem.

Proposition 7.13. *For any $\alpha \in \{1, 2, \dots, p\}$ we define $\bar{\mathbf{B}}^\alpha := L_{\bar{\mathbf{A}}} L^{-1}_{\bar{\mathbf{A}}^2}(\partial_\alpha \bar{\mathbf{A}}) \in L^2(\Omega, S_q(\mathbf{R}))$ and we set*

$$\bar{\mathbf{v}}^\alpha(\xi, x) := \bar{\mathbf{B}}^\alpha(\xi)x \in \mathbf{R}^q.$$

for $\xi \in \Omega$ and $x \in D$. Then $\bar{\mathbf{v}} \in L^2_{\bar{\boldsymbol{\mu}}}(\Omega \times D, \mathbf{R}^{pq})$ is the tangent velocity field to $\bar{\boldsymbol{\mu}}$.

Proof. Take $\psi \in C^1_c(\Omega \times D, \mathbf{R}^p)$ a test function. If we define $\tilde{\psi} \in H^1(\Omega, \mathbf{R}^p)$ by

$$\tilde{\psi}(\xi) := \int_D \psi(\xi, x)\bar{\boldsymbol{\mu}}(\xi, \mathrm{d}x) = \int_D \psi(\xi, \bar{\mathbf{A}}(\xi)x)\rho(x)\mathrm{d}x,$$

then we see that $\tilde{\psi}$ is compactly supported in Ω, in particular the integral of $\nabla \cdot \tilde{\psi}$ over Ω vanishes. It reads

$$\int_{\Omega \times D} (\nabla_\Omega \cdot \psi)(\xi, \bar{\mathbf{A}}(\xi)x)\rho(x)\mathrm{d}x + \int_{\Omega \times D} \sum_{\alpha=1}^{p} (\partial_\alpha \bar{\mathbf{A}}(\xi)x) \cdot (\nabla_D \psi^\alpha)(\xi, \bar{\mathbf{A}}(\xi)x)\rho(x)\mathrm{d}x = 0.$$

By doing for a fixed $\xi \in \Omega$ the change of variables $y = \bar{\mathbf{A}}(\xi)x$, one can see that $(\bar{\boldsymbol{\mu}}, \mathbf{w}\bar{\boldsymbol{\mu}})$ satisfies the continuity equation where $\mathbf{w} : \Omega \times D \to \mathbf{R}^p$ is given by

$$\mathbf{w}^\alpha(\xi, y) := \partial_\alpha \bar{\mathbf{A}}(\xi)\bar{\mathbf{A}}(\xi)^{-1}y.$$

Notice that $\mathbf{w}(\xi, \cdot)$ is not a gradient because $\partial_\alpha \bar{\mathbf{A}}(\xi)$ and $\bar{\mathbf{A}}(\xi)^{-1}$ do not necessarily commute. On the contrary, as the matrices $\bar{\mathbf{B}}^\alpha(\xi)$ for $\alpha \in \{1, 2, \dots, p\}$ are symmetric, $\bar{\mathbf{v}}(\xi, \cdot)$ is a gradient.

Fix $\xi \in \Omega$ and $\alpha \in \{1, 2, \dots, p\}$. We claim that the velocity field $\bar{\mathbf{v}}^\alpha(\xi, \cdot)$ is the orthogonal projection in $L^2_{\bar{\boldsymbol{\mu}}(\xi)}(D, \mathbf{R}^q)$ of $\mathbf{w}^\alpha(\xi, \cdot)$ on the space of gradients (actually, this is exactly how $\bar{\mathbf{v}}^\alpha$ was chosen). Not to overburden the notations, we drop momentarily the dependence on ξ, i.e. $\bar{\mathbf{A}} := \bar{\mathbf{A}}(\xi)$, $\bar{\mathbf{B}}^\alpha := \bar{\mathbf{B}}^\alpha(\xi)$ and $\partial_\alpha \bar{\mathbf{A}} := \partial_\alpha \bar{\mathbf{A}}(\xi)$ are considered as given matrices. Take $f \in C^1(D)$ a test function defined on D and compute:

$$\begin{aligned}\int_D \nabla f(x) \cdot (\mathbf{w}^\alpha(\xi, x) - \bar{\mathbf{v}}^\alpha(\xi, x))\bar{\boldsymbol{\mu}}(\xi, \mathrm{d}x) &= \int_D (\nabla f)(\bar{\mathbf{A}}x) \cdot \left((\partial_\alpha \bar{\mathbf{A}}\bar{\mathbf{A}}^{-1} - \bar{\mathbf{B}}^\alpha)\bar{\mathbf{A}}x\right)\rho(x)\mathrm{d}x \\ &= \int_D (\nabla \tilde{f})(x) \cdot \left(\bar{\mathbf{A}}^{-1}(\partial_\alpha \bar{\mathbf{A}}\bar{\mathbf{A}}^{-1} - \bar{\mathbf{B}}^\alpha)\bar{\mathbf{A}}x\right)\rho(x)\mathrm{d}x,\end{aligned}$$

where $\tilde{f}(x) := f(\bar{\mathbf{A}}x)$. On the other hand, as the reader can check, $\bar{\mathbf{B}}^\alpha$ is the projection on the set of symmetric matrices of $\partial_\alpha \bar{\mathbf{A}}\bar{\mathbf{A}}^{-1}$ where the scalar product between two matrices C and D is given by $\mathrm{Tr}(\bar{\mathbf{A}}^t C D \bar{\mathbf{A}})$. In particular, the matrix $(\partial_\alpha \bar{\mathbf{A}}\bar{\mathbf{A}}^{-1} - \bar{\mathbf{B}}^\alpha)\bar{\mathbf{A}}^2$ is skew-symmetric, thus the matrix $\bar{\mathbf{A}}^{-1}(\partial_\alpha \bar{\mathbf{A}}\bar{\mathbf{A}}^{-1} - \bar{\mathbf{B}}^\alpha)\bar{\mathbf{A}}$ is also skew-symmetric. As ρ is radial, it implies that the function

$$x \in D \mapsto \left(\bar{\mathbf{A}}^{-1}(\partial_\alpha \bar{\mathbf{A}}\bar{\mathbf{A}}^{-1} - \bar{\mathbf{B}}^\alpha)\bar{\mathbf{A}}x\right)\rho(x)$$

is divergence-free. It allows us to conclude that

$$\int_D \nabla f(x) \cdot (\mathbf{w}^\alpha(\xi, x) - \bar{\mathbf{v}}^\alpha(\xi, x))\bar{\boldsymbol{\mu}}(\xi, \mathrm{d}x) = \int_D \tilde{f}(x)\nabla_D \cdot \left[\left(\bar{\mathbf{A}}^{-1}(\partial_\alpha \bar{\mathbf{A}}\bar{\mathbf{A}}^{-1} - \bar{\mathbf{B}}^\alpha)\bar{\mathbf{A}}x\right)\rho(x)\right]\mathrm{d}x = 0,$$

hence the claim is proved as f is arbitrary.

The claim implies that $(\bar{\boldsymbol{\mu}}, \bar{\mathbf{v}}\bar{\boldsymbol{\mu}})$ also satisfies the continuity equation: for any $\psi \in C^1_c(\Omega \times D, \mathbf{R}^p)$,

$$\int_{\Omega \times D} \nabla_\Omega \cdot \psi \mathrm{d}\bar{\boldsymbol{\mu}} + \int_{\Omega \times D} \nabla_D \psi \cdot \bar{\mathbf{v}} \mathrm{d}\bar{\boldsymbol{\mu}} = \int_{\Omega \times D} \nabla_\Omega \cdot \psi \mathrm{d}\bar{\boldsymbol{\mu}} + \int_{\Omega \times D} \nabla_D \psi \cdot \mathbf{w} \mathrm{d}\bar{\boldsymbol{\mu}} + \int_{\Omega \times D} \nabla_D \psi \cdot (\bar{\mathbf{v}} - \mathbf{w})\mathrm{d}\bar{\boldsymbol{\mu}} = 0,$$

as the last integral vanishes because of the claim.

As $\bar{\mathbf{v}}(\xi, \cdot)$ is a gradient (because the $\bar{\mathbf{B}}^\alpha$ are symmetric), Corollary 3.21 implies that $\bar{\mathbf{v}}$ is the tangent velocity field to $\bar{\boldsymbol{\mu}}$. □

Notice that if we define $\bar\varphi : \Omega \times D \to \mathbf{R}^p$ by, for any $\xi \in \Omega, x \in D$ and $\alpha \in \{1, 2, \dots, p\}$,

$$\bar\varphi^\alpha(\xi, x) := \frac{1}{2}\bar{\mathbf{B}}^\alpha(\xi)x \cdot x;$$

then $\bar{\mathbf{v}} = \nabla_D\varphi$. More precisely, for a.e. $\xi \in \Omega$, $\bar\varphi(\xi, \cdot)$ (resp. $\bar{\mathbf{v}}(\xi, \cdot)$) is defined *everywhere* on D as a smooth function belonging to $C^1(D, \mathbf{R}^p)$ (resp. $C^1(D, \mathbf{R}^{pq})$). Moreover the Euler-Lagrange equation (7.5), which can be written

$$\sum_{\alpha=1}^p \partial_\alpha\bar{\mathbf{B}}^\alpha + \sum_{\alpha=1}^p (\bar{\mathbf{B}}^\alpha)^2 = 0, \tag{7.8}$$

translates at the level of $\bar\varphi$ in

$$\nabla_\Omega \cdot \bar\varphi + \frac{1}{2}|\nabla_D\bar\varphi|^2 = 0.$$

Hence, at least formally (because of the lack of smoothness of $\bar\varphi$), the function $\bar\varphi$ is a solution of the dual problem. We will use $\bar\varphi$ to show that the tangent velocity field of any other solution of the Dirichlet problem with boundary values $\boldsymbol{\mu}_l$ must coincide with $\bar{\mathbf{v}}$. About the smoothness of the objects involved, notice that for any $\alpha \in \{1, 2, \dots, p\}$ one has $\bar{\mathbf{B}}^\alpha \in L^2(\Omega, S_q(\mathbf{R}))$ and, given (7.8), the function

$$\sum_{\alpha=1}^p \partial_\alpha\bar{\mathbf{B}}^\alpha$$

belongs to $L^1(\Omega, S_q(\mathbf{R}))$.

Proposition 7.14. *Let $\boldsymbol{\mu}$ a solution of the Dirichlet problem with boundary conditions $\boldsymbol{\mu}_l$ and $\mathbf{v}$ its tangent velocity field. Then, for a.e. $\xi \in \Omega$, one has $\mathbf{v}(\xi, x) = \bar{\mathbf{v}}(\xi, x)$ for $\boldsymbol{\mu}(\xi)$-a.e. x.*

Proof. If $\varphi \in C^1(\Omega \times D, \mathbf{R}^p)$ then, as $\boldsymbol{\mu}$ and $\bar{\boldsymbol{\mu}}$ share the same boundary conditions,

$$\int_{\Omega\times D} (\nabla_\Omega \cdot \varphi + \nabla_D\varphi \cdot \mathbf{v})\mathrm{d}\boldsymbol{\mu} = \mathrm{BT}_{\boldsymbol{\mu}_l}(\varphi) = \int_{\Omega\times D} (\nabla_\Omega \cdot \varphi + \nabla_D\varphi \cdot \bar{\mathbf{v}})\mathrm{d}\bar{\boldsymbol{\mu}}.$$

We claim that we can insert $\varphi = \bar\varphi$ even though $\bar\varphi$ is *a priori* not regular enough. In other words, given (7.2) and the fact that $\bar{\mathbf{v}} = \nabla_D\bar\varphi$, we claim that

$$\int_{\Omega\times D} \left(-\frac{1}{2}|\bar{\mathbf{v}}|^2 + \bar{\mathbf{v}} \cdot \mathbf{v}\right) \mathrm{d}\boldsymbol{\mu} = \int_{\Omega\times D} \frac{1}{2}|\bar{\mathbf{v}}|^2\mathrm{d}\bar{\boldsymbol{\mu}}. \tag{7.9}$$

Notice that the r.h.s. is (formally) equal to both $\mathrm{BT}_{\boldsymbol{\mu}_l}(\bar\varphi)$ and $\mathrm{Dir}(\bar{\boldsymbol{\mu}})$: it is not surprising as $\bar\varphi$ is a solution of the dual problem.

To prove such an equality we regularize $\bar\varphi$ in the following way. For each $\alpha \in \{1, 2, \dots, p\}$ we apply to the matrix field $\bar{\mathbf{B}}^\alpha$ the standard truncation and convolution procedure (see [12, Theorem 3 of Section 4.2]) to produce a sequence $(\bar{\mathbf{B}}^\alpha_n)_{n\in\mathbf{N}}$ which belongs to $C^1(\Omega, S_q(\mathbf{R}))$ and which converges to $\bar{\mathbf{B}}^\alpha$ in $L^2(\Omega, S_q(\mathbf{R}))$. Moreover, as derivatives commute with convolution, we can say that

$$\lim_{n\to+\infty} \sum_{\alpha=1}^p \partial_\alpha\bar{\mathbf{B}}^\alpha_n = \sum_{\alpha=1}^p \partial_\alpha\bar{\mathbf{B}}^\alpha = -\sum_{\alpha=1}^p (\bar{\mathbf{B}}^\alpha)^2,$$

and the limit takes place in $L^1(\Omega, S_q(\mathbf{R}))$ as we already know that the r.h.s. belongs to such a space. In particular, up to extraction the convergences hold a.e. on Ω. Then we set

$$\varphi^\alpha_n(\xi, x) := \frac{1}{2}\bar{\mathbf{B}}^\alpha_n(\xi)x \cdot x.$$

for $\xi \in \Omega$ and $x \in D$. By construction $\varphi_n \in C^1(\Omega \times D, \mathbf{R})$ so that

$$\int_{\Omega\times D} (\nabla_\Omega \cdot \varphi_n + \nabla_D \varphi_n \cdot \mathbf{v}) \mathrm{d}\boldsymbol{\mu} = \mathrm{BT}_{\boldsymbol{\mu}_l}(\varphi_n) = \int_{\Omega\times D} (\nabla_\Omega \cdot \varphi_n + \nabla_D \varphi_n \cdot \bar{\mathbf{v}}) \mathrm{d}\bar{\boldsymbol{\mu}}. \tag{7.10}$$

It remains to show that we can pass to the limit $n \to +\infty$. Given the convergence a.e. of the $\bar{\mathbf{B}}_n^\alpha$ and of $\sum \partial_\alpha \bar{\mathbf{B}}_n^\alpha$, we can assume that for a.e. $\xi \in \Omega$, the functions $\nabla_\Omega \cdot \varphi_n(\xi, \cdot)$ and $\nabla_D \varphi_n(\xi, \cdot)$ converge to respectively $-\frac{1}{2}|\bar{\mathbf{v}}|^2(\xi, \cdot)$ and $\bar{\mathbf{v}}(\xi, \cdot)$ in respectively $C(D)$ and $C(D, \mathbf{R}^{pq})$ respectively (notice that we use the fact that D is bounded). Hence for a.e. $\xi \in \Omega$,

$$\lim_{n\to+\infty} \int_D (\nabla_\Omega \cdot \varphi_n(\xi, x) + \nabla_D \varphi_n(\xi, x) \cdot \mathbf{v}(\xi, x)) \, \boldsymbol{\mu}(\xi, \mathrm{d}x) = \int_D \left(-\frac{1}{2}|\bar{\mathbf{v}}|^2(\xi, x) + \bar{\mathbf{v}}(\xi, x) \cdot \mathbf{v}(\xi, x) \right) \boldsymbol{\mu}(\xi, \mathrm{d}x). \tag{7.11}$$

It remains to integrate this limit over Ω. The natural upper bound for the l.h.s. of (7.11) is obtained by Cauchy-Schwarz and the boundedness of D: for any $n \in \mathbf{N}$,

$$\left| \int_D (\nabla_\Omega \cdot \varphi_n(\xi, x) + \nabla_D \varphi_n(\xi, x) \cdot \mathbf{v}(\xi, x)) \boldsymbol{\mu}(\xi, \mathrm{d}x) \right| \leqslant C \left(\sum_{\alpha=1}^p |\bar{\mathbf{B}}_n^\alpha(\xi)|^2 + \sqrt{\int_D |\mathbf{v}(\xi, x)|^2 \boldsymbol{\mu}(\xi, \mathrm{d}x)} \sqrt{\sum_{\alpha=1}^p |\bar{\mathbf{B}}_n^\alpha(\xi)|^2} \right),$$

where C depends only on D. The r.h.s. is not bounded uniformly w.r.t. $n \in \mathbf{N}$ but on the other hand it converges in $L^1(\Omega)$ which is enough to say that the l.h.s. is uniformly integrable. Hence, up to extraction we can integrate (7.11) w.r.t. Ω:

$$\lim_{n\to+\infty} \int_{\Omega\times D} (\nabla_\Omega \cdot \varphi_n + \nabla_D \varphi_n \cdot \mathbf{v}) \mathrm{d}\boldsymbol{\mu} = \int_{\Omega\times D} \left(-\frac{1}{2}|\bar{\mathbf{v}}|^2 + \bar{\mathbf{v}} \cdot \mathbf{v} \right) \mathrm{d}\boldsymbol{\mu}.$$

Of course, the result still holds if we take $(\boldsymbol{\mu}, \mathbf{v}) = (\bar{\boldsymbol{\mu}}, \bar{\mathbf{v}})$. Thus, passing in the limit in (7.10) we get (7.9).

Until now we have not used the optimality of $\boldsymbol{\mu}$. We notice that the r.h.s. of (7.9) is nothing else than $\mathrm{Dir}(\bar{\boldsymbol{\mu}})$ which coincides with $\mathrm{Dir}(\boldsymbol{\mu}) = \int_{\Omega\times D} \frac{1}{2}|\mathbf{v}|^2 \mathrm{d}\boldsymbol{\mu}$ by optimality of $\boldsymbol{\mu}$. From there, an algebraic manipulation leads to

$$\int_{\Omega\times D} \frac{1}{2}|\mathbf{v} - \bar{\mathbf{v}}|^2 \mathrm{d}\boldsymbol{\mu} = 0,$$

which easily implies the result: recall that for a.e. $\xi \in \Omega$, the velocity field $\bar{\mathbf{v}}$ is continuous on D. □

Proposition 7.15. *Let $\boldsymbol{\mu}$ a solution of the Dirichlet problem with boundary conditions $\boldsymbol{\mu}_l$. Then $\boldsymbol{\mu} = \bar{\boldsymbol{\mu}}$.*

Proof. Take $\boldsymbol{\mu}$ a solution of the Dirichlet problem with boundary conditions $\boldsymbol{\mu}_l$ and define $\boldsymbol{\nu} = \boldsymbol{\mu} - \bar{\boldsymbol{\mu}}$. We extend $\boldsymbol{\nu}$ on $\mathbf{R}^p \backslash \Omega$ by the 0: with such a choice $\boldsymbol{\nu} \in L^2(\mathbf{R}^p, \mathcal{M}(D))$ is a (signed) measure-valued mapping defined on the whole space $\mathbf{R}^p$ which vanishes outside a compact set. We also define $\bar{\mathbf{v}}$ as a function $\mathbf{R}^p \times D \to \mathbf{R}^{pq}$ by extending it to 0 outside $\Omega \times D$. If $\varphi \in C^1(\mathbf{R}^p \times D, \mathbf{R}^p)$ is any smooth function then

$$\begin{aligned} \int_{\mathbf{R}^p\times D} (\nabla_\Omega \cdot \varphi + \nabla_D \varphi \cdot \bar{\mathbf{v}}) \, \mathrm{d}\boldsymbol{\nu} &= \int_{\Omega\times D} (\nabla_\Omega \cdot \varphi + \nabla_D \varphi \cdot \bar{\mathbf{v}}) \, \mathrm{d}\boldsymbol{\nu} \\ &= \int_{\Omega\times D} (\nabla_\Omega \cdot \varphi + \nabla_D \varphi \cdot \bar{\mathbf{v}}) \, \mathrm{d}\boldsymbol{\mu} - \int_{\Omega\times D} (\nabla_\Omega \cdot \varphi + \nabla_D \varphi \cdot \bar{\mathbf{v}}) \, \mathrm{d}\bar{\boldsymbol{\mu}} \\ &= \mathrm{BT}_{\boldsymbol{\mu}_l}(\varphi) - \mathrm{BT}_{\boldsymbol{\mu}_l}(\varphi) = 0, \end{aligned}$$

where we have used the fact that both $(\boldsymbol{\mu}, \bar{\mathbf{v}}\boldsymbol{\mu})$ and $(\bar{\boldsymbol{\mu}}, \bar{\mathbf{v}}\bar{\boldsymbol{\mu}})$ satisfy the continuity equation. In other words, $(\boldsymbol{\nu}, \bar{\mathbf{v}}\boldsymbol{\nu})$ satisfy the continuity equation on the whole space $\mathbf{R}^p \times D$.

We take an arbitrary direction in $\mathbf{R}^p$: we fix $\alpha = 1$. As we have seen in the proof of Proposition 3.20, the (generalized) continuity equation implies that for a.e. $\xi \in \mathbf{R}^{p-1} = (e_\alpha)^\perp$, the curve $t \in \mathbf{R} \mapsto \boldsymbol{\nu}((t,\xi))$ satisfies the (1-dimensional) continuity equation with a velocity field given by $\mathbf{w}(t,x) = \bar{\mathbf{v}}^\alpha((t,\xi),x)$. Notice that for a fixed t the velocity field $\mathbf{w}(t,\cdot)$ is Lipschitz and bounded with Lipschitz constant and upper bound controlled by $C\mathbb{1}_{(t,\xi)\in\Omega}|\bar{\mathbf{B}}^\alpha((t,\xi))|$ where $C < +\infty$ depends only on D. Given that $\bar{\mathbf{B}}^\alpha \in L^2(\Omega)$, for a.e. $\xi \in \mathbf{R}^{p-1}$ one has that

$$\int_{\mathbf{R}} \mathbb{1}_{(t,\xi)\in\Omega}|\bar{\mathbf{B}}^\alpha((t,\xi))|\mathrm{d}t < +\infty.$$

Hence for a.e. $\xi \in \mathbf{R}^{p-1}$ the assumptions of [4, Proposition 8.1.7] are satisfied: the curve $t \in \mathbf{R} \mapsto \boldsymbol{\nu}((t,\xi))$ is solution of a continuity equation which has at most one solution. As the curve identically equal to 0 is a solution (recall that $\boldsymbol{\nu}((t,\xi)) = 0$ for $|t|$ large enough), so must be $\boldsymbol{\nu}((\cdot,\xi))$. As this result holds for a.e. $\xi \in \mathbf{R}^{p-1}$, it implies that $\boldsymbol{\nu}$ is identically zero, which is the desired result. □

Appendix A. Measurable selection of the $\arg\min$

We want to show a result which states that if $F : X \times Y \to \mathbf{R}$ is a function which is measurable w.r.t. X, then one can find a selection $m : X \to Y$ such that $F(x, m(x)) = \min_Y F(x,\cdot)$ for every $x \in X$, i.e. such that $m(x) \in \arg\min_Y F(x,\cdot)$. Recall the following result which can be found in [2, Theorem 18.19].

Proposition A.1. *Let X be a measured space and Y a polish space. Let $F : X \times Y \to \mathbf{R}$ a function such that $F(x,\cdot) : Y \to \mathbf{R}$ is continuous for every $x \in X$, and $F(\cdot, y) : X \to \mathbf{R}$ is measurable for every $y \in Y$. Assume that for every $x \in X$, the function $F(x,\cdot)$ has a minimizer over Y.*

Then there exists $m : X \to Y$ a measurable function such that for all $x \in X$,

$$F(x, m(x)) = \min_{y\in Y} F(x,y).$$

However, in particular for Proposition 6.8, we need a case where $F(x,\cdot)$ is only l.s.c.. Thus, we prove some *ad hoc* result relying on the particular structure of our problem which allows to treat lower semi-continuity.

Lemma A.2. *Let X be a measured space and Y a compact metrizable space. Let $F : X \times Y \to \mathbf{R}$ a function such that $F(x,\cdot) : Y \to \mathbf{R}$ is continuous for every $x \in X$, and $F(\cdot, y) : X \to \mathbf{R}$ is measurable for every $y \in Y$; and let $G : Y \to \mathbf{R}$ a l.s.c. function.*

Then the function $H : X \to \mathbf{R}$ defined by

$$H(x) := \min\{F(x,y) + G(y) \ : \ y \in Y\}$$

is measurable.

Proof. Notice that Y is separable as it is compact and metrizable. For any rational number a, the exists a sequence dense in $\{y \in Y \ : \ G(y) \leqslant a\}$. Hence, we can construct a sequence $(y_n)_{n\in\mathbf{N}}$ such that for any rational number a there is a subsequence of $(y_n)_{n\in\mathbf{N}}$ which is included and dense in $\{y \in Y \ : \ G(y) \leqslant a\}$.

Set $\tilde{H}(x) := \inf_n F(x, y_n) + G(y_n)$ which is measurable and larger than H. Let us prove that it is equal to H. Indeed, if $x \in X$, by standard arguments of calculus of variations, there exists $\bar{y}$ such that $H(x) = F(x,\bar{y}) + G(\bar{y})$. For any $a > G(\bar{y})$ rational, take a subsequence $(y_{n_k})_{k\in\mathbf{N}}$ which belongs to $\{y \in Y \ : \ G(y) \leqslant a\}$ and which converges to $\bar{y}$. By continuity of F, one has

$$\tilde{H}(x) \leqslant \liminf_{k\to+\infty} (F(x, y_{n_k}) + G(y_{n_k})) \leqslant F(x,\bar{y}) + a.$$

As a can be chosen arbitrary close to $G(\bar{y})$, we have that $\tilde{H}(x) \leqslant F(x,\bar{y}) + G(\bar{y}) = H(x)$. □

Proposition A.3. *Let X be a measured space and Y a compact metrizable space. Let $F : X \times Y \to \mathbf{R}$ a function such that $F(x,\cdot) : Y \to \mathbf{R}$ is continuous for every $x \in X$, and $F(\cdot,y) : X \to \mathbf{R}$ is measurable for every $y \in Y$; and let $G : Y \to \mathbf{R}$ a l.s.c. function.*

Then there exists $m : X \to Y$ a measurable function such that for any $x \in X$,

$$F(x,m(x)) + G(m(x)) = \min\{F(x,y) + G(y) \ : \ y \in Y\}.$$

Proof. As in the previous lemma, define $H(x) := \min\{F(x,y) + G(y) \ : \ y \in Y\}$, it is a measurable function valued in $\mathbf{R}$. Let Γ be the mapping going from X and valued in the compact subsets of Y defined by $\Gamma(x) = \arg\min_x(F(x,\cdot) + G(\cdot))$ which means

$$\Gamma(x) := \{y \in Y \ : \ F(x,y) + G(y) = H(x)\}.$$

Notice that $\Gamma(x)$ is never empty thanks to standard arguments of calculus of variations. To prove the existence of a measurable selection of Γ, we rely on [2, Theorem 18.13]: it is sufficient to show that Γ is measurable, which means that $\{x \in X \ : \ \Gamma(x) \cap Z \neq \varnothing\}$ is a measurable set of X for any closed set $Z \subset Y$. But one can be convinced that, for a fixed $Z \subset Y$ closed,

$$\Gamma(x) \cap Z \neq \varnothing \ \Leftrightarrow \ H(x) = H_Z(x),$$

where $H_Z(x) := \min\{F(x,z)+G(z) \ : \ z \in Z\}$. Thanks to Lemma A.2, both H and H_Z are measurable, thus the set on which they coincide is measurable, which concludes the proof. □

Appendix B. $H^{1/2}$ determination of the square root

In this appendix we want to prove Lemma 5.4, which states that, with $\mathbf{S}^1$ the unit circle of the complex plane $\mathbf{C}$ and $\mathbf{U}$ its unit disk, there is no function $f \in H^{1/2}(\mathbf{S}^1,\mathbf{S}^1)$ such that $f(\xi)^2 = \xi$ for a.e. $\xi \in \mathbf{S}^1$ (where the multiplication is understood as a complex multiplication). We take for granted that there is no continuous function $f \in C(\mathbf{S}^1,\mathbf{S}^1)$ such that $f(\xi)^2 = \xi$ for all $\xi \in \mathbf{S}^1$. Hence, it is enough to reason by contradiction and to prove that a function $f \in H^{1/2}(\mathbf{S}^1,\mathbf{S}^1)$ such that $f(\xi)^2 = \xi$ for a.e. $\xi \in \mathbf{S}^1$ admits a continuous representative.

We start with some easy lemma which states that $H^{1/2}(\mathbf{S}^1,\mathbf{U})$ is stable by composition with Lipschitz function.

Lemma B.1. *Let $u : \mathbf{S}^1 \to \mathbf{R}$ a Lipschitz function and $f \in H^{1/2}(\mathbf{S}^1,\mathbf{S}^1)$. Then $(u \circ f) \in H^{1/2}(\mathbf{S}^1,\mathbf{R})$.*

Proof. It is well known (see [23, Chapter 3]) that there exists $\tilde{f} \in H^1(\mathbf{U},\mathbf{U})$ whose trace on $\mathbf{S}^1$ is f. Clearly, the function $u \circ \tilde{f}$ stays in $H^1(\mathbf{U},\mathbf{R})$, hence its trace, which is nothing else than $u \circ f$, is in $H^{1/2}(\mathbf{S}^1,\mathbf{R})$. □

Then, let us prove that an $H^{1/2}$ function cannot have a jump.

Proposition B.2. *Let $f \in H^{1/2}([0,1],\mathbf{R})$ such that $f(\xi) \in \{0,1\}$ for a.e. $\xi \in [0,1]$. Then there is a representative of f which is constant.*

Proof. We reason by contraposition: we assume that f is not constant, which translates in $0 < \int_0^1 f < 1$ and we want to show that $f \notin H^{1/2}([0,1],\mathbf{R})$. Recall that it is sufficient to prove, given the definition of the $H^{1/2}$ norm [23, Chapter 3], that

$$\int_{[0,1]\times[0,1]} \frac{|f(\eta) - f(\theta)|}{|\theta - \eta|^2} \mathrm{d}\eta\mathrm{d}\theta = +\infty.$$

Take $t > 0$ large enough. The function

$$\xi \mapsto \frac{1}{\sqrt{t}} \int_{\xi - t^{-1/2}/2}^{\xi + t^{-1/2}/2} f(\eta)\mathrm{d}\eta$$

is continuous on $[t^{-1/2}/2, 1-t^{-1/2}/2]$ and has a means which belongs to $[c, 1-c]$, where $0<c<1$ is independent of t (provided it is large enough) and is related to $0<\int_0^1 f<1$. Hence, there exists ξ_t such that

$$\int_{\xi_t - t^{-1/2}/2}^{\xi_t + t^{-1/2}/2} f(\eta)\mathrm{d}\eta \in \left[\frac{c}{\sqrt{t}}, 1-\frac{c}{\sqrt{t}}\right].$$

Heuristically, ξ_t is close to a point where f "jumps". In particular, it implies that

$$\mathcal{L}_{[0,1]}\otimes\mathcal{L}_{[0,1]}\left(\left\{(\eta,\theta)\in\left[\xi_t-\frac{1}{2\sqrt{t}},\xi_t+\frac{1}{2\sqrt{t}}\right]^2 \ : \ f(\eta)=1 \text{ and } f(\theta)=0\right\}\right)\geqslant\frac{c^2}{t}.$$

As a consequence,

$$\mathcal{L}_{[0,1]}\otimes\mathcal{L}_{[0,1]}\left(\left\{(\eta,\theta)\in[0,1]^2 \ : \ \frac{|f(\eta)-f(\theta)|}{|\theta-\eta|^2}\geqslant t\right\}\right)\geqslant\frac{c^2}{t}.$$

This estimate leads to

$$\int_{[0,1]\times[0,1]}\frac{|f(\eta)-f(\theta)|}{|\theta-\eta|^2}\mathrm{d}\eta\mathrm{d}\theta=\int_0^{+\infty}\left[\mathcal{L}_{[0,1]}\otimes\mathcal{L}_{[0,1]}\left(\left\{(\eta,\theta)\in[0,1]^2 \ : \ \frac{|f(\eta)-f(\theta)|}{|\theta-\eta|^2}\geqslant t\right\}\right)\right]\mathrm{d}t=+\infty.$$

□

With these two lemmas, we can easily arrive to our conclusion.

Proof of Lemma 5.4. Let $f\in H^{1/2}(\mathbf{S}^1,\mathbf{S}^1)$ such that $f(\xi)^2=\xi$ for a.e. $\xi\in\mathbf{S}^1$. We want to show that f is continuous. Take X an arc of circle of $\mathbf{S}^1$. If X is small enough, there are two continuous functions f_0 and f_1 (the complex square roots) defined on X such that for all $\xi\in X$, $z^2=\xi$ if and only if $z\in\{f_0(\xi),f_1(\xi)\}$. Moreover, if X is small, the ranges of f_0 and f_1 are far apart, hence we can find a Lipschitz function $u:\mathbf{U}\to\{0,1\}$ such that $u\circ f_0=0$ and $u\circ f_1=1$ on X. Thus, $(u\circ f)(\xi)\in\{0,1\}$ for $\xi\in X$. The previous lemmas allow us to conclude that the function is in $H^{1/2}(X,\{0,1\})$, hence constant, which means that f is continuous on X. As X is arbitrary, f is continuous on $\mathbf{S}^1$, which is a contradiction. □

ACKNOWLEDGMENTS

The author acknowledges the financial support of the French ANR ISOTACE (ANR-12-MONU-0013) and benefited from the support of the FMJH "Program Gaspard Monge for Optimization and operations research and their interactions with data science" and EDF via the PGMO project VarPDEMFG.

He also thanks Yann Brenier, Guillaume Carlier, Filippo Santambrogio and Thierry De Pauw for fruitful discussions, precious advice and pointing him out relevant references. In particular, Theorem 4.4 was provided by Filippo Santambrogio and Thierry de Pauw.

REFERENCES

[1] Agueh, Martial, and Guillaume Carlier. Barycenters in the Wasserstein space. *SIAM Journal on Mathematical Analysis* 43.2 (2011): 904-924.

[2] Aliprantis, Charalambos D., and Kim Border. *Infinite dimensional analysis: a hitchhiker's guide*. Springer Science, 2006.

[3] Ambrosio, Luigi, and Paolo Tilli. *Topics on analysis in metric spaces*. Vol. 25. Oxford University Press, 2004.

[4] Ambrosio, Luigi, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Springer Science , 2008.

[5] Scheffer, Vladimir, and Jean E. Taylor, eds. *Almgren's big regularity paper: Q-valued functions minimizing Dirichlet's integral and the regularity of area-minimizing rectifiable currents up to codimension 2*. Vol. 1. World scientific, 2000.
[6] Arendt, Wolfgang. Semigroups and evolution equations: functional calculus, regularity and kernel estimates. *Handbook of Differential Equations: Evolutionary Equations* 1 (2002): 1-85.
[7] Benamou, Jean-David, and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik* 84.3 (2000): 375-393.
[8] Bigot, Jérémie, and Thierry Klein. Consistent estimation of a population barycenter in the Wasserstein space. ArXiv e-prints (2012).
[9] Brenier, Yann. Extended Monge-Kantorovich theory. Lecture Notes in mathematics – Springer Verlag – (2003): 91-122.
[10] De Lellis, Camillo, and Emanuele Nunzio Spadaro. *Q-valued functions revisited*. American Mathematical Soc., 2011.
[11] Eells, James, and Joseph H. Sampson. Harmonic mappings of Riemannian manifolds. *American journal of mathematics* 86.1 (1964): 109-160.
[12] Evans, Lawrence Craig, and Ronald F. Gariepy. *Measure theory and fine properties of functions*. CRC press, 1992.
[13] Gelbrich, Matthias. On a formula for the L2 Wasserstein metric between measures on Euclidean and Hilbert spaces. *Mathematische Nachrichten* 147.1 (1990): 185-203.
[14] Hélein, Frédéric, and John C. Wood. Harmonic maps. *Handbook of global analysis* 1213 (2008): 417-491.
[15] Ishihara, Tôru. A mapping of Riemannian manifolds which preserves harmonic functions. *Journal of Mathematics of Kyoto University* 19.2 (1979): 215-229.
[16] Jost, Jürgen. Equilibrium maps between metric spaces. *Calculus of Variations and Partial Differential Equations* 2.2 (1994): 173-204.
[17] Jost, Jürgen. *Riemannian geometry and geometric analysis*. Springer Science & Business Media, 2008.
[18] Kim, Young-Heon, and Brendan Pass. Wasserstein barycenters over Riemannian manifolds. *Advances in Mathematics* 307 (2017): 640-683.
[19] Korevaar, Nicholas J., and Richard M. Schoen. Sobolev spaces and harmonic maps for metric space targets. *Comm. Anal. Geom* 1.3-4 (1993): 561-659.
[20] Lang, U., and Schroeder, V. Kirszbraun's Theorem and Metric Spaces of Bounded Curvature. *Geometric & Functional Analysis GAFA* (1997) 7: 535.
[21] Lu, Zhuoran. *Properties of Soft Maps on Riemannian Manifolds*. ProQuest Dissertations Publishing 10617234. (2017).
[22] Matthes, Daniel, Robert J. McCann, and Giuseppe Savaré. A family of nonlinear fourth order equations of gradient flow type. *Communications in Partial Differential Equations* 34.11 (2009): 1352-1397.
[23] McLean, William Charles Hector. *Strongly elliptic systems and boundary integral equations*. Cambridge university press, 2000.
[24] Ohta, Shin-ichi. Extending Lipschitz and Hölder maps between metric spaces. *Positivity* 13.2 (2009): 407-425.
[25] Reshetnyak, Yu G. Sobolev classes of functions with values in a metric space. *Sibirskii Matematicheskii Zhurnal* 38.3 (1997): 657-675.
[26] Reshetnyak, Yu G. Sobolev-type classes of functions with values in a metric space. II. *Siberian Mathematical Journal* 45.4 (2004): 709-721.
[27] Santambrogio, Filippo. *Optimal transport for applied mathematicians: calculus of variations, pdes, and modeling*. Vol. 87. Birkhäuser, 2015.

[28] Solomon, Justin, Leonidas Guibas, and Adrian Butscher. Dirichlet energy for analysis and synthesis of soft maps. *Computer Graphics Forum*. Vol. 32. No. 5. Blackwell Publishing Ltd, 2013.
[29] Solomon, Justin, et al. Wasserstein propagation for semi-supervised learning. *Proceedings of the 31st International Conference on Machine Learning* (ICML-14). 2014.
[30] Sturm, Karl-Theodor. A semigroup approach to harmonic maps. *Potential Analysis* 23.3 (2005): 225-277.
[31] Villani, Cédric. *Topics in optimal transportation*. No. 58. American Mathematical Soc., 2003.
[32] Vogt, Thomas, and Lellmann, Jan. Measure-Valued Variational Models with Applications to Diffusion-Weighted Imaging. *arXiv preprint* arXiv:1710.00798, 2017.