A new class of transport distances between measures

Jean Dolbeault · **Bruno Nazaret** · **Giuseppe Savaré**

March 11, 2008

Abstract We introduce a new class of distances between nonnegative Radon measures in $\mathbb{R}^d$. They are modeled on the dynamical characterization of the Kantorovich-Rubinstein-Wasserstein distances proposed by BENAMOU-BRENIER [7] and provide a wide family interpolating between the Wasserstein and the homogeneous $W_\gamma^{-1,p}$-Sobolev distances.
From the point of view of optimal transport theory, these distances minimize a dynamical cost to move a given initial distribution of mass to a final configuration. An important difference with the classical setting in mass transport theory is that the cost not only depends on the velocity of the moving particles but also on the densities of the intermediate configurations with respect to a given reference measure γ.
We study the topological and geometric properties of these new distances, comparing them with the notion of weak convergence of measures and the well established Kantorovich-Rubinstein-Wasserstein theory. An example of possible applications to the geometric theory of gradient flows is also given.

Keywords Optimal transport · Kantorovich-Rubinstein-Wasserstein distance · Continuity equation · Gradient flows

1 Introduction

Starting from the contributions by Y. BRENIER, R. MCCANN, W. GANGBO, L.C. EVANS, F. OTTO, C. VILLANI [9,18,25,17,27], the theory of Optimal Transportation has received a lot of attention and many deep applications to various mathematical fields, such as PDE's, Calculus of Variations, functional and geometric inequalities, geometry of metric-measure spaces, have been found (we refer here to the monographs [28,16,30,3,31]). Among all possible transportation costs, those inducing the so-called L^p-KANTOROVICH-RUBINSTEIN-

J. Dolbeault, B. Nazaret
Ceremade (UMR CNRS no. 7534), Université Paris-Dauphine, place de Lattre de Tassigny, 75775 Paris Cédex 16, France.
E-mail: {dolbeaul, nazaret}@ceremade.dauphine.fr

G. Savaré
Università degli studi di Pavia, Department of Mathematics, Via Ferrata 1, 27100, Pavia, Italy. E-mail: giuseppe.savare@unipv.it

WASSERSTEIN distances $W_p(\mu_0,\mu_1)$, $p\in(1,+\infty)$, between two probability measures $\mu,\nu\in\mathcal{P}(\mathbb{R}^d)$

$$W_p(\mu_0,\mu_1):=\inf\left\{\left(\int_{\mathbb{R}^d\times\mathbb{R}^d}|y-x|^p\,\mathrm{d}\Sigma\right)^{\frac{1}{p}}:\ \Sigma\in\Gamma(\mu_0,\mu_1)\right\} \tag{1.1}$$

play a distinguished role. Here $\Gamma(\mu_0,\mu_1)$ is the set of all *couplings* between μ_0 and μ_1: they are probability measures Σ on $\mathbb{R}^d\times\mathbb{R}^d$ whose first and second marginals are respectively μ_0 and μ_1, i.e. $\Sigma(B\times\mathbb{R}^d)=\mu_0(B)$ and $\Sigma(\mathbb{R}^d\times B)=\mu_1(B)$ for all Borel sets $B\in\mathcal{B}(\mathbb{R}^d)$.

It was one of the most surprising achievement of [24,25,19,26] that many evolution partial differential equations of the form

$$\partial_t\rho+\nabla\cdot\big(\rho\,|\boldsymbol{\xi}|^{q-2}\boldsymbol{\xi}\big)=0,\qquad \boldsymbol{\xi}=-\nabla\Big(\frac{\delta\mathscr{F}}{\delta\rho}\Big)\qquad \text{in }\mathbb{R}^d\times(0,+\infty), \tag{1.2}$$

can be, at least formally, interpreted as *gradient flows* of suitable integral functionals $\mathscr{F}$ with respect to W_p (see also the general approach developed in [30,3,31]). In (1.2) $\delta\mathscr{F}/\delta\rho$ is the Euler first variation of $\mathscr{F}$, $q:=p/(p-1)$ is the Hölder's conjugate exponent of p, and $t\mapsto\rho_t$ (a time dependent solution of (1.2)) can be interpreted as a flow of probability measures $\mu_t=\rho_t\mathscr{L}^d$ with density ρ_t with respect to the Lebesgue measure $\mathscr{L}^d$ in $\mathbb{R}^d$.

Besides showing deep relations with entropy estimates and functional inequalities [27], this point of view provides a powerful variational method to prove existence of solutions to (1.2), by the so-called *Minimizing movement* scheme [19,13,3]: given a time step $\tau>0$ and an initial datum $\mu_0=\rho_0\mathscr{L}^d$, the solution $\mu_t=\rho_t\mathscr{L}^d$ at time $t\approx n\tau$ can be approximated by the discrete solution μ_τ^n obtained by a recursive minimization of the functional

$$\mu\mapsto\frac{1}{p\tau^{p-1}}W_p^p(\mu,\mu_\tau^k)+\mathscr{F}(\mu),\qquad k=0,1,\cdots \tag{1.3}$$

The link between the Wasserstein distance and equations exhibiting the characteristic structure of (1.2) (in particular the presence of the diffusion coefficient ρ, the fact that $\boldsymbol{\xi}$ is a gradient vector field, and the presence of the q-duality map $\boldsymbol{\xi}\mapsto|\boldsymbol{\xi}|^{q-2}\boldsymbol{\xi}$), is well explained by the *dynamic characterization* of W_p introduced by BENAMOU-BRENIER [7]: it relies in the minimization of the "action" integral functional

$$\begin{aligned}W_p^p(\mu_0,\mu_1)=\inf\Big\{&\int_0^1\int_{\mathbb{R}^d}\rho_t(x)\,|\mathbf{v}_t(x)|^p\,\mathrm{d}x\,\mathrm{d}t\ :\\ &\partial_t\rho_t+\nabla\cdot(\rho_t\mathbf{v}_t)=0\text{ in }\mathbb{R}^d\times(0,1),\quad \mu_0=\rho_{|_{t=0}}\mathscr{L}^d,\quad \mu_1=\rho_{|_{t=1}}\mathscr{L}^d\Big\}.\end{aligned} \tag{1.4}$$

Towards more general cost functionals. If one is interested to study the more general class of diffusion equations

$$\partial_t\rho+\nabla\cdot\big(h(\rho)\,|\boldsymbol{\xi}|^{q-2}\boldsymbol{\xi}\big)=0,\qquad \boldsymbol{\xi}=-\nabla\Big(\frac{\delta\mathscr{F}}{\delta\rho}\Big)\qquad \text{in }\mathbb{R}^d\times(0,+\infty), \tag{1.5}$$

obtained from (1.2) replacing the mobility coefficient ρ by an increasing nonlinear function $h(\rho)$, $h:[0,+\infty)\to[0,+\infty)$ whose typical examples are the functions $h(\rho)=\rho^\alpha$, $\alpha\geq 0$, it is then natural to investigate the properties of the "distance"

$$\begin{aligned}\widetilde{W}_p^p(\mu_0,\mu_1)=\inf\Big\{&\int_0^1\int_{\mathbb{R}^d}h\big(\rho_t(x)\big)|\mathbf{v}_t(x)|^p\,\mathrm{d}x\,\mathrm{d}t:\\ &\partial_t\rho_t+\nabla\cdot(h(\rho_t)\,\mathbf{v}_t)=0\text{ in }\mathbb{R}^d\times(0,1),\quad \mu_0=\rho_{|_{t=0}}\mathscr{L}^d,\quad \mu_1=\rho_{|_{t=1}}\mathscr{L}^d\Big\}.\end{aligned} \tag{1.6}$$

In the limiting case $\alpha = 0$, $h(\rho) \equiv 1$, one can easily recognize that (1.6) provides an equivalent description of the homogeneous (dual) $\dot W^{-1,p}(\mathbb{R}^d)$ Sobolev (pseudo)-distance

$$\|\mu_0-\mu_1\|_{\dot W^{-1,p}(\mathbb{R}^d)} := \sup\Big\{\int_{\mathbb{R}^d} \zeta \,\mathrm{d}(\mu_0-\mu_1) : \zeta \in C^1_{\mathrm{c}}(\mathbb{R}^d),\ \int_{\mathbb{R}^d} |\mathrm{D}\zeta|^q \,\mathrm{d}x \le 1\Big\}. \tag{1.7}$$

Thus the distances defined by (1.6) for $0 \le \alpha \le 1$ (we shall see that this is the natural range for the parameter α) can be considered as a natural "interpolating" family between the Wasserstein and the (dual) Sobolev ones.

Notice that if one wants to keep the usual transport interpretation given by a "dynamic cost" to be minimized along the solution of the continuity equation, one can simply introduce the velocity vector field $\tilde{\mathbf{v}}_t := \rho_t^{-1} h(\rho_t)\mathbf{v}_t$ and minimize the cost

$$\int_0^1 \int_{\mathbb{R}^d} \rho\, f(\rho)\, |\tilde{\mathbf{v}}_t|^p \,\mathrm{d}x\,\mathrm{d}t, \quad \text{where} \quad f(\rho) := \Big(\frac{\rho}{h(\rho)}\Big)^{p-1}. \tag{1.8}$$

Therefore, in this model the usual p-energy $\int_{\mathbb{R}^d} \rho_t |\tilde{\mathbf{v}}_t|^p \,\mathrm{d}x$ of the moving masses ρ_t with velocity $\tilde{\mathbf{v}}_t$ results locally modified by a factor $f(\rho_t)$ depending on the local density of the mass occupied at the time t. Different non-local models have been considered in [8,4].

In the present paper we try to present a systematic study of these families of intermediate distances, in view of possible applications, e.g., to the study of evolution equations like (1.5), the Minimizing movement approach (1.3), and functional inequalities.

Examples: PDE's as gradient flows. Let us show a few examples evolution equations which can be *formally* interpreted as gradient flows of suitable integral functionals in this setting: the scalar conservation law

$$\partial_t \rho - \nabla\cdot\big(\rho^\alpha \nabla V\big) = 0 \quad \text{corresponds to the linear functional} \quad \mathscr{F}(\rho) := \int_{\mathbb{R}^d} V(x)\,\rho\,\mathrm{d}x,$$

for some smooth potential $V : \mathbb{R}^d \to \mathbb{R}$ and $p = 2$. Choosing for $m > 0$

$$p = 2, \quad \mathscr{F}(\rho) = c_{\alpha,m} \int \rho^{m+1-\alpha}\,\mathrm{d}x, \quad c_{\alpha,m} := \frac{m}{(m+1-\alpha)(m-\alpha)},$$

one gets the porous media/fast diffusion equation

$$\partial_t \rho - \frac{m}{m-\alpha}\nabla\cdot\big(\rho^\alpha \nabla \rho^{m-\alpha}\big) = \partial_t\rho - \Delta\rho^m = 0, \tag{1.9}$$

and in particular the heat equation for the entropy functional $\frac{1}{(2-\alpha)(1-\alpha)} \int \rho^{2-\alpha}\,\mathrm{d}x$. Choosing

$$\mathscr{F}(\rho) = c_{\alpha,m,q} \int \rho^{\frac{m+2q-3-\alpha}{q-1}}\,\mathrm{d}x, \quad c_{\alpha,m,q} := \frac{m(q-1)^q}{(m+2q-3-\alpha)(m+q-2-\alpha)},$$

one obtains the doubly nonlinear equation

$$\partial_t \rho - m\nabla\cdot(\rho^{m-1}|\nabla\rho|^{q-2}\nabla\rho) = 0 \tag{1.10}$$

and in particular the evolution equation for the q-Laplacian when $m = 1$. The Dirichlet integral for $p = 2$

$$\mathscr{F}(\rho) = \frac{1}{2}\int |\nabla\rho|^2\,\mathrm{d}x \quad \text{yields} \quad \partial_t\rho + \nabla\cdot\big(\rho^\alpha \nabla\Delta\rho\big) = 0, \tag{1.11}$$

a thin-film like equation.

The measure-theoretic point of view: Wasserstein distance. We present now the main points of our approach (see also, in a different context, [10]). First of all, even if the language of densities and vector fields (as ρ and $\mathbf{v},\tilde{\mathbf{v}}$ in (1.4) or (1.6)) is simpler and suggests interesting interpretations, the natural framework for considering the variational problems (1.4) and (1.6) is provided by time dependent families of Radon measures in $\mathbb{R}^d$. Following this point of view, one can replace ρ_t by a continuous curve $t\in[0,1]\mapsto\mu_t$ ($\mu_t=\rho_t\mathscr{L}^d$ in the absolutely continuous case) in the space $\mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ of nonnegative Radon measures in $\mathbb{R}^d$ endowed with the usual weak* topology induced by the duality with functions in $C^0_{\rm c}(\mathbb{R}^d)$. The (Borel) vector field $\mathbf{v}_t$ in (1.4) induces a time dependent family of vector measures $\boldsymbol{\nu}_t:=\mu_t\mathbf{v}_t\ll\mu_t$. In terms of the couple $(\mu,\boldsymbol{\nu})$ the continuity equation (1.4) reads

$$\partial_t\mu_t+\nabla\cdot\boldsymbol{\nu}_t=0\quad\text{in the sense of distributions in } \mathscr{D}'(\mathbb{R}^d\times(0,1)), \tag{1.12}$$

and it is now a *linear* equation. Since $\mathbf{v}_t=\mathrm{d}\boldsymbol{\nu}_t/\mathrm{d}\mu_t$ is the density of $\boldsymbol{\nu}_t$ w.r.t. μ_t, the action functional which has to be minimized in (1.4) can be written as

$$\mathscr{E}_{p,1}(\mu,\boldsymbol{\nu})=\int_0^1\Phi_{p,1}(\mu_t,\boldsymbol{\nu}_t)\,\mathrm{d}t,\quad \Phi_{p,1}(\mu,\boldsymbol{\nu}):=\int_{\mathbb{R}^d}\left|\frac{\mathrm{d}\boldsymbol{\nu}}{\mathrm{d}\mu}\right|^p\mathrm{d}\mu. \tag{1.13}$$

Notice that in the case of absolutely continuous measures with respect to $\mathscr{L}^d$, i.e. $\mu=\rho\mathscr{L}^d$ and $\boldsymbol{\nu}=\mathbf{w}\mathscr{L}^d$, the functional $\Phi_{p,1}$ can also be expressed as

$$\Phi_{p,1}(\mu,\boldsymbol{\nu}):=\int_{\mathbb{R}^d}\phi_{p,1}(\rho,\mathbf{w})\,\mathrm{d}\mathscr{L}^d(x),\quad \phi_{p,1}(\rho,\mathbf{w}):=\rho\left|\frac{\mathbf{w}}{\rho}\right|^p. \tag{1.14}$$

Denoting by $\mathcal{CE}(0,1)$ the class of measure-valued distributional solutions $\mu,\boldsymbol{\nu}$ of the continuity equation (1.12), we end up with the equivalent characterization of the Kantorovich-Rubinstein-Wasserstein distance

$$W_p^p(\mu_0,\mu_1):=\inf\Big\{\mathscr{E}_{p,1}(\mu,\boldsymbol{\nu}):(\mu,\boldsymbol{\nu})\in\mathcal{CE}(0,1),\ \mu|_{t=0}=\mu_0,\ \mu|_{t=1}=\mu_1\Big\}. \tag{1.15}$$

Structural properties and convexity issues. The density function $\phi=\phi_{p,1}:(0,+\infty)\times\mathbb{R}^d\to\mathbb{R}^d$ appearing in (1.14) exhibits some crucial features

1. $\mathbf{w}\mapsto\phi(\cdot,\mathbf{w})$ is symmetric, positive (when $\mathbf{w}\neq0$), and p-homogeneous with respect to the vector variable $\mathbf{w}$: this ensures that W_p is symmetric and satisfies the triangular inequality.
2. ϕ is jointly convex in $(0,+\infty)\times\mathbb{R}^d$: this ensures that the functional $\Phi_{p,1}$ (and therefore also $\mathscr{E}$) defined in (1.13) is lower semicontinuous with respect to the weak* convergence of Radon measures. It is then possible to show that the infimum in (1.15) is attained, as soon as it is finite (i.e. when there exists at least one curve $(\mu,\boldsymbol{\nu})\in\mathcal{CE}(0,1)$ with finite energy $\mathscr{E}(\mu,\boldsymbol{\nu})$ joining μ_0 to μ_1); in particular $W_p(\mu_0,\mu_1)=0$ yields $\mu_0=\mu_1$. Moreover, the distance map $(\mu_0,\mu_1)\mapsto W_p(\mu_0,\mu_1)$ is lower semicontinuous with respect to the weak* convergence, a crucial property in many variational problems involving W_p, as (1.3).
3. ϕ is jointly positively 1-homogeneous: this a distinguished feature of the Wasserstein case, which shows that the functional $\Phi_{p,1}$ depends only on $\mu,\boldsymbol{\nu}$ and not on the Lebesgue measure $\mathscr{L}^d$, even if it can be represented as in (1.14). In other words, suppose that $\mu=\tilde{\rho}\gamma$ and $\boldsymbol{\nu}=\tilde{\mathbf{w}}\gamma$, where γ is another reference (Radon, nonnegative) measure in $\mathbb{R}^d$. Then
$$\Phi_{p,1}(\mu,\boldsymbol{\nu})=\int_{\mathbb{R}^d}\phi_{p,1}(\tilde{\rho},\tilde{\mathbf{w}})\,\mathrm{d}\gamma. \tag{1.16}$$

As we will show in this paper, the 1-homogeneity assumption yields also two "quantitative" properties: if μ_0 is a probability measure, then any solution $(\mu,\boldsymbol{\nu})$ of the continuity equation (1.12) with finite energy $\mathscr{E}(\mu,\boldsymbol{\nu}) < +\infty$ still preserves the mass $\mu_t(\mathbb{R}^d) \equiv 1$ for every time $t \geq 0$ (and it is therefore equivalent to assume this condition in the definition of $\mathcal{CE}(0,1)$, see e.g. [3, Chap. 8]). Moreover, if the p-moment of μ_0 $\mathsf{m}_p(\mu_0) := \int_{\mathbb{R}^d} |x|^p \,\mathrm{d}\mu_0(x)$ is finite, then $W_p(\mu_0,\mu_1) < +\infty$ if and only if $\mathsf{m}_p(\mu_1) < +\infty$.

Main definitions. Starting from the above remarks, it is then natural to consider the more general case when the density functional $\phi : (0,+\infty)\times\mathbb{R}^d \to [0,+\infty)$ still satisfies 1. (p-homogeneity w.r.t. $\mathbf{w}$) and 2. (convexity), but not 3. (1-homogeneity). Due to this last choice, the associated integral functional $\boldsymbol{\Phi}$ is no more independent of a reference measure γ and it seems therefore too restrictive to consider only the case of the Lebesgue measure $\gamma = \mathscr{L}^d$.

In the present paper we will thus introduce a further nonnegative *reference* Radon measure $\gamma \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ and a general convex functional $\phi : (0,+\infty)\times\mathbb{R}^d \to [0,+\infty)$ which is p-homogeneous w.r.t. its second (vector) variable and non degenerate (i.e. $\phi(\rho,\mathbf{w}) > 0$ if $\mathbf{w} \neq 0$). Particularly interesting examples of density functionals ϕ, corresponding to (1.6), are given by

$$\phi(\rho,\mathbf{w}) := h(\rho)\left|\frac{\mathbf{w}}{h(\rho)}\right|^p, \tag{1.17}$$

where $h : (0,+\infty) \to (0,+\infty)$ is an increasing and *concave* function; the concavity of h is a necessary and sufficient condition for the convexity of ϕ in (1.17) (see [29] and §3). Choosing $h(\rho) := \rho^\alpha$, $\alpha \in (0,1)$, one obtains

$$\phi_{p,\alpha}(\rho,\mathbf{w}) := \rho^\alpha\left|\frac{\mathbf{w}}{\rho^\alpha}\right|^p = \rho^{\theta-p}\,|\mathbf{w}|^p, \qquad \theta := (1-\alpha)\,p + \alpha \in (1,p), \tag{1.18}$$

which is jointly θ-homogeneous in $(\rho,\mathbf{w})$.

In the case, e.g., when $\alpha < 1$ in (1.18) or more generally $\lim_{\rho\uparrow\infty} h(\rho)/\rho = 0$, the *recession function* of ϕ satisfies

$$\phi^\infty(\rho,\mathbf{w}) = \lim_{\lambda\uparrow+\infty} \lambda^{-1}\phi(\lambda\rho,\lambda\mathbf{w}) = +\infty \qquad \text{if } \rho,\mathbf{w} \neq 0, \tag{1.19}$$

so that the associated integral functional reads as

$$\boldsymbol{\Phi}(\mu,\boldsymbol{\nu}|\gamma) := \int_{\mathbb{R}^d} \phi(\rho,\mathbf{w})\,\mathrm{d}\gamma \quad \mu = \rho\gamma + \mu^\perp, \quad \boldsymbol{\nu} = \mathbf{w}\gamma \ll \gamma, \tag{1.20}$$

extended to $+\infty$ when $\boldsymbol{\nu}$ is not absolutely continuous with respect to γ or $\mathrm{supp}(\mu) \not\subset \mathrm{supp}(\gamma)$. Notice that only the density ρ of the γ-absolutely continuous part of μ enters in the functional, but the functional could be finite even if μ has a singular part $\mu^\perp$. This choice is crucial in order to obtain a lower semicontinuous functional w.r.t. weak* convergence of measures. The associated (ϕ,γ)-Wasserstein distance is therefore

$$\mathcal{W}^p_{\phi,\gamma}(\mu_0,\mu_1) := \inf\Big\{\mathscr{E}_{\phi,\gamma}(\mu,\boldsymbol{\nu}) : (\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,1),\ \mu_{|t=0} = \mu_0,\ \mu_{|t=1} = \mu_1\Big\}, \tag{1.21}$$

where the energy $\mathscr{E}_{\phi,\gamma}$ of a curve $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,1)$ is

$$\mathscr{E}_{\phi,\gamma}(\mu,\boldsymbol{\nu}) := \int_0^1 \boldsymbol{\Phi}(\mu_t,\boldsymbol{\nu}_t|\gamma)\,\mathrm{d}t. \tag{1.22}$$

The most important case associated to the functional (1.18) deserves the distinguished notation

$$W_{p,\alpha;\gamma}(\cdot,\cdot) := \mathcal{W}_{\phi_{p,\alpha},\gamma}(\cdot,\cdot). \tag{1.23}$$

The limiting case $\alpha = \theta = 1$ corresponds to the L^p-Wasserstein distance, the Sobolev $\dot{W}_\gamma^{-1,p}$ corresponds to $\alpha = 0$, $\theta = p$. The choice of γ allows for a great flexibility: besides the Lebesgue measure in $\mathbb{R}^d$, we quote

– $\gamma := \mathscr{L}^d|_\Omega$, Ω being an open subset of $\mathbb{R}^d$. The measures are then supported in $\bar{\Omega}$ and, with the choice (1.17) and $\mathbf{v} = \mathbf{w}/h(\rho)$, (1.12) is a weak formulation of the continuity equation ($\mathbf{n}_{\partial\Omega}$ being the exterior unit normal to $\partial\Omega$)

$$\partial_t \rho_t + \nabla \cdot \big(h(\rho_t)\,\mathbf{v}_t\big) = 0 \quad \text{in } \Omega \times (0,1), \quad \mathbf{v}_t \cdot \mathbf{n}_{\partial\Omega} = 0 \quad \text{on } \partial\Omega. \tag{1.24}$$

This choice is useful for studying equations (1.9) (see [11]), (1.10), (1.11) in bounded domains with Neumann boundary conditions.

– $\gamma := e^{-V}\mathscr{L}^d$ for some C^1 potential $V : \mathbb{R}^d \to \mathbb{R}$. With the choice (1.17) and $\mathbf{v} = \mathbf{w}/h(\rho)$ (1.12) is a weak formulation of the equation

$$\partial_t \rho_t + \nabla \cdot \big(h(\rho_t)\,\mathbf{v}_t\big) - h(\rho_t)\nabla V \cdot \mathbf{v}_t = 0 \quad \text{in } \mathbb{R}^d \times (0,1). \tag{1.25}$$

When $h(\rho) = \rho^\alpha$, $p = 2$, the gradient flow of $\mathscr{F}(\mu) := \frac{1}{(2-\alpha)(1-\alpha)} \int_{\mathbb{R}^d} \rho^{2-\alpha}\, d\gamma$ is the Kolmogorov-Fokker-Planck equations [15]

$$\partial_t \mu - \Delta\mu - \nabla \cdot (\mu \nabla V) = 0, \quad \partial_t \rho - \Delta\rho + \nabla V \cdot \nabla \rho = 0,$$

which in the Wasserstein framework is generated by the logarithmic entropy ([19, 3, 5]).

– $\gamma := \mathscr{H}^k|_{\mathbb{M}}$, $\mathbb{M}$ being a smooth k-dimensional manifold embedded in $\mathbb{R}^d$ with the Riemannian metric induced by the Euclidean distance; $\mathscr{H}^k$ denotes the k-dimensional Hausdorff measure. (1.12) is a weak formulation of

$$\partial_t \rho_t + \mathrm{div}_{\mathbb{M}}\big(h(\rho)\mathbf{v}_t\big) = 0 \quad \text{on } \mathbb{M} \times (0,1). \tag{1.26}$$

Thanks to Nash embedding theorems [22, 23], the study of the continuity equation and of the weighted Wasserstein distances on arbitrary Riemannian manifolds can be reduced to this case, which could be therefore applied to study equations (1.9), (1.10), (1.11) on Riemannian manifolds.

Main results. Let us now summarize some of the main properties of $W_{p,\alpha;\gamma}(\cdot,\cdot)$ we will prove in the last section of the present paper. In order to deal with distances (instead of pseudo-distances, possibly assuming the value $+\infty$), for a nonnegative Radon measure σ we will denote by $\mathcal{M}_{p,\alpha;\gamma}[\sigma]$ the set of all measures μ with $W_{p,\alpha;\gamma}(\mu,\sigma) < +\infty$ endowed with the $W_{p,\alpha;\gamma}$-distance.

1. $\mathcal{M}_{p,\alpha;\gamma}[\sigma]$ is a complete metric space (Theorem 5.7).
2. $W_{p,\alpha;\gamma}$ induces a stronger convergence than the usual weak* one (Theorem 5.5).
3. Bounded sets in $\mathcal{M}_{p,\alpha;\gamma}[\sigma]$ are weakly* relatively compact (Theorem 5.5).
4. The map $(\mu_0,\mu_1) \mapsto W_{p,\alpha;\gamma}(\mu_0,\mu_1)$ is weakly* lower semicontinuous (Theorem 5.6), convex (Theorem 5.11), and subadditive (Theorem 5.12). It enjoys some useful monotonicity properties with respect to γ (Proposition 5.14) and to convolution (Theorem 5.15).

5. The infimum in (1.15) is attained, $\mathcal{M}_{p,\alpha;\gamma}[\sigma]$ is a geodesic space (Theorem 5.4), and constant speed geodesics connecting two measures μ_0, μ_1 are unique (Theorem 5.11).
6. If
$$\int_{|x|\geq 1} |x|^{-p/(\theta-1)}\,\mathrm{d}\gamma(x) < +\infty \qquad \theta = (1-\alpha)p+\alpha, \quad \frac{p}{\theta-1} = \frac{q}{1-\alpha}, \tag{1.27}$$
and $\sigma \in \mathcal{P}(\mathbb{R}^d)$, then $\mathcal{M}_{p,\alpha;\gamma}[\sigma] \subset \mathcal{P}(\mathbb{R}^d)$ (Theorem 5.8). If moreover γ satisfies stronger summability assumptions, then the distances $W_{p,\alpha;\gamma}$ provide a control of various moments of the measures (Theorem 5.9). Comparison results with W_p and $\dot{W}^{-1,p}$ are also discussed in §5.4.
7. Absolutely continuous curves w.r.t. $W_{p,\alpha;\gamma}$ can be characterized in completely analogous ways as in the Wasserstein case (§5.3).
8. In the case $\gamma = \mathscr{L}^d$ the functional
$$\Psi_\alpha(\mu|\gamma) := \frac{1}{(2-\alpha)(1-\alpha)} \int_{\mathbb{R}^d} \rho^{2-\alpha}\,\mathrm{d}x \qquad \mu = \rho\mathscr{L}^d \ll \mathscr{L}^d, \tag{1.28}$$
is geodesically convex w.r.t. the distance $W_{2,\alpha;\mathscr{L}^d}$ and the heat equation in $\mathbb{R}^d$ is its gradient flow, as formally suggested by (1.9) (§5.5: we prove this property in the case $\alpha > 1-2/d$, when $\mathcal{P}(\mathbb{R}^d)$ is complete w.r.t. $W_{2,\alpha;\mathscr{L}^d}$.)

Plan of the paper. Section 2 recalls some basic notation and preliminary facts about weak* convergence and integral functionals of Radon measures; 2.3 recalls a simple duality result in convex analysis, which plays a crucial role in the analysis of the integrand $\phi(\rho,\mathbf{w})$.

The third section is devoted to the class of admissible action integral functionals Φ like (1.20) and their density ϕ. Starting from a few basic structural assumptions on ϕ we deduce its main properties and we present some important examples in Section 3.2. The corresponding properties of Φ (in particular, lower semicontinuity and relaxation with respect to weak* convergence, monotonicity, etc) are considered in Section 3.3.

Section 4 is devoted to the study of measure-valued solutions of the continuity equation (1.12). It starts with some preliminary basic results, which extend the theory presented in [3] to the case of general Radon measures: this extension is motivated by the fact that the class of probability measures (and therefore with finite mass) is too restrictive to study the distances $W_{p,\alpha;\gamma}$, in particular when $\gamma(\mathbb{R}^d) = +\infty$ as in the case of the Lebesgue measure. We shall see (Remark 5.27) that $\mathcal{P}(\mathbb{R}^d)$ with the distance $W_{p,\alpha;\mathscr{L}^d}$ is not complete if $d > p/(\theta-1) = q/(1-\alpha)$. We consider in Section 4.2 the class of solutions of (1.12) with finite energy $\mathscr{E}_{\phi,\gamma}$ (1.22), deriving all basic estimates to control their mass and momentum.

As we briefly showed, Section 5 contains all main results of the paper concerning the modified Wasserstein distances.

2 Notation and preliminaries

Here is a list of the main notation used throughout the paper:

B_R	The open ball (in some $\mathbb{R}^h$) of radius R centered at 0
$\mathcal{B}(\mathbb{R}^h)$ (resp. $\mathcal{B}_c(\mathbb{R}^h)$)	Borel subsets of $\mathbb{R}^h$ (resp. with compact closure)
$\mathcal{P}(\mathbb{R}^h)$	Borel probability measures in $\mathbb{R}^h$
$\mathcal{M}^+(\mathbb{R}^h)$ (resp. $\mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^h)$)	Finite (resp. Radon), nonnegative Borel measures on $\mathbb{R}^h$
$\mathcal{P}(\mathbb{R}^h)$	Borel probability measures in $\mathbb{R}^h$

$\mathcal{M}(\mathbb{R}^h;\mathbb{R}^m)$	$\mathbb{R}^m$-valued Borel measures with finite variation
$\mathcal{M}_{\rm loc}(\mathbb{R}^h;\mathbb{R}^m)$	$\mathbb{R}^m$-valued Radon measures
$\lVert\boldsymbol{\mu}\rVert$	Total variation of $\boldsymbol{\mu}\in\mathcal{M}_{\rm loc}(\mathbb{R}^h;\mathbb{R}^m)$, see (2.2)
$C^0_b(\mathbb{R}^h)$	Continuous and *bounded* real functions
$\mathsf{m}_\mathsf{p}(\mu)$	p-moment $\int_{\mathbb{R}^d}\lvert x\rvert^p\,\mathrm{d}\mu$ of $\mu\in\mathcal{M}^+(\mathbb{R}^h)$
ψ^∞	Recession function of ψ, see (2.4)
$\boldsymbol{\Psi}(\boldsymbol{\mu}\lvert\gamma),\ \boldsymbol{\Phi}(\mu,\boldsymbol{\nu}\lvert\gamma)$	Integral functionals on measures, see 2.2 and 3.3
$\mu(\zeta),\ \langle\mu,\zeta\rangle,\ \langle\boldsymbol{\mu},\boldsymbol{\zeta}\rangle$	the integrals $\int_{\mathbb{R}^d}\zeta\,\mathrm{d}\mu$, $\int_{\mathbb{R}^d}\boldsymbol{\zeta}\cdot\mathrm{d}\boldsymbol{\mu}$
$\mathcal{CE}(0,T),\mathcal{CE}_{\phi,\gamma}(0,T),$ $\mathcal{CE}(0,T;\mu_0\to\mu_1)$	Classes of measure-valued solutions of the continuity equation, see Def. 4.2 and Sec. 4.2.

2.1 Measures and weak convergence

We recall some basic notation and properties of weak convergence of (vector) radon measures (see e.g. [2]). A Radon vector measure in $\mathcal{M}_{\rm loc}(\mathbb{R}^h;\mathbb{R}^m)$ is a $\mathbb{R}^m$-valued map $\boldsymbol{\mu}:\mathcal{B}_{\rm c}(\mathbb{R}^h)\to\mathbb{R}^m$ defined on the Borel sets of $\mathbb{R}^h$ with compact closure. We identify $\boldsymbol{\mu}\in\mathcal{M}_{\rm loc}(\mathbb{R}^h;\mathbb{R}^m)$ with a vector $(\boldsymbol{\mu}^1,\boldsymbol{\mu}^2,\cdots,\boldsymbol{\mu}^m)$ of m measures in $\mathcal{M}_{\rm loc}(\mathbb{R}^h)$: its integral with a continuous vector valued function with compact support $\boldsymbol{\zeta}\in C^0_{\rm c}(\mathbb{R}^h;\mathbb{R}^m)$ is given by

$$\langle\boldsymbol{\mu},\boldsymbol{\zeta}\rangle:=\int_{\mathbb{R}^h}\boldsymbol{\zeta}\cdot\mathrm{d}\boldsymbol{\mu}=\sum_{i=1}^m\int_{\mathbb{R}^h}\boldsymbol{\zeta}^i(x)\,\mathrm{d}\boldsymbol{\mu}^i(x). \tag{2.1}$$

It is well known that $\mathcal{M}_{\rm loc}(\mathbb{R}^h;\mathbb{R}^m)$ can be identified with the dual of $C^0_{\rm c}(\mathbb{R}^h;\mathbb{R}^m)$ by the above duality pairing and it is therefore endowed with the corresponding of weak* topology. If $\lVert\cdot\rVert$ is a norm in $\mathbb{R}^d$ with dual $\lVert\cdot\rVert_*$ (in particular the euclidean norm $\lvert\cdot\rvert$) for every open subset $A\subset\mathbb{R}^h$ we have

$$\lVert\boldsymbol{\mu}\rVert(A)=\sup\Big\{\int_{\mathbb{R}^h}\boldsymbol{\zeta}\cdot\mathrm{d}\boldsymbol{\mu}:\quad \mathrm{supp}(\boldsymbol{\zeta})\subset A,\quad \lVert\boldsymbol{\zeta}(x)\rVert_*\le 1\quad\forall x\in\mathbb{R}^h\Big\}. \tag{2.2}$$

$\lVert\boldsymbol{\mu}\rVert$ is in fact a Radon positive measure in $\mathcal{M}^+_{\rm loc}(\mathbb{R}^h)$ and $\boldsymbol{\mu}$ admits the polar decomposition $\boldsymbol{\mu}=\mathbf{w}\lVert\boldsymbol{\mu}\rVert$ where the Borel vector field $\mathbf{w}$ belongs to $L^1_{\rm loc}(\lVert\boldsymbol{\mu}\rVert;\mathbb{R}^m)$. We thus have

$$\langle\boldsymbol{\mu},\boldsymbol{\zeta}\rangle=\int_{\mathbb{R}^h}\boldsymbol{\zeta}\cdot\mathrm{d}\boldsymbol{\mu}=\int_{\mathbb{R}^h}\boldsymbol{\zeta}\cdot\mathbf{w}\,\mathrm{d}\lVert\boldsymbol{\mu}\rVert. \tag{2.3}$$

If $(\boldsymbol{\mu}_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{M}_{\rm loc}(\mathbb{R}^h;\mathbb{R}^m)$ with $\sup_n\lVert\boldsymbol{\mu}\rVert(B_R)<+\infty$ for every open ball B_R, then it is possible to extract a subsequence $\boldsymbol{\mu}_{k_n}$ weakly* convergent to $\boldsymbol{\mu}\in\mathcal{M}(\mathbb{R}^h;\mathbb{R}^m)$, whose total variation $\lVert\boldsymbol{\mu}_{k_n}\rVert$ weakly* converges to $\lambda\in\mathcal{M}^+(\mathbb{R}^h)$ with $\lVert\boldsymbol{\mu}\rVert\le\lambda$.

2.2 Convex functionals defined on Radon measures

Let $\psi:\mathbb{R}^m\to[0,+\infty]$ be a convex and lower semicontinuous function with $\psi(0)=0$, whose proper domain $D(\psi):=\{x\in\mathbb{R}^m:\psi(x)<+\infty\}$ has non empty interior. Its *recession function* (see e.g. [2]) $\psi^\infty:\mathbb{R}^m\to[0,+\infty]$ is defined as

$$\psi^\infty(y):=\lim_{r\to+\infty}\frac{\psi(ry)}{r}=\sup_{r>0}\frac{\psi(ry)}{r}. \tag{2.4}$$

ψ^∞ is still convex, lower semicontinuous, and positively 1-homogeneous, so that its proper domain $D(\psi^\infty)$ is a convex cone always containing 0. We say that

$$\begin{gathered}\psi \text{ has a } \textit{superlinear growth} \text{ if } \psi^\infty(y)=\infty \text{ for every } y\neq 0\text{: } D(\psi^\infty)=\{0\},\\ \psi \text{ has a } \textit{sublinear growth} \text{ if } \psi^\infty(y)\equiv 0 \text{ for every } y\in\mathbb{R}^m.\end{gathered}\tag{2.5}$$

Let now $\gamma\in\mathcal{M}^+_{\rm loc}(\mathbb{R}^h)$ and $\boldsymbol{\mu}\in\mathcal{M}_{\rm loc}(\mathbb{R}^h;\mathbb{R}^m)$ with $\operatorname{supp}(\boldsymbol{\mu})\subset\operatorname{supp}(\gamma)$; the Lebesgue decomposition of $\boldsymbol{\mu}$ w.r.t. γ reads $\boldsymbol{\mu}=\boldsymbol{\vartheta}\gamma+\boldsymbol{\mu}^\perp$, where $\boldsymbol{\vartheta}=\mathrm{d}\boldsymbol{\mu}/\mathrm{d}\gamma$. We can introduce a nonnegative Radon measure $\sigma\in\mathcal{M}^+_{\rm loc}(\mathbb{R}^h)$ such that $\boldsymbol{\mu}^\perp=\boldsymbol{\vartheta}^\perp\sigma\ll\sigma$, e.g. $\sigma=|\boldsymbol{\mu}^\perp|$ and we set

$$\Psi^a(\boldsymbol{\mu}|\gamma):=\int_{\mathbb{R}^h}\psi(\boldsymbol{\vartheta}(x))\,\mathrm{d}\gamma(x),\qquad \Psi^\infty(\boldsymbol{\mu}|\gamma):=\int_{\mathbb{R}^h}\psi^\infty(\boldsymbol{\vartheta}^\perp(y))\,\mathrm{d}\sigma(y),\tag{2.6}$$

and finally

$$\Psi(\boldsymbol{\mu}|\gamma):=\Psi^a(\boldsymbol{\mu}|\gamma)+\Psi^\infty(\boldsymbol{\mu}|\gamma);\quad \Psi^\infty(\boldsymbol{\mu}|\gamma)=+\infty \text{ if } \operatorname{supp}(\boldsymbol{\mu})\not\subset\operatorname{supp}(\gamma).\tag{2.7}$$

Since ψ^∞ is 1-homogeneous, the definition of Ψ^∞ depends on γ only through its support and it is independent of the particular choice of σ in (2.6). When ψ has a superlinear growth then the functional Ψ is finite iff $\boldsymbol{\mu}\ll\gamma$ and $\Psi^a(\boldsymbol{\mu}|\gamma)$ is finite; in this case $\Psi(\mu|\gamma)=\Psi^a(\boldsymbol{\mu}|\gamma)$.

Theorem 2.1 (L.s.c. and relaxation of integral functionals of measures [1,2]) *Let us consider two sequences $\gamma_n\in\mathcal{M}^+_{\rm loc}(\mathbb{R}^h),\boldsymbol{\mu}_n\in\mathcal{M}_{\rm loc}(\mathbb{R}^h;\mathbb{R}^m)$ weakly* converging to $\gamma\in\mathcal{M}^+_{\rm loc}(\mathbb{R}^h)$ and $\boldsymbol{\mu}\in\mathcal{M}_{\rm loc}(\mathbb{R}^h;\mathbb{R}^m)$ respectively. We have*

$$\liminf_{n\uparrow+\infty}\Psi(\boldsymbol{\mu}_n|\gamma_n)\geq\Psi(\boldsymbol{\mu}|\gamma).\tag{2.8}$$

Let conversely $\boldsymbol{\mu},\gamma$ be such that $\Psi(\boldsymbol{\mu}|\gamma)<+\infty$. Then there exists a sequence $\boldsymbol{\mu}_n=\boldsymbol{\vartheta}_n\gamma\ll\gamma$ weakly* converging to $\boldsymbol{\mu}$ such that

$$\lim_{n\uparrow+\infty}\Psi^a(\boldsymbol{\mu}_n|\gamma)=\lim_{n\uparrow+\infty}\int_{\mathbb{R}^h}\psi(\boldsymbol{\vartheta}_n(x))\,\mathrm{d}\gamma(x)=\Psi(\boldsymbol{\mu}|\gamma).\tag{2.9}$$

Theorem 2.2 (Montonicity w.r.t. γ) *If $\gamma_1\leq\gamma_2$ then*

$$\Psi(\boldsymbol{\mu}|\gamma_2)\leq\Psi(\boldsymbol{\mu}|\gamma_1).\tag{2.10}$$

Proof Thanks to Theorem 2.1, it is sufficient to prove the above inequality for $\boldsymbol{\mu}\ll\gamma^1$. Since $\gamma_1=\theta\gamma_2$, with density $\theta\leq 1$ γ_2-a.e., we have $\boldsymbol{\mu}=\boldsymbol{\vartheta}^i\gamma^i$ with $\boldsymbol{\vartheta}^2=\theta\,\boldsymbol{\vartheta}^1$, and therefore

$$\int_{\mathbb{R}^d}\psi(\boldsymbol{\vartheta}_1)\,\mathrm{d}\gamma_1=\int_{\mathbb{R}^d}\psi(\theta^{-1}\boldsymbol{\vartheta}_2)\theta\,\mathrm{d}\gamma_2\geq\int_{\mathbb{R}^d}\psi(\boldsymbol{\vartheta}_2)\,\mathrm{d}\gamma_2,\tag{2.11}$$

where we used the property $\theta\psi(\theta^{-1}x)\geq\psi(x)$ for $\theta\leq 1$, being $\psi(0)=0$. □

Theorem 2.3 (Monotonicity with respect to convolution) *If $k\in C^\infty_{\rm c}(\mathbb{R}^d)$ is a convolution kernel satisfying $k(x)\geq 0,\quad \int_{\mathbb{R}^d}k(x)\,\mathrm{d}x=1$, then*

$$\Psi(\boldsymbol{\mu}*k|\gamma*k)\leq\Psi(\boldsymbol{\mu}|\gamma).\tag{2.12}$$

The *proof* follows the same argument of [3, Lemma 8.1.10], by observing that the map $(x,y)\mapsto x\psi(y/x)$ is convex and positively 1-homogeneous in $(0,+\infty)\times\mathbb{R}^d$.

2.3 **A duality result in convex analysis**

Let X,Y be Banach spaces and let A be an open convex subset of X. We consider a convex (and a fortiori continuous) function $\phi : A\times Y \to \mathbb{R}$ and its partial Legendre transform

$$\tilde\phi(x,y^*) := \sup_{y\in Y} \langle y^*,y\rangle - \phi(x,y) \in (-\infty,+\infty], \quad \forall x\in A,\ y^*\in Y^*. \tag{2.13}$$

The following duality result is well known in the framework of minimax problems [29].

Theorem 2.4 *$\tilde\phi$ is a l.s.c. function and there exists a convex set $Y^*_o \subset Y^*$ such that*

$$\tilde\phi(x,y^*) < +\infty \quad \Leftrightarrow \quad y^*\in Y^*_o, \tag{2.14}$$

so that $\tilde\phi(\cdot,y^*)\equiv +\infty$ for every $y^*\in Y^*\setminus Y^*_o$ and ϕ admits the dual representation formula

$$\phi(x,y) = \sup_{y^*\in Y^*_o} \langle y,y^*\rangle - \tilde\phi(x,y^*) \quad \forall x\in A,\ y\in Y. \tag{2.15}$$

For every $y^*\in Y^*_o$ we have

$$\text{the map}\quad x\mapsto \tilde\phi(x,y^*) \quad \text{is concave (and continuous) in } A. \tag{2.16}$$

Conversely, a function $\phi : A\times Y\to\mathbb{R}$ is convex if it admits the dual representation (2.15) *for a function $\tilde\phi$ satisfying* (2.16).

Proof Let us first show that (2.16) holds. For a fixed $y^*\in Y^*$, $x_0,x_1>0$, $\theta\in[0,1]$, and arbitrary $y_i\in Y$, we get

$$\begin{aligned}\tilde\phi((1-\vartheta)x_0+\vartheta x_1,y^*) &\ge \langle y^*,(1-\vartheta)y_0+\vartheta y_1\rangle - \phi((1-\vartheta)x_0+\vartheta x_1,(1-\vartheta)y_0+\vartheta y_1)\\ &\ge (1-\vartheta)\Big(\langle y^*,y_0\rangle - \phi(x_0,y_0)\Big) + \vartheta\Big(\langle y^*,y_1\rangle-\phi(x_1,y_1)\Big).\end{aligned}$$

Taking the supremum with respect to y_0,y_1 we eventually get

$$\tilde\phi((1-\vartheta)x_0+\vartheta x_1,y^*) \ge (1-\vartheta)\tilde\phi(x_0,y^*)+\vartheta\tilde\phi(x_1,y^*) \tag{2.17}$$

and we conclude that $\tilde\phi(\cdot,y^*)$ is concave. In particular, if it takes the value $+\infty$ at some point it should be identically $+\infty$ so that (2.14) holds.

The converse implication is even easier, since (2.15) exhibits ϕ as a supremum of continuous and convex functions (jointly in $x\in A, y\in Y$). □

3 **Action functionals**

The aim of this section is to study some property of integral functionals of the type

$$\boldsymbol{\Phi}^a(\mu,\boldsymbol{\nu}|\gamma) := \int_{\mathbb{R}^d}\phi(\rho,\mathbf{w})\,\mathrm{d}\gamma, \quad \mu=\rho\gamma\in\mathcal{M}^+_{\rm loc}(\mathbb{R}^d),\ \boldsymbol{\nu}=\mathbf{w}\gamma\in\mathcal{M}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d) \tag{3.1}$$

and their relaxation, when ϕ satisfies suitable convexity and homogeneity properties.

3.1 Action density functions

Let us therefore consider a nonnegative density function $\phi : (0,+\infty)\times\mathbb{R}^d \to [0,+\infty)$ and an exponent $p \in (1,+\infty)$ satisfying the following assumptions

$$\phi \quad \text{is convex and (a fortiori) continuous,} \tag{3.2a}$$

$$\begin{gathered}\mathbf{w} \mapsto \phi(\cdot,\mathbf{w}) \text{ is homogeneous of degree } p \text{, i.e.}\\ \phi(\rho,\lambda\mathbf{w}) = |\lambda|^p\phi(\rho,\mathbf{w}) \quad \forall\rho>0,\,\lambda\in\mathbb{R},\,\mathbf{w}\in\mathbb{R}^d,\end{gathered} \tag{3.2b}$$

$$\exists\rho_0>0: \quad \phi(\rho_0,\cdot) \quad \text{is non degenerate, i.e.} \qquad \phi(\rho_0,\mathbf{w})>0 \quad \forall\mathbf{w}\in\mathbb{R}^d\setminus\{0\}. \tag{3.2c}$$

Let $q = p/(p-1) \in (1,+\infty)$ be the usual conjugate exponent of p. We denote by $\tilde\phi : (0,+\infty)\times\mathbb{R}^d \to (-\infty,+\infty]$ the partial Legendre transform

$$\frac{1}{q}\tilde\phi(\rho,\mathbf{z}) := \sup_{\mathbf{w}\in\mathbb{R}^d} \mathbf{z}\cdot\mathbf{w} - \frac{1}{p}\phi(\rho,\mathbf{w}) \quad \forall\rho>0,\mathbf{z}\in\mathbb{R}^d. \tag{3.2d}$$

We collect some useful properties of such functions in the following result.

Theorem 3.1 *Let $\phi : (0,+\infty)\times\mathbb{R}^d \to \mathbb{R}^d$ satisfy* (3.2a,b,c). *Then*

1. *For every $\rho>0$ the function $\mathbf{w}\mapsto\phi(\rho,\mathbf{w})^{1/p}$ is a norm of $\mathbb{R}^d$ whose dual norm is given by $\mathbf{z}\mapsto\tilde\phi(\rho,\mathbf{z})^{1/q}$, i.e.*

$$\tilde\phi(\rho,\mathbf{z})^{1/q} = \sup_{\mathbf{w}\neq0}\frac{\mathbf{w}\cdot\mathbf{z}}{\phi(\rho,\mathbf{w})^{1/p}}, \qquad \phi(\rho,\mathbf{w})^{1/p} = \sup_{\mathbf{z}\neq0}\frac{\mathbf{w}\cdot\mathbf{z}}{\tilde\phi(\rho,\mathbf{z})^{1/q}}. \tag{3.3}$$

   In particular $\tilde\phi(\cdot,\mathbf{z})$ is q-homogeneous with respect to $\mathbf{z}$.

2. *The marginal conjugate function $\tilde\phi$ takes its values in $[0,+\infty)$ and for every $\mathbf{z}\in\mathbb{R}^d$*

$$\textit{the map } \rho\mapsto\tilde\phi(\rho,\mathbf{z}) \quad \textit{is} \text{ concave and non decreasing } \textit{in } (0,+\infty). \tag{3.4}$$

   In particular, for every $\mathbf{w}\in\mathbb{R}^d$

$$\textit{the map } \rho\mapsto\phi(\rho,\mathbf{w}) \quad \textit{is} \text{ convex and non increasing } \textit{in } (0,+\infty). \tag{3.5}$$

3. *There exist constants $a,b\geq0$ such that*

$$\tilde\phi(\rho,\mathbf{z}) \leq \big(a+b\rho\big)|\mathbf{z}|^q, \quad \phi(\rho,\mathbf{z}) \geq \big(a+b\rho\big)^{1-p}|\mathbf{w}|^p \qquad \forall\rho>0,\ \mathbf{z},\mathbf{w}\in\mathbb{R}^d. \tag{3.6}$$

4. *For every closed interval $[\rho_0,\rho_1]\subset(0,+\infty)$ there exists a constant $C=C_{\rho_0,\rho_1}>0$ such that for every $\rho\in[\rho_0,\rho_1]$*

$$C^{-1}|\mathbf{w}|^p \leq \phi(\rho,\mathbf{w}) \leq C|\mathbf{w}|^p, \quad C^{-1}|\mathbf{z}|^q \leq \tilde\phi(\rho,\mathbf{z}) \leq C|\mathbf{z}|^q \quad \forall\mathbf{w},\mathbf{z}\in\mathbb{R}^d. \tag{3.7}$$

Equivalently, a function ϕ satisfies (3.2a,b,c) *if and only if it admits the dual representation formula*

$$\frac{1}{p}\phi(\rho,\mathbf{w}) = \sup_{\mathbf{z}\in\mathbb{R}^d}\mathbf{w}\cdot\mathbf{z} - \frac{1}{q}\tilde\phi(\rho,\mathbf{z}) \quad \forall\rho>0,\mathbf{w}\in\mathbb{R}^d, \tag{3.8}$$

where $\tilde\phi : (0,+\infty)\times\mathbb{R}^d \to (0,+\infty)$ is a nonnegative function which is convex and q-homogeneous w.r.t. $\mathbf{z}$ and concave *with respect to ρ.*

Proof Let us first assume that ϕ satisfies (3.2a,b,c). The function $\mathbf{w} \mapsto \phi(\rho,\mathbf{w})^{1/p}$ is 1-homogeneous and its sublevels are convex, i.e. it is the gauge function of a (symmetric) convex set and therefore it is a (semi)-norm. The concavity of $\tilde\phi$ follows from Theorem 2.4; taking $\mathbf{w} = 0$ in (3.2d), we easily get that $\tilde\phi$ is nonnegative; (3.2c) yields, for a suitable constant $c_0 > 0$,

$$\phi(\rho_0,\mathbf{w}) \geq c_0|\mathbf{w}|^p \quad \forall \mathbf{w} \in \mathbb{R}^d, \quad \text{so that} \quad \tilde\phi(\rho_0,\mathbf{z}) \leq c_0|\mathbf{z}|^q < +\infty \quad \forall \mathbf{z} \in \mathbb{R}^d. \tag{3.9}$$

Still applying Theorem 2.4, we obtain that $\rho \mapsto \tilde\phi(\rho,\mathbf{z})$ is finite, strictly positive and nondecreasing in the interval $(0,+\infty)$. Since $\tilde\phi(\rho,0) = 0$ we easily get

$$\tilde\phi(\rho,\mathbf{z}) \leq \tilde\phi(\rho_0,\mathbf{z}) \leq c_0|\mathbf{z}|^q \quad \forall \mathbf{z} \in \mathbb{R}^d,\ \rho \in (0,\rho_0); \tag{3.10}$$

$$\tilde\phi(\rho,\mathbf{z}) \leq \frac{\rho}{\rho_0}\tilde\phi(\rho_0,\mathbf{z}) \leq \frac{c_0}{\rho_0}\rho|\mathbf{z}|^q \quad \forall \mathbf{z} \in \mathbb{R}^d,\ \rho \in (\rho_0,+\infty). \tag{3.11}$$

Combining the last two bounds we get (3.6). (3.7) follows by homogeneity and by the fact that the continuous map ϕ has a maximum and a strictly positive minimum on the compact set $[\rho_0,\rho_1] \times \{\mathbf{w} \in \mathbb{R}^d : |\mathbf{w}| = 1\}$.

The final assertion concerning (3.8) still follows by Theorem 2.4. □

3.2 **Examples**

Example 3.2 Our main example is provided by the function

$$\phi_{2,\alpha}(\rho,\mathbf{w}) = \frac{|\mathbf{w}|^2}{\rho^\alpha}, \quad \tilde\phi_{2,\alpha}(\rho,\mathbf{z}) := \rho^\alpha|\mathbf{z}|^2, \qquad 0 \leq \alpha \leq 1. \tag{3.12}$$

Observe that $\phi_{2,\alpha}$ is positively θ-homogeneous, $\theta := 2-\alpha$, i.e.

$$\phi_{2,\alpha}(\lambda\rho,\lambda\mathbf{w}) = \lambda^\theta\phi(\rho,\mathbf{w}) \quad \forall \lambda,\rho > 0,\ \mathbf{w} \in \mathbb{R}^d. \tag{3.13}$$

It can be considered as a family of interpolating densities between the case $\alpha = 0$, when

$$\phi_{2,0}(\rho,\mathbf{w}) := |\mathbf{w}|^2, \tag{3.14}$$

and $\alpha = 1$, corresponding to the 1-homogeneous functional

$$\phi_{2,1}(\rho,\mathbf{w}) := \frac{|\mathbf{w}|^2}{\rho}. \tag{3.15}$$

Example 3.3 More generally, we introduce a concave function $h : (0,+\infty) \to (0,+\infty)$, which is a fortiori continuous and nondecreasing, and we consider the density function

$$\phi(\rho,\mathbf{w}) := \frac{|\mathbf{w}|^2}{h(\rho)}, \quad \tilde\phi(\rho,\mathbf{z}) := h(\rho)|\mathbf{w}|^2. \tag{3.16}$$

If h is of class C^2, we can express the concavity condition in terms of the function $g(\rho) := 1/h(\rho)$ as

$$h \text{ is concave} \quad \Leftrightarrow \quad g''(\rho)g(\rho) \geq 2\big(g'(\rho)\big)^2 \quad \forall \rho > 0, \tag{3.17}$$

which is related to a condition introduced in [6, Section 2.2, (2.12c)] to study entropy functionals.

Example 3.4 We consider matrix-valued functions $\mathsf{H},\mathsf{G} : (0,+\infty) \to \mathbb{M}^{d\times d}$ such that

$$\mathsf{H}(\rho),\mathsf{G}(\rho) \quad \text{are symmetric and positive definite, } \mathsf{H}(\rho) = \mathsf{G}^{-1}(\rho) \quad \forall \rho > 0. \tag{3.18}$$

They induce the action density $\phi : (0,+\infty)\times\mathbb{R}^d \to [0,+\infty)$ defined as

$$\phi(\rho,\mathbf{w}) := \langle \mathsf{G}(\rho)\mathbf{w},\mathbf{w}\rangle = \left\langle \mathsf{H}^{-1}(\rho)\mathbf{w},\mathbf{w}\right\rangle. \tag{3.19}$$

Taking into account Theorem 3.1, ϕ satisfies conditions (3.2) if and only if the maps

$$\rho \mapsto \langle \mathsf{H}(\rho)\mathbf{w},\mathbf{w}\rangle \quad \text{are concave in } (0,+\infty) \quad \forall \mathbf{w}\in\mathbb{R}^d. \tag{3.20}$$

Equivalently,

$$\mathsf{H}((1-\vartheta)\rho_0+\vartheta\rho_1) \ge (1-\vartheta)\mathsf{H}(\rho_0)+\vartheta\mathsf{H}(\rho_1) \quad \text{as quadratic forms.} \tag{3.21}$$

When G is of class C^2 this is also equivalent to ask that

$$\mathsf{G}''(\rho) \ge 2\mathsf{G}'(\rho)\mathsf{H}(\rho)\mathsf{G}'(\rho) \qquad \forall \rho > 0, \tag{3.22}$$

in the sense of the associated quadratic forms. In fact, differentiating $\mathsf{H} = \mathsf{G}^{-1}$ with respect to ρ we get

$$\mathsf{H}' = -\mathsf{H}\,\mathsf{G}'\,\mathsf{H}, \qquad \mathsf{H}'' = -\mathsf{H}\,\mathsf{G}''\,\mathsf{H}+2\mathsf{H}\,\mathsf{G}'\,\mathsf{H}\,\mathsf{G}'\,\mathsf{H},$$

so that

$$\frac{\mathrm{d}^2}{\mathrm{d}^2\rho}\langle \mathsf{H}(\rho)\mathbf{w},\mathbf{w}\rangle = -\left\langle \mathsf{G}''\tilde{\mathbf{w}},\tilde{\mathbf{w}}\right\rangle + 2\left\langle \mathsf{G}'\mathsf{H}\mathsf{G}'\tilde{\mathbf{w}},\tilde{\mathbf{w}}\right\rangle \quad \text{where} \quad \tilde{\mathbf{w}} := \mathsf{H}\mathbf{w};$$

we eventually recall that $\mathsf{H}(\rho)$ is invertible for every $\rho > 0$.

Example 3.5 Let $\|\cdot\|$ be any norm in $\mathbb{R}^d$ with dual norm $\|\cdot\|_*$, and let $h : (0,+\infty)\to(0,+\infty)$ be a concave (continuous, nondecreasing) function as in Example 3.3. We can thus consider

$$\phi(\rho,\mathbf{w}) := h(\rho)\left\|\frac{\mathbf{w}}{h(\rho)}\right\|^p, \qquad \tilde{\phi}(\rho,\mathbf{z}) := h(\rho)\|\mathbf{z}\|_*^q. \tag{3.23}$$

See [20,21] for a in-depth study of this class of functions.

Example 3.6 ($(\alpha$-$\theta)$-homogeneous functionals) In the particular case $h(\rho) := \rho^\alpha$ the functional ϕ of the previous example is jointly positively θ-homogeneous, with $\theta := \alpha+(1-\alpha)p$. This is in fact the most general example of θ-homogeneous functional, since if ϕ is θ-positively homogeneous, $1\le\theta\le p$, then

$$\phi(\rho,\mathbf{w}) = \rho^\theta\phi(1,\mathbf{w}/\rho) = \rho^{\theta-p}\phi(1,\mathbf{w}) = \rho^\alpha\|\mathbf{w}/\rho^\alpha\|^p, \quad \alpha = \frac{p-\theta}{p-1}, \tag{3.24}$$

where $\|\mathbf{w}\| := \phi(1,\mathbf{w})^{1/p}$ is a norm in $\mathbb{R}^d$ by Theorem 3.1. The dual marginal density $\tilde{\phi}$ in this case takes the form

$$\tilde{\phi}(\rho,\mathbf{z}) = \rho^\alpha\|\mathbf{z}\|_*^q \quad \forall \rho>0,\ \mathbf{z}\in\mathbb{R}^d, \tag{3.25}$$

and it is $q+\alpha$-homogeneous. Notice that α and θ are related by

$$\frac{\theta}{p}+\frac{\alpha}{q} = 1. \tag{3.26}$$

In the particular case when $\|\cdot\| = \|\cdot\|_* = |\cdot|$ is the Euclidean norm, we set as in (3.16)

$$\phi_{p,\alpha}(\rho,\mathbf{w}) := \rho^\alpha\left|\frac{\mathbf{w}}{\rho^\alpha}\right|^p, \quad \tilde{\phi}_{q,\alpha}(\rho,\mathbf{z}) := \rho^\alpha|\mathbf{z}|^q, \qquad 0\le\alpha\le 1. \tag{3.27}$$

3.3 **The action functional on measures**

Lower semicontinuity envelope and recession function. Thanks to the monotonicity property (3.5), we can extend ϕ also for $\rho = 0$ by setting for every $\mathbf{w} \in \mathbb{R}^d$

$$\phi(0,\mathbf{w}) = \sup_{\rho>0} \phi(\rho,\mathbf{w}) = \lim_{\rho\downarrow 0} \phi(\rho,\mathbf{w}); \quad \text{in particular } \begin{cases} \phi(0,\mathbf{0}) = 0, \\ \phi(0,\mathbf{w}) > 0 \quad \text{if } \mathbf{w} \neq \mathbf{0}. \end{cases} \tag{3.28}$$

When $\rho < 0$ we simply set $\phi(\rho,\mathbf{w}) = +\infty$, observing that this extension is lower semicontinuous in $\mathbb{R}\times\mathbb{R}^d$. It is not difficult to check that $\tilde{\phi}(0,\cdot)$ satisfies an analogous formula

$$\tilde{\phi}(0,\mathbf{z}) = \sup_{\mathbf{w}\in\mathbb{R}^d} \mathbf{z}\cdot\mathbf{w} - \phi(0,\mathbf{w}) = \inf_{\rho>0} \tilde{\phi}(\rho,\mathbf{z}) = \lim_{\rho\downarrow 0} \tilde{\phi}(\rho,\mathbf{z}) \qquad \forall \mathbf{z} \in \mathbb{R}^d. \tag{3.29}$$

Observe that, as in the $(\alpha$-$\theta)$-homogeneous case of Example 3.6 with $\alpha > 0$,

$$\tilde{\phi}(0,\mathbf{z}) \equiv 0 \quad \Rightarrow \quad \phi(0,\mathbf{w}) = \begin{cases} +\infty & \text{if } \mathbf{w} \neq 0 \\ 0 & \text{if } \mathbf{w} = 0. \end{cases} \tag{3.30}$$

As in (2.4), we also introduce the *recession functional*

$$\phi^\infty(\rho,\mathbf{w}) = \sup_{\lambda>0} \frac{1}{\lambda}\phi(\lambda\rho,\lambda\mathbf{w}) = \lim_{\lambda\uparrow+\infty} \frac{1}{\lambda}\phi(\lambda\rho,\lambda\mathbf{w}) = \lim_{\lambda\uparrow+\infty} \lambda^{p-1}\phi(\lambda\rho,\mathbf{w}). \tag{3.31}$$

ϕ^∞ is still convex, p-homogeneous w.r.t. $\mathbf{w}$, and l.s.c. with values in $[0,+\infty]$; moreover, it is 1-homogeneous so that it can be expressed as

$$\phi^\infty(\rho,\mathbf{w}) = \begin{cases} \frac{\varphi^\infty(\mathbf{w})}{\rho^{p-1}} = \rho\,\varphi^\infty(\mathbf{w}/\rho) & \text{if } \rho \neq 0, \\ +\infty & \text{if } \rho = 0 \text{ and } \mathbf{w} \neq 0, \end{cases} \tag{3.32}$$

where $\varphi^\infty : \mathbb{R}^d \to [0,+\infty]$ is a convex and p-homogeneous function which is non degenerate, i.e. $\varphi^\infty(\mathbf{w}) > 0$ if $\mathbf{w} \neq 0$. φ^∞ admits a dual representation, based on

$$\tilde{\varphi}^\infty(\mathbf{z}) := \inf_{\lambda>0} \frac{1}{\lambda}\tilde{\phi}(\lambda,\mathbf{z}) = \lim_{\lambda\uparrow+\infty} \frac{1}{\lambda}\tilde{\phi}(\lambda\rho,\mathbf{z}). \tag{3.33}$$

$\tilde{\varphi}^\infty$ is finite, convex, nonnegative, and q-homogeneous, so that $\tilde{\varphi}^\infty(\mathbf{z})^{1/q}$ is a seminorm, which does not vanish at $\mathbf{z} \in \mathbb{R}^d$ if and only if $\rho \mapsto \tilde{\phi}(\rho,\mathbf{z})$ has a linear growth when $\rho \uparrow +\infty$. It is easy to check that

$$\varphi^\infty(\mathbf{w})^{1/p} = \sup\Big\{\mathbf{w}\cdot\mathbf{z} : \tilde{\varphi}^\infty(\mathbf{z}) \le 1\Big\}. \tag{3.34}$$

In the case $\tilde{\phi}$ has a sublinear growth w.r.t. ρ, as for $(\alpha$-$\theta)$-homogeneous functionals with $\alpha < 1$ (see Example 3.6), we have in particular

$$\text{when } \tilde{\varphi}^\infty(\mathbf{z}) \equiv 0, \qquad \varphi^\infty(\mathbf{w}) = \begin{cases} +\infty & \text{if } \mathbf{w} \neq 0, \\ 0 & \text{if } \mathbf{w} = 0. \end{cases} \tag{3.35}$$

The action functional. Let $\gamma,\mu\in\mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ be nonnegative Radon measures and let $\boldsymbol{\nu}\in\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)$ be a vector Radon measure on $\mathbb{R}^d$. We assume that $\mathrm{supp}(\mu),\mathrm{supp}(\boldsymbol{\nu})\subset\mathrm{supp}(\gamma)$, and we write their Lebesgue decomposition with respect to the reference measure γ

$$\mu:=\rho\gamma+\mu^\perp,\quad \boldsymbol{\nu}:=\mathbf{w}\gamma+\boldsymbol{\nu}^\perp. \tag{3.36}$$

We can always introduce a nonnegative Radon measure $\sigma\in\mathcal{M}^+(\overline{\Omega})$ such that $\mu^\perp=\rho^\perp\sigma\ll\sigma,\boldsymbol{\nu}^\perp=\mathbf{w}^\perp\sigma\ll\sigma$, e.g. $\sigma:=\mu^\perp+|\boldsymbol{\nu}^\perp|$. We can thus define the *action functional*

$$\Phi(\mu,\boldsymbol{\nu}|\gamma)=\Phi^a(\mu,\boldsymbol{\nu}|\gamma)+\Phi^\infty(\mu,\boldsymbol{\nu}|\gamma):=\int_{\mathbb{R}^d}\phi(\rho,\mathbf{w})\,\mathrm{d}\gamma+\int_{\mathbb{R}^d}\phi^\infty(\rho^\perp,\mathbf{w}^\perp)\,\mathrm{d}\sigma. \tag{3.37}$$

Observe that, being ϕ^∞ 1-homogeneous, this definition is independent of σ. We will also use a localized version of Φ: if $B\in\mathcal{B}(\mathbb{R}^d)$ we set

$$\Phi(\mu,\boldsymbol{\nu}|\gamma,B):=\int_B\phi(\rho,\mathbf{w})\,\mathrm{d}\gamma+\int_B\phi^\infty(\rho^\perp,\mathbf{w}^\perp)\,\mathrm{d}\sigma. \tag{3.38}$$

Lemma 3.7 *Let $\mu=\rho\gamma+\mu^\perp,\boldsymbol{\nu}=\mathbf{w}\gamma+\boldsymbol{\nu}^\perp$ be such that $\Phi(\mu,\boldsymbol{\nu}|\gamma)$ is finite. Then $\boldsymbol{\nu}^\perp=\mathbf{w}^\perp\mu^\perp\ll\mu^\perp$ and*

$$\Phi^\infty(\mu,\boldsymbol{\nu}|\gamma)=\int_{\mathbb{R}^d}\varphi^\infty(\mathbf{w}^\perp)\,\mathrm{d}\mu^\perp,\ \Phi(\mu,\boldsymbol{\nu}|\gamma)=\int_{\mathbb{R}^d}\phi(\rho,\mathbf{w})\,\mathrm{d}\gamma+\int_{\mathbb{R}^d}\varphi^\infty(\mathbf{w}^\perp)\,\mathrm{d}\mu^\perp. \tag{3.39}$$

Moreover, if $\tilde\phi$ has a sublinear growth with respect to ρ *(e.g. in the $(\alpha$-$\theta)$-homogeneous case of Example 3.6, with $\alpha<1$) then $\tilde\varphi^\infty(\cdot)\equiv 0$ and*

$$\Phi(\mu,\boldsymbol{\nu})<+\infty\ \Rightarrow\ \boldsymbol{\nu}=\mathbf{w}\cdot\gamma\ll\gamma,\quad \Phi(\mu,\boldsymbol{\nu})=\Phi^a(\mu,\boldsymbol{\nu})=\int_{\mathbb{R}^d}\phi(\rho,\mathbf{w})\,\mathrm{d}\gamma, \tag{3.40}$$

independently on the singular part $\mu^\perp$.

Proof Let $\tilde\sigma\in\mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ any measure such that $\mu^\perp\ll\tilde\sigma,|\boldsymbol{\nu}^\perp|\ll\tilde\sigma$ so that $\Phi^\infty(\mu,\boldsymbol{\nu}|\gamma)$ can be represented as

$$\Phi^\infty(\mu,\boldsymbol{\nu}|\gamma)=\int_{\mathbb{R}^d}\phi^\infty(\tilde\rho^\perp,\tilde{\mathbf{w}}^\perp)\,\mathrm{d}\tilde\sigma,\quad \tilde\rho^\perp=\frac{\mathrm{d}\mu^\perp}{\mathrm{d}\tilde\sigma},\ \tilde{\mathbf{w}}^\perp=\frac{\mathrm{d}\boldsymbol{\nu}^\perp}{\mathrm{d}\tilde\sigma}.$$

When $\Phi^\infty(\mu,\boldsymbol{\nu}|\gamma)<+\infty$, (3.32) yields $\tilde{\mathbf{w}}^\perp(x)=0$ for $\tilde\sigma$-a.e. x such that $\tilde\rho^\perp(x)=0$. It follows that

$$\Phi(\mu,\boldsymbol{\nu})<+\infty\quad\Rightarrow\quad\boldsymbol{\nu}^\perp\ll\mu^\perp, \tag{3.41}$$

so that one can always choose $\tilde\sigma=\mu^\perp$, $\tilde\rho^\perp=1$, and decompose $\boldsymbol{\nu}^\perp$ as $\mathbf{w}^\perp\mu^\perp$ obtaining (3.39). (3.40) is then an immediate consequence of (3.35). □

Remark 3.8 When $\tilde\phi(0,\mathbf{z})\equiv 0$ (e.g. in the $(\alpha$-$\theta)$-homogeneous case of Example 3.6, with $\alpha>0$) the density $\mathbf{w}$ of $\boldsymbol{\nu}$ w.r.t. γ vanishes if ρ vanishes, i.e.

$$\Phi(\mu,\boldsymbol{\nu}|\gamma)<+\infty\quad\Rightarrow\quad\mathbf{w}(x)=0\text{ if }\rho(x)=0,\text{ for }\gamma\text{-a.e. }x\in\mathbb{R}^d. \tag{3.42}$$

In particular $\boldsymbol{\nu}^a$ is absolutely continuous also with respect to μ.

Applying Theorem 2.1 we immediately get

Lemma 3.9 (Lower semicontinuity and approximation of the action functional) *The action functional is lower semicontinuous with respect to weak* convergence of measures, i.e. if*

$$\mu_n\rightharpoonup^*\mu,\ \gamma_n\rightharpoonup^*\gamma\quad\textit{weakly}^*\textit{ in }\mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d),\qquad\boldsymbol{\nu}_n\rightharpoonup^*\boldsymbol{\nu}\quad\textit{in }\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)\textit{ as }n\uparrow+\infty,$$

then

$$\liminf_{n\uparrow\infty}\Phi(\mu_n,\boldsymbol{\nu}_n|\gamma_n)\ge\Phi(\mu,\boldsymbol{\nu}|\gamma).$$

Equiintegrability estimate. We collect in this section some basic estimates on ϕ which will turn to be useful in the sequel. Let us first introduce the notation

$$\|\mathbf{z}\|_* := \tilde{\phi}(1,\mathbf{z})^{1/q}, \quad \|\mathbf{w}\| := \phi(1,\mathbf{w})^{1/p}, \quad \eta^{-1}|\mathbf{z}| \le \|\mathbf{z}\|_* \le \eta|\mathbf{z}|, \tag{3.43}$$

$$\Gamma_\phi := \Big\{(a,b) : \sup_{\|\mathbf{z}\|_*=1} \tilde{\phi}(\rho,\mathbf{z}) \le a+b\rho\Big\}, \quad \mathsf{h}(\rho) := \inf\Big\{a+b\rho : (a,b)\in\Gamma_\phi\Big\}, \tag{3.44}$$

$$H(s,\rho) := s\mathsf{h}(\rho/s) = \inf\Big\{as+b\rho : (a,b)\in\Gamma_\phi\Big\}. \tag{3.45}$$

Observe that h is a concave increasing function defined in $[0,+\infty)$, satisfying, in the homogeneous case $\mathsf{h}(\rho) = h(\rho) = \rho^\alpha$. It provides the bounds

$$\begin{aligned} &\tilde{\phi}(\rho,\mathbf{z}) \le \mathsf{h}(\rho)\|\mathbf{z}\|_*^q, \qquad \|\mathbf{w}\| \le \mathsf{h}(\rho)^{1/q}\phi(\rho,\mathbf{w})^{1/p},\\ &\tilde{\varphi}^\infty(\mathbf{z}) \le \mathsf{h}^\infty\|\mathbf{z}\|_*^q, \qquad \|\mathbf{w}\| \le \big(\mathsf{h}^\infty\big)^{1/q}\varphi^\infty(\mathbf{w})^{1/p}, \quad \text{if } \mathsf{h}^\infty := \lim_{\lambda\uparrow+\infty}\lambda^{-1}\mathsf{h}(\lambda) > 0. \end{aligned} \tag{3.46}$$

Observe that when $\mathsf{h}^\infty = 0$ then $\tilde{\varphi}^\infty \equiv 0$ and $\varphi^\infty(\mathbf{w})$ is given by (3.35).

Proposition 3.10 (Integrability estimates) *Let ζ be a nonnegative Borel function such that*

$$\mu(\zeta^q) := \int_{\mathbb{R}^d}\zeta^q\,\mathrm{d}\mu \quad \textit{and} \quad \gamma(\zeta^q) := \int_{\mathbb{R}^d}\zeta^q\,\mathrm{d}\gamma \quad \textit{are finite},$$

and let $Z := \big\{x\in\mathbb{R}^d : \zeta(x) > 0\big\}$. If $\Phi(\mu,\boldsymbol{\nu}|\gamma) < +\infty$ we have

$$\int_{\mathbb{R}^d}\zeta(x)\,\mathrm{d}\|\boldsymbol{\nu}\|(x) \le \Phi^{1/p}\big(\mu,\boldsymbol{\nu}|\gamma,Z\big)\,H^{1/q}\big(\gamma(\zeta^q),\mu(\zeta^q)\big). \tag{3.47}$$

In particular, for every Borel set $A\in\mathcal{B}(\mathbb{R}^d)$ we have

$$\|\boldsymbol{\nu}\|(A) \le \Phi^{1/p}\big(\mu,\boldsymbol{\nu}|\gamma,A\big)\,H^{1/q}\big(\gamma(A),\mu(A)\big). \tag{3.48}$$

Proof It is sufficient to prove (3.47). Observe that if $(a,b)\in\Gamma_\phi$ then $a\ge 0$, and $\mathsf{h}^\infty \le b$ so that by (3.46) we have

$$\begin{aligned} &\int_{\mathbb{R}^d}\zeta(x)\,\mathrm{d}\|\boldsymbol{\nu}\|(x) \le \int_Z \zeta\|\mathbf{w}\|\,\mathrm{d}\gamma + \int_Z\zeta\|\mathbf{w}^\perp\|\,\mathrm{d}\mu^\perp\\ &\le \Big(\int_Z\phi(\rho,\mathbf{w})\,\mathrm{d}\gamma\Big)^{1/p}\Big(\int_Z\zeta^q\mathsf{h}(\rho)\,\mathrm{d}\gamma\Big)^{1/q} + \Big(\int_Z\varphi^\infty(\mathbf{w}^\perp)\,\mathrm{d}\mu^\perp\Big)^{1/p}\Big(\mathsf{h}^\infty\!\int_Z\zeta^q\,\mathrm{d}\mu^\perp\Big)^{1/q}\\ &\le \Big(\Phi(\mu,\boldsymbol{\nu}|\gamma,Z)\Big)^{1/p}\Big(a\int_{\mathbb{R}^d}\zeta^q\,\mathrm{d}\gamma + b\int_{\mathbb{R}^d}\zeta^q\,\mathrm{d}\mu\Big)^{1/q}, \end{aligned}$$

Taking the infimum of the last term over all the couples $(a,b)\in\Gamma_\phi$ we obtain (3.48). □

4 Measure valued solutions of the continuity equation in $\mathbb{R}^d$

In this section we collect some results on the continuity equation

$$\partial_t \mu_t + \nabla \cdot \boldsymbol{v}_t = 0 \qquad \text{in } \mathbb{R}^d \times (0,T), \tag{4.1}$$

which we will need in the sequel. Here $\mu_t, \boldsymbol{v}_t$ are Borel families of measures (see e.g. [3]) in $\mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ and $\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)$ respectively, defined for t in the open interval $(0,T)$, such that

$$\int_0^T \mu_t(B_R)\,\mathrm{d}t < +\infty, \qquad V_R := \int_0^T |\boldsymbol{v}_t|(B_R)\,\mathrm{d}t < +\infty \qquad \forall R > 0, \tag{4.2}$$

and we suppose that (4.1) holds in the sense of distributions, i.e.

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \zeta(x,t)\,\mathrm{d}\mu_t(x)\,\mathrm{d}t + \int_0^T \int_{\mathbb{R}^d} \nabla_x \zeta(x,t) \cdot \mathrm{d}\boldsymbol{v}_t(x)\,\mathrm{d}t = 0 \tag{4.3}$$

for every $\zeta \in C^1_c(\mathbb{R}^d \times (0,T))$. Thanks to the disintegration theorem [14, 4, III-70], we can identify $(\boldsymbol{v}_t)_{t\in(0,T)}$ with the measure $\boldsymbol{v} = \int_0^T \boldsymbol{v}_t\,\mathrm{d}t \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d \times (0,T);\mathbb{R}^d)$ defined by the formula

$$\langle \boldsymbol{v}, \boldsymbol{\zeta} \rangle = \int_0^T \left(\int_{\mathbb{R}^d} \boldsymbol{\zeta}(x,t) \cdot \mathrm{d}\boldsymbol{v}_t(x) \right) \mathrm{d}t \qquad \forall \boldsymbol{\zeta} \in C^0_{\mathrm{c}}(\mathbb{R}^d \times (0,T);\mathbb{R}^d). \tag{4.4}$$

4.1 **Preliminaries**

Let us first adapt the results of [3, Chap. 8] (concerning a family of *probability* measures μ_t) to the more general case of Radon measures. First of all we recall some (technical) preliminaries.

Lemma 4.1 (Continuous representative) *Let $\mu_t, \boldsymbol{v}_t$ be Borel families of measures satisfying* (4.2) *and* (4.3). *Then there exists a unique* weakly* continuous *curve $t \in [0,T] \mapsto \tilde\mu_t \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ such that $\mu_t = \tilde\mu_t$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$; if $\zeta \in C^1_c(\mathbb{R}^d \times [0,T])$ and $t_1 \le t_2 \in [0,T]$, we have*

$$\int_{\mathbb{R}^d} \zeta_{t_2}\,\mathrm{d}\tilde\mu_{t_2} - \int_{\mathbb{R}^d} \zeta_{t_1}\,\mathrm{d}\tilde\mu_{t_1} = \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \partial_t \zeta\,\mathrm{d}\mu_t(x)\,\mathrm{d}t + \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \nabla \zeta \cdot \mathrm{d}\boldsymbol{v}_t(x)\,\mathrm{d}t, \tag{4.5}$$

and the mass of $\tilde\mu_t$ can be uniformly bounded by

$$\sup_{t\in[0,T]} \tilde\mu_t(B_R) \le \tilde\mu_s(B_{2R}) + 2R^{-1}V_{2R} \qquad \forall s \in [0,T]. \tag{4.6}$$

Moreover, if $\tilde\mu_s(\mathbb{R}^d) < +\infty$ for some $s \in [0,T]$ and $\lim_{R\uparrow+\infty} R^{-1}V_R = 0$, then the total mass $\tilde\mu_t(\mathbb{R}^d)$ is (finite and) constant.

Proof Let us take $\zeta(x,t) = \eta(t)\zeta(x)$, $\eta \in C^\infty_c(0,T)$ and $\zeta \in C^\infty_c(\mathbb{R}^d)$ with $\operatorname{supp}\zeta \subset B_R$; we have

$$-\int_0^T \eta'(t) \Big(\int_{\mathbb{R}^d} \zeta(x)\,\mathrm{d}\mu_t(x) \Big)\,\mathrm{d}t = \int_0^T \eta(t) \Big(\int_{\mathbb{R}^d} \nabla\zeta(x) \cdot \mathrm{d}\boldsymbol{v}_t(x) \Big)\,\mathrm{d}t,$$

so that the map $t \mapsto \mu_t(\zeta) = \int_{\mathbb{R}^d} \zeta\,\mathrm{d}\mu_t$ belongs to $W^{1,1}(0,T)$ with distributional derivative

$$\dot\mu_t(\zeta) = \int_{\mathbb{R}^d} \nabla\zeta(x) \cdot \mathrm{d}\boldsymbol{v}_t(x) \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T), \tag{4.7}$$

satisfying

$$|\dot{\mu}_t(\zeta)| \le V_R(t) \sup_{\mathbb{R}^d} |\nabla\zeta|, \quad V_R(t) := |\boldsymbol{\nu}_t|(B_R), \quad \int_0^T V_R(t)\,\mathrm{d}t = V_R < +\infty. \tag{4.8}$$

If L_ζ is the set of its Lebesgue points, we know that $\mathscr{L}^1((0,T)\setminus L_\zeta) = 0$. Let us now take an increasing sequence $R_n := 2^n \uparrow +\infty$ and countable sets $Z_n \subset C^\infty_{\rm c}(B_{R_n})$ which are dense in $C^1_0(B_{R_n}) := \{\zeta \in C^1(\mathbb{R}^d) : \mathrm{supp}(\zeta) \subset \overline{B_{R_n}}\}$, the closure of $C^1_{\rm c}(B_{R_n})$ with respect the usual C^1 norm $\|\zeta\|_{C^1} = \sup_{\mathbb{R}^d}(|\zeta|,|\nabla\zeta|)$. We also set $L_Z := \cap_{n\in\mathbb{N},\zeta\in Z_n} L_\zeta$. The restriction of the curve μ to L_Z provides a uniformly continuous family of functionals on each space $C^1_0(B_{R_n})$, since (4.8) shows

$$|\mu_t(\zeta) - \mu_s(\zeta)| \le \|\zeta\|_{C^1} \int_s^t V_{R_n}(\lambda)\,\mathrm{d}\lambda \quad \forall s,t \in L_Z \quad \forall\, \zeta \in Z_n.$$

Therefore, for every $n \in \mathbb{N}$ it can be extended in a unique way to a continuous curve $\{\tilde{\mu}^n_t\}_{t\in[0,T]}$ in $[C^1_0(B_{R_n})]'$ which is uniformly bounded and satisfies the compatibility condition

$$\tilde{\mu}^m_t(\zeta) = \tilde{\mu}^n(\zeta) \quad \text{if } m \le n \text{ and } \zeta \in C^1_{\rm c}(B_{R_m}). \tag{4.9}$$

If $\zeta \in C^1_{\rm c}(\mathbb{R}^d)$ we can thus define

$$\tilde{\mu}_t(\zeta) := \tilde{\mu}^n_t(\zeta) \quad \text{for every } n \in \mathbb{N} \text{ such that } \mathrm{supp}(\zeta) \subset B_{R_n}. \tag{4.10}$$

If we show that $\{\mu_t(B_{R_n})\}_{t\in L_Z}$ is uniformly bounded for every $n \in \mathbb{N}$, the extension provides a continuous curve in $\mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$. To this aim, let us consider nonnegative, smooth functions

$$\zeta_k : \mathbb{R}^d \to [0,1], \text{ such that} \quad \zeta_k(x) := \zeta_0(x/2^k), \tag{4.11a}$$

$$\zeta_k(x) = 1 \text{ if } |x| \le 2^k, \quad \zeta_k(x) = 0 \text{ if } |x| \ge 2^{k+1}, \quad |\nabla\zeta_k(x)| \le A\,2^{-k}, \tag{4.11b}$$

for some constant $A > 1$. It is not restrictive to suppose that $\zeta_k \in Z_{k+1}$. Applying the previous formula (4.7), for $t,\, s \in L_Z$ we have

$$|\mu_t(\zeta_k) - \mu_s(\zeta_k)| \le a_k := 2^{1-k} \int_0^T |\boldsymbol{\nu}_r|\big(B_{2R_k} \setminus B_{R_k}\big)\,\mathrm{d}r \le A\,2^{-k} V_{2R_k}. \tag{4.12}$$

It follows that

$$\mu_t(B_{R_k}) \le \mu_t(\zeta_k) \le \mu_s(\zeta_k) + A\,2^{-k} V_{2R_k} \le \mu_s(B_{2R_k}) + A\,2^{-k} V_{2R_k} \quad \forall t \in L_Z. \tag{4.13}$$

Integrating with respect to s we end up with the uniform bound

$$\mu_t(B_{R_k}) \le A\,2^{-k} V_{R_{k+1}} + \int_0^T \mu_s(B_{2R_k})\,\mathrm{d}s < +\infty \quad \forall t \in L_Z.$$

Observe that the extension $\tilde{\mu}_t$ satisfies (4.13) (and therefore, in a completely analogous way, (4.6)) and (4.12) for every $s,t \in [0,T]$.

Now we show (4.5). Let us choose $\zeta \in C^1_c(\mathbb{R}^d \times [0,T])$ and $\eta_\varepsilon \in C^\infty_c(t_1,t_2)$ such that

$$0 \le \eta_\varepsilon(t) \le 1, \quad \lim_{\varepsilon\downarrow 0} \eta_\varepsilon(t) = \chi_{(t_1,t_2)}(t) \quad \forall t \in [0,T], \quad \lim_{\varepsilon\downarrow 0} \eta'_\varepsilon = \delta_{t_1} - \delta_{t_2}$$

in the duality with continuous functions in $[0,T]$. We get

$$\begin{aligned} 0 &= \int_0^T \int_{\mathbb{R}^d} \partial_t(\eta_\varepsilon \zeta)\, \mathrm{d}\mu_t(x)\, \mathrm{d}t + \int_0^T \int_{\mathbb{R}^d} \nabla_x(\eta_\varepsilon \zeta)\cdot \mathrm{d}\boldsymbol{\nu}_t\, \mathrm{d}t \\ &= \int_0^T \eta_\varepsilon(t) \int_{\mathbb{R}^d} \partial_t \zeta\, \mathrm{d}\mu_t\, \mathrm{d}t + \int_0^T \eta_\varepsilon(t) \int_{\mathbb{R}^d} \nabla_x \zeta \cdot \mathrm{d}\boldsymbol{\nu}_t\, \mathrm{d}t + \int_0^T \eta_\varepsilon'(t) \int_{\mathbb{R}^d} \zeta\, \mathrm{d}\tilde{\mu}_t\, \mathrm{d}t. \end{aligned}$$

Passing to the limit as ε vanishes and invoking the continuity of $\tilde{\mu}_t$, we get (4.5).

Finally, if $\lim_{R\uparrow+\infty} R^{-1}V_R = 0$ we can pass to the limit as $R_k \uparrow +\infty$ in the inequality (4.12), which also holds for every $t,s\in[0,T]$ if we replace μ by $\tilde{\mu}$, by choosing s so that

$$m := \tilde{\mu}_s(\mathbb{R}^d) = \lim_{k\uparrow+\infty} \tilde{\mu}_s(\zeta_k) < +\infty.$$

It follows that $\tilde{\mu}_t(\mathbb{R}^d) = \lim_{k\uparrow+\infty} \tilde{\mu}_s(\zeta_k) = m$ for every $t\in[0,T]$. □

Thanks to Lemma 4.1 we can introduce the following class of solutions of the continuity equation.

Definition 4.2 (Solutions of the continuity equation) We denote by $\mathcal{CE}(0,T)$ the set of time dependent measures $(\mu_t)_{t\in[0,T]}, (\boldsymbol{\nu}_t)_{t\in(0,T)}$ such that

1. $t\mapsto\mu_t$ is weakly* continuous in $\mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ (in particular, $\sup_{t\in[0,T]}\mu_t(B_R) < +\infty$ for every $R>0$),
2. $(\boldsymbol{\nu}_t)_{t\in(0,T)}$ is a Borel family with $\int_0^T |\boldsymbol{\nu}_t|(B_R)\,\mathrm{d}t < +\infty \qquad \forall R>0$;
3. $(\mu,\boldsymbol{\nu})$ is a distributional solution of (4.1).

$\mathcal{CE}(0,T;\sigma\to\eta)$ denotes the subset of $(\mu,\boldsymbol{\nu})\in\mathcal{CE}(0,T)$ such that $\mu_0=\sigma,\ \mu_1=\eta$.

Solutions of the continuity equation can be rescaled in time:

Lemma 4.3 (Time rescaling) *Let* $\mathsf{t}: s\in[0,T']\to\mathsf{t}(s)\in[0,T]$ *be a strictly increasing absolutely continuous map with absolutely continuous inverse* $\mathsf{s}:=\mathsf{t}^{-1}$. *Then* $(\mu,\boldsymbol{\nu})$ *is a distributional solution of* (4.1) *if and only if*

$$\hat{\mu} := \mu\circ\mathsf{t},\ \hat{\boldsymbol{\nu}} := \mathsf{t}'\big(\boldsymbol{\nu}\circ\mathsf{t}\big), \quad \textit{is a distributional solution of } (4.1) \textit{ on } (0,T').$$

We refer to [3, Lemma 8.1.3] for the proof.

The proof of the next lemma follows directly from (4.5).

Lemma 4.4 (Glueing solutions) *Let* $(\mu^i,\boldsymbol{\nu}^i)\in\mathcal{CE}(0,T_i)$, $i=1,2$, *with* $\mu^1_{T_1}=\mu^2_0$. *Then the new family* $(\mu_t,\boldsymbol{\nu}_t)_{t\in(0,T_1+T_2)}$ *defined as*

$$\mu_t := \begin{cases} \mu^1_t & \text{if } 0\le t\le T_1 \\ \mu^2_{t-T_1} & \text{if } T_1\le t\le T_1+T_2 \end{cases} \qquad \boldsymbol{\nu}_t := \begin{cases} \boldsymbol{\nu}^1_t & \text{if } 0\le t\le T_1 \\ \boldsymbol{\nu}^2_{t-T_1} & \text{if } T_1\le t\le T_1+T_2 \end{cases} \tag{4.14}$$

belongs to $\mathcal{CE}(0,T_1+T_2)$.

Lemma 4.5 (Compactness for solutions of the continuity equation (I)) *Let* $(\mu^n,\boldsymbol{\nu}^n)$ *be a sequence in* $\mathcal{CE}(0,T)$ *such that*

1. *for some* $s\in[0,T]$ $\quad\sup_{n\in\mathbb{N}}\mu^n_s(B_R) < +\infty \quad \forall R>0$;
2. *the sequence of maps* $t\mapsto|\boldsymbol{\nu}^n_t|(B_R)$ *is equiintegrable in* $(0,T)$, *for every* $R>0$.

Then there exists a subsequence (still indexed by n) and a couple $(\mu_t, \boldsymbol{\nu}_t) \in \mathcal{CE}(0,T)$ *such that (recall* (4.4)*)*

$$\begin{aligned} \mu_t^n \rightharpoonup^* \mu_t \quad & \textit{weakly}^* \textit{ in } \mathcal{M}^+_{\rm loc}(\mathbb{R}^d) \quad \forall t \in [0,T], \\ \boldsymbol{\nu}^n \rightharpoonup^* \boldsymbol{\nu} \quad & \textit{weakly}^* \textit{ in } \mathcal{M}_{\rm loc}(\mathbb{R}^d \times (0,T); \mathbb{R}^d). \end{aligned} \tag{4.15}$$

(4.15) *yields in particular*

$$\int_0^T \Phi(\mu_t, \boldsymbol{\nu}_t | \gamma)\,\mathrm{d}t \le \liminf_{n\uparrow+\infty} \int_0^T \Phi(\mu_t^n, \boldsymbol{\nu}_t^n | \gamma^n)\,\mathrm{d}t \tag{4.16}$$

for every sequence of Radon measures $\gamma^n \rightharpoonup^* \gamma$ *in* $\mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$*, where* Φ *is an integral functional as in* (3.37).

Proof Since $\boldsymbol{\nu}^n := \int_0^T \boldsymbol{\nu}_t^n\,\mathrm{d}t$ and μ_s^n have total variation uniformly bounded on each compact subset of $\mathbb{R}^d \times [0,T]$, we can extract a subsequence (still denoted by $\mu_s^n, \boldsymbol{\nu}^n$) such that $\mu_s^n \rightharpoonup^* \mu_s$ in $\mathcal{M}_{\rm loc}(\mathbb{R}^d)$ and $\boldsymbol{\nu}^n \rightharpoonup^* \boldsymbol{\nu}$ in $\mathcal{M}_{\rm loc}(\mathbb{R}^d \times [0,T]; \mathbb{R}^d)$. The estimate (4.6) shows that

$$\sup_{n\in\mathbb{N}} \mu_t^n(B_R) < +\infty \quad \forall t \in [0,T],\ R > 0. \tag{4.17}$$

The equiintegrability condition on $\boldsymbol{\nu}^n$ shows that $\boldsymbol{\nu}$ satisfies

$$|\boldsymbol{\nu}|(B_R \times I) = \int_I m_R(t)\,\mathrm{d}t \qquad \forall I \in \mathcal{B}(0,T),\ R > 0, \quad \text{for some } m_R \in L^1(0,T),$$

so that by the disintegration theorem we can represent it as $\boldsymbol{\nu} = \int_0^T \boldsymbol{\nu}_t$ for a Borel family $\{\boldsymbol{\nu}_t\}_{t\in(0,T)}$ still satisfying (4.2). Let us now consider a function $\zeta \in C^1_c(\mathbb{R}^d)$ and for a given interval $I = [t_0, t_1] \subset [0,T]$ the time dependent function $\boldsymbol{\zeta}(t,x) := \chi_I(t)\nabla\zeta(x)$. Since the discontinuity set of $\boldsymbol{\zeta}$ is concentrated on $N = \mathbb{R}^d \times \{t_0, t_1\}$ and $|\boldsymbol{\nu}|(N) = 0$, general convergence theorems (see e.g. [3, Prop. 5.1.10] yields

$$\begin{aligned} \lim_{n\to\infty} \int_I \int_{\mathbb{R}^d} \nabla\zeta(x) \cdot \mathrm{d}\boldsymbol{\nu}_t^n(x)\,\mathrm{d}t &= \lim_{n\to\infty} \int_{\mathbb{R}^d\times(0,T)} \boldsymbol{\zeta} \cdot \mathrm{d}\boldsymbol{\nu}^n(t,x) \\ &= \int_{\mathbb{R}^d\times(0,T)} \boldsymbol{\zeta} \cdot \mathrm{d}\boldsymbol{\nu}(t,x) = \int_I \int_{\mathbb{R}^d} \nabla\zeta(x) \cdot \mathrm{d}\boldsymbol{\nu}_t(x)\,\mathrm{d}t. \end{aligned} \tag{4.18}$$

Applying (4.5) with $\zeta(t,x) := \zeta(x)$ and $t_0 := s$ and the estimate (4.17) we thus obtain the weak convergence of μ_t^n to a measure $\mu_t \in \mathcal{M}^+(\mathbb{R}^d)$ for every $t \in [0,T]$. It is immediate to check that the couple $(\mu_t, \boldsymbol{\nu}_t)$ belongs to $\mathcal{CE}(0,T)$. (4.16) follows now by the representation

$$\int_0^T \Phi(\mu_t, \boldsymbol{\nu}_t | \gamma)\,\mathrm{d}t = \Phi(\mu, \boldsymbol{\nu} | \overline{\gamma}), \quad \mu := \int_0^T \mu_t\,\mathrm{d}t,\ \overline{\gamma} = \gamma \otimes \mathscr{L}^1 \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d \times (0,T))$$

and the lower semicontinuity property stated in Theorem 2.1. □

4.2 Solutions of the continuity equation with finite Φ-energy

For all this section we will assume that $\phi : (0,+\infty)\times\mathbb{R}^d \to (0,+\infty)$ is an admissible action density function as in (3.2a,b,c) for some $p\in(1,+\infty)$, $\gamma\in\mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ is a given reference Radon measure, and Φ is the corresponding integral functional as in (3.37). We want to study the properties of measure valued solutions $(\mu,\boldsymbol{\nu})$ of the continuity equation (4.1) with finite Φ-energy

$$E := \int_0^T \Phi(\mu_t,\boldsymbol{\nu}_t|\gamma)\,\mathrm{d}t < +\infty. \tag{4.19}$$

We denote by $\mathcal{CE}_{\phi,\gamma}(0,T)$ the subset of $\mathcal{CE}(0,T)$ whose elements $(\mu,\boldsymbol{\nu})$ satisfies (4.19).

Remark 4.6 If $(\mu_t)_{t\in[0,T]}$ is weakly* continuous in $\mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ and (4.19) holds, then $\mu_t,\boldsymbol{\nu}_t$ also satisfy (4.2): in fact, the weak* continuity of μ_t yields for every $R>0$ $\sup_{t\in[0,T]}\mu_t(B_R)=M_R<+\infty$, and the estimate (3.48) yields (recall (3.43))

$$V_R \le \eta \int_0^T \|\boldsymbol{\nu}_t\|(B_R)\,\mathrm{d}t \le \eta\, T^{1/q} E^{1/p} H(\gamma(B_R),M_R)^{1/q} < +\infty. \tag{4.20}$$

Recalling that the function h is defined by (3.44), we also introduce the concave function

$$\omega(s) := \int_0^s \frac{1}{\mathsf{h}(r)^{1/q}}\,\mathrm{d}r,\quad \omega(0)=0,\quad \omega'(s)=\frac{1}{\mathsf{h}(s)^{1/q}},\quad \lim_{s\to\infty}\omega(s)=+\infty. \tag{4.21}$$

In the homogeneous case $\phi(\rho,\mathbf{z})=\rho^\alpha\|\mathbf{z}\|_*^q$ we have

$$\omega(s) = \int_0^s r^{-\alpha/q}\,\mathrm{d}r = \frac{q}{q-\alpha}s^{1-\alpha/q} = \frac{p}{\theta}s^{\theta/p}. \tag{4.22}$$

For given nonnegative $\zeta\in C^1_{\rm c}(\mathbb{R}^d)$ and $\mu\in\mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ we will use the short notation

$$Z := \mathrm{supp}(\mathrm{D}\zeta)\subset\mathbb{R}^d,\quad G_p(\zeta) := \int_Z \zeta^p\,\mathrm{d}\gamma,\quad D(\zeta) := \sup_{\mathbb{R}^d}\|\mathrm{D}\zeta\|_*. \tag{4.23}$$

Theorem 4.7 *Let $\zeta\in C^1_{\rm c}(\mathbb{R}^d)$ be a nonnegative function with $Z, G(\zeta), D(\zeta)$ defined as in* (4.23)*, and let $\mu,\boldsymbol{\nu}\in\mathcal{CE}_{\phi,\gamma}(0,T)$. Setting*

$$E_Z := \int_0^T \Phi(\mu_t,\boldsymbol{\nu}_t|\gamma,Z)\,\mathrm{d}t \le E < +\infty, \tag{4.24}$$

we have

$$\left|\tfrac{\mathrm{d}}{\mathrm{d}t}\mu_t(\zeta^p)\right| \le p D(\zeta)\,\Phi(\mu_t,\boldsymbol{\nu}_t|\gamma,Z)^{1/p} H\big(G_p(\zeta),\mu_t(\zeta^p)\big)^{1/q}. \tag{4.25}$$

In particular, there exists a constant $\mathsf{C}_1>0$ only depending (in a monotone way) on h,p,T such that

$$\sup_{t\in[0,T]}\mu_t(\zeta^p) \le \mathsf{C}_1\Big(\mu_0(\zeta^p) + D(\zeta)G_p(\zeta)^{1/q}E_Z^{1/p} + D^p(\zeta)E_Z\Big). \tag{4.26}$$

Moreover, if $G_p(\zeta)>0$,

$$\left|\tfrac{\mathrm{d}}{\mathrm{d}t}\omega\left(\mu_t(\zeta^p)/G_p(\zeta)\right)\right| \le \frac{pD(\zeta)}{G_p(\zeta)^{1/p}}\Phi(\mu_t,\boldsymbol{\nu}_t|\gamma,Z)^{1/p}\quad \textit{for a.e. } t\in(0,T). \tag{4.27}$$

In particular, in the $(\alpha$-$\theta)$-homogeneous case, for every $0\le s\le t\le T$ we have

$$\Big|\|\zeta\|^\theta_{L^p(\mu_t)} - \|\zeta\|^\theta_{L^p(\mu_s)}\Big| \le \theta D(\zeta)\,\|\zeta\|^{\theta-1}_{L^p(\gamma,Z)}\int_s^t \Phi(\mu_r,\boldsymbol{\nu}_r|\gamma,Z)^{1/p}\,\mathrm{d}r. \tag{4.28}$$

Proof Setting $m_t := \mu_t(\zeta^p),\ G = G_p(\zeta),\ D = D(\zeta)$ we easily have by (3.47)

$$\frac{\mathrm{d}}{\mathrm{d}t}m_t = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}\zeta^p\,\mathrm{d}\mu_t = p\int_Z \zeta^{p-1}\nabla\zeta\cdot\mathrm{d}\boldsymbol{\nu}_t \le pD\,\Phi(\mu_t,\boldsymbol{\nu}_t|\gamma,Z)^{1/p}H\big(G,m_t\big)^{1/q},$$

since $(\zeta^{p-1})^q = \zeta^p$. Since $H\big(G,m_t\big) = G\mathsf{h}(m_t/G)$, we get

$$\mathsf{h}^{-1/q}(m_t/G)\frac{\mathrm{d}}{\mathrm{d}t}m_t \le pDG^{1/q}\Phi(\mu_t,\boldsymbol{\nu}_t|\gamma,Z)^{1/p}.$$

Recalling that $\frac{\mathrm{d}}{\mathrm{d}r}\omega(r) = \mathsf{h}^{-1/q}(r)$ we get (4.27).

In order to prove (4.26) we set $M := \sup_{t\in[0,T]} m_t$ and we choose constants $(a,b)\in\Gamma_\phi$; integrating (4.25) we get

$$\sup_{t\in[0,T]}\big|m_t - m_0\big| \le pDT^{1/q}\Big(\big(aG\big)^{1/q}E_Z^{1/p} + \big(bM\big)^{1/q}E_Z^{1/p}\Big). \tag{4.29}$$

By using the inequality $xy \le p^{-1}x^p + q^{-1}y^q$ we obtain

$$M \le m_0 + pD\,\big(a\,T\,G\big)^{1/q}E_Z^{1/p} + \frac{1}{q}M + p^{p-1}D^p\,\big(b\,T\big)^{p/q}E_Z \tag{4.30}$$

which yields (4.26) with $\mathsf{C}_1 := p\max\big(1, p(aT)^{1/q}, p^{p-1}(bT)^{p/q}\big)$.

Finally, let us assume that ϕ satisfies the $(\alpha\text{-}\theta)$-homogeneity condition, so that $\omega(s) = \frac{p}{\theta}s^{\theta/p}$ as in (4.22). It follows that

$$\omega(G^{-1}m_t) = \frac{p}{\theta}\|\zeta\|^{\theta}_{L^p(\mu_t)}\|\zeta\|^{-\theta}_{L^p(\gamma,Z)}. \tag{4.31}$$

Integrating (4.27) we conclude. $\square$

We extend the definition of $\mathsf{m}_r(\mu)$ also for negative values of r by setting

$$\tilde{\mathsf{m}}_r(\mu) := \mu(B_1) + \int_{\mathbb{R}^d\setminus B_1}|x|^r\,\mathrm{d}\mu(x) = \int_{\mathbb{R}^d}\big(1\vee|x|\big)^r\,\mathrm{d}\mu(x)\quad \forall r\in\mathbb{R}. \tag{4.32}$$

Notice that $\tilde{\mathsf{m}}_0(\mu) = \mu(\mathbb{R}^d)$ and $\mathsf{m}_r(\mu) \le \tilde{\mathsf{m}}_r(\mu) \le \mu(B_1) + \mathsf{m}_r(\mu)$ when $r > 0$.

Theorem 4.8 *Let us assume that $\tilde{\mathsf{m}}_r(\gamma) < +\infty$ for some $r \le p$ and let $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,T)$ satisfy* (4.19)*. For every $\delta \le 1 + r/q$, if $\tilde{\mathsf{m}}_\delta(\mu_0) < +\infty$ then also $\tilde{\mathsf{m}}_\delta(\mu_t) < +\infty$ and there exists a constant C_2 only depending in a monotone way on $\mathsf{h}, p, T, A, |\delta|$ such that*

$$\tilde{\mathsf{m}}_\delta(\mu_t) \le \mathsf{C}_2\Big(\tilde{\mathsf{m}}_\delta(\mu_0) + \tilde{\mathsf{m}}_r(\gamma)^{1/q}E^{1/p} + E\Big). \tag{4.33}$$

Moreover, if $r \ge -q$ and $\mu_0(\mathbb{R}^d) < +\infty$, then $\mu_t(\mathbb{R}^d)$ is finite and constant for every $t\in[0,T]$.

Proof Let us first set

$$K_n := 2^{nr}\gamma(B_{2^{n+1}}\setminus B_{2^n}) \tag{4.34}$$

observing that

$$K_n \le \sum_{j=0}^{+\infty}K_j \le 2^{r^-}\,\tilde{\mathsf{m}}_r(\gamma),\qquad \limsup_{n\uparrow+\infty}K_n = 0. \tag{4.35}$$

We consider the usual cutoff functions $\zeta_n \in C^\infty_{\mathrm{c}}(\mathbb{R}^d)$ as in (4.11a,b) and we set

$$D_n = D(\zeta_n) = \sup\|\mathrm{D}\zeta_n\|_* \le A2^{-n},\quad G_n = G_p(\zeta_n) \le \gamma(B_{2^{n+1}}\setminus B_{2^n}) = 2^{-nr}K_n. \tag{4.36}$$

By (4.26) we obtain

$$\sup_{t\in[0,T]} \mu_t(B_{2^n}) \le \mathsf{C}_1\Big(\mu_0(B_{2^{n+1}}) + A\,2^{-n(1+r/q)}K_n^{1/q}E^{1/q} + A^p\,2^{-np}E\Big); \tag{4.37}$$

in particular, if $r \ge -q$ and $\mu_0(\mathbb{R}^d) < +\infty$, we can derive the uniform upper bound $\mu_t(\mathbb{R}^d) \le \mathsf{C}_1\,\mu_0(\mathbb{R}^d)$ letting $n \uparrow +\infty$. We can then deduce that $\mu_t(\mathbb{R}^d)$ is constant by applying the estimate (4.25), which yields after an integration in time and for every $(a,b) \in \Gamma_\phi$

$$\sup_{t\in[0,T]} \big|\mu_t(\zeta_n^p) - \mu_0(\zeta_n^p)\big| \le pAT^{1/q}E^{1/p}\,2^{-n}\big(a2^{-nr}K_n + b\mathsf{C}_1\mu_0(\mathbb{R}^d)\big)^{1/q}.$$

In order to show (4.33), we argue as before, by introducing the new family of test functions induced by $\upsilon_n(x) := \upsilon_0(x/2^n) \in C_c^\infty(\mathbb{R}^d)$

$$0 \le \upsilon_n \le 1, \quad \begin{cases} \upsilon_n(x) \equiv 1 & \text{if } 2^n \le |x| \le 2^{n+1}, \\ \upsilon_n(x) \equiv 0 & \text{if } |x| \le 2^{n-1} \text{ or } |x| \ge 2^{n+2}, \end{cases} \quad \|\mathrm{D}\upsilon_n\|_* \le A\,2^{-n}. \tag{4.38}$$

Observe that $1 \le \sum_{n=1}^{+\infty}\big(\upsilon_n(x)\big)^p \le 3$ and for some constant $A_\delta > 1$

$$A_\delta^{-1}\,|x|^\delta \le \sum_{n=1}^{+\infty} 2^{\delta n}\big(\upsilon_n(x)\big)^p \le A_\delta\,|x|^\delta \quad \forall x \in \mathbb{R}^d,\ |x| \ge 2. \tag{4.39}$$

As before, setting $K_n' := K_{n+1} + K_{n-1}$, we have $D(\upsilon_n) \le A\,2^{-n}$ and

$$G_p(\upsilon_n) \le \Big(2^{-(n+1)r}K_{n+1} + 2^{-(n-1)r}K_{n-1}\Big) \le 2^{|r|}\,2^{-nr}\,K_n'. \tag{4.40}$$

Applying (4.26) we get for every $t \in [0,T]$

$$2^{n\delta}\mu_t(\upsilon_n^p) \le \mathsf{C}_1\Big(2^{n\delta}\mu_0(\upsilon_n^p) + A\,2^{(\delta-1-r/q)n}(K_n')^{1/q}\,(E_n')^{1/p} + A^p\,2^{(\delta-p)n}E_n'\Big),$$

where

$$E_n := \int_0^T \Phi(\mu_t,\boldsymbol{\nu}_t|\gamma, B_{2^{n+1}}\setminus B_{2^n})\,\mathrm{d}t, \quad E_n' := E_{n+1} + E_{n-1}. \tag{4.41}$$

Since $\delta \le 1 + r/q$ and $\delta \le p$, summing up with respect to n and recalling (4.37) we get

$$\tilde{\mathsf{m}}_\delta(\mu_t) \le \mathsf{C}_2\Big(\tilde{\mathsf{m}}_\delta(\mu_0) + (\tilde{\mathsf{m}}_r(\gamma))^{1/q}E^{1/p} + E\Big). \quad \square \tag{4.42}$$

In the the θ-homogeneous case we have a more refined estimate:

Theorem 4.9 *Let us assume that ϕ is θ-homogeneous for some $\theta \in (1,p]$, the measure γ satisfies the r-moment condition $\tilde{\mathsf{m}}_r(\gamma) < +\infty$, and let $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,T)$ satisfy* (4.19). *For every $\delta \le \overline{\delta} := \frac{1}{\theta}p + (1-\frac{1}{\theta})r$, if $\tilde{\mathsf{m}}_\delta(\mu_0) < +\infty$ then $\tilde{\mathsf{m}}_\delta(\mu_t)$ is finite and there exists a constant $\mathsf{C}_3 > 0$ such that*

$$\tilde{\mathsf{m}}_\delta(\mu_t) \le \mathsf{C}_3\Big(\tilde{\mathsf{m}}_\delta(\mu_0) + \tilde{\mathsf{m}}_r(\gamma)^{1-1/\theta}E^{1/\theta}\Big). \tag{4.43}$$

Moreover, if $\overline{\delta} \ge 0$ (i.e. $r \ge -p/(\theta-1)$) and $\mu_0(\mathbb{R}^d) < +\infty$ then $\mu_t(\mathbb{R}^d)$ is finite and constant for $t \in [0,T]$.

Proof We argue as in the proof of Theorem (4.9), keeping the same notation and using the crucial estimate (4.28). If ζ_n are the test functions of (4.11a,b),

$$\|\zeta_n\|^{\theta-1}_{L^p(\gamma,Z_n)} = G_n^{(\theta-1)/p} \overset{(4.36)}{=} 2^{-nr(\theta-1)/p}\big(K_n\big)^{(\theta-1)/p}, \tag{4.44}$$

so that, since $\bar\delta\theta/p = 1+(\theta-1)r/p$, (4.28) yields

$$\Big|\big(\mu_t(\zeta_n^p)\big)^{\theta/p} - \big(\mu_0(\zeta_n^p)\big)^{\theta/p}\Big| \le A\,\theta 2^{-\bar\delta\theta n/p}\big(K_n\big)^{(\theta-1)/p}E^{1/p}. \tag{4.45}$$

Since $\bar\delta \ge 0$, passing to the limit as $n\uparrow\infty$ and recalling (4.35), we get $\mu_t(\mathbb{R}^d)\equiv\mu_0(\mathbb{R}^d)$. Concerning the moment estimate, we replace ζ_n by υ_n, defined by in (4.38), obtaining

$$\Big|\big(\mu_t(\upsilon_n^p)\big)^{\theta/p} - \big(\mu_0(\upsilon_n^p)\big)^{\theta/p}\Big| \le \mathsf{C}_{3.1}2^{-\bar\delta\theta n/p}\big(K_n'\big)^{(\theta-1)/p}\big(E_n'\big)^{1/p}, \tag{4.46}$$

and therefore

$$\mu_t(\upsilon_n^p) \le \mathsf{C}_{3.2}\Big(\mu_0(\upsilon_n^p) + 2^{-\bar\delta n}\big(K_n'\big)^{1-1/\theta}\big(E_n'\big)^{1/\theta}\Big). \tag{4.47}$$

Multiplying this inequality by $2^{n\delta}$, summing up w.r.t. n, and recalling (4.39), we obtain

$$\tilde{\mathsf{m}}_\delta(\mu_t) \le \mathsf{C}_3\Big(\tilde{\mathsf{m}}_\delta(\mu_0) + \tilde{\mathsf{m}}_r(\gamma)^{1-1/\theta}E^{1/\theta}\Big). \quad \square \tag{4.48}$$

Corollary 4.10 (Compactness for solutions of the continuity equation (II)) *Let $(\mu^n,\boldsymbol{\nu}^n)$ be a sequence in $\mathcal{CE}_{\phi,\gamma}(0,T)$ and let $\gamma^n \rightharpoonup^* \gamma$ in $\mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ such that*

$$\sup_{n\in\mathbb{N}} \mu_0^n(B_R) < +\infty \quad \forall R>0, \qquad \sup_{n\in\mathbb{N}} \int_0^T \Phi(\mu_t^n,\boldsymbol{\nu}_t^n|\gamma^n)\,\mathrm{d}t < +\infty. \tag{4.49}$$

Then conditions 1. and 2. of Lemma 4.5 are satisfied and therefore there exists a subsequence (still indexed by n) and a couple $(\mu_t,\boldsymbol{\nu}_t)\in\mathcal{CE}_{\phi,\gamma}(0,T)$ such that

$$\begin{aligned} \mu_t^n &\rightharpoonup^* \mu_t \quad \textit{weakly}^* \textit{ in } \mathcal{M}^+_{\rm loc}(\mathbb{R}^d) \quad \forall t\in[0,T],\\ \boldsymbol{\nu}^n &\rightharpoonup^* \boldsymbol{\nu} \quad \textit{weakly}^* \textit{ in } \mathcal{M}_{\rm loc}(\mathbb{R}^d\times(0,T);\mathbb{R}^d), \end{aligned} \tag{4.50}$$

$$\int_0^T \Phi(\mu_t,\boldsymbol{\nu}_t|\gamma)\,\mathrm{d}t \le \liminf_{n\uparrow+\infty}\int_0^T \Phi(\mu_t^n,\boldsymbol{\nu}_t^n|\gamma^n)\,\mathrm{d}t. \tag{4.51}$$

Suppose moreover that $\mu_0^n(\mathbb{R}^d)\to\mu_0(\mathbb{R}^d)$ and $\sup_n \tilde{\mathsf{m}}_\kappa(\gamma^n) < +\infty$ where $\kappa=-q$ or $\kappa = -p/(\theta-1)$ in the θ-homogeneous case, then (along the same subsequence) $\mu_t^n(\mathbb{R}^d)\to\mu_t(\mathbb{R}^d)$ for every $t\in[0,T]$.

Proof Since $P_R := \sup_n \gamma^n(B_R) < +\infty$ for every $R>0$, the estimate (4.33) for $\delta=0$ and the assumption (4.49) show that $M_R = \sup_{n\in\mathbb{N},t\in[0,T]}\mu_t^n(B_R) < +\infty$ for every $R>0$. We can therefore obtain a bound of $\|\boldsymbol{\nu}_t^n\|(B_R)$ by (3.48), which yields

$$\|\boldsymbol{\nu}_t^n\|(B_R) \le H(P_R,M_R)^{1/q}\Phi(\mu_t,\boldsymbol{\nu}_t|\gamma)^{1/p},$$

so that the maps $t\mapsto\|\boldsymbol{\nu}_t^n\|(B_R)$ are uniformly bounded by a function in $L^p(0,T)$. The last assertion follows by the fact that $t\mapsto\mu_t^n(\mathbb{R}^d)$ is independent of time, thanks to Theorem 4.9 (in the $(\alpha$-$\theta)$-homogeneous case) or Theorem 4.8 (for general density functions ϕ). $\square$

5 The (ϕ-γ)-weighted Wasserstein distance

As we already mentioned in the Introduction, BENAMOU-BRENIER [7] showed that the Wasserstein distance W_p (1.1) can be equivalently characterized by a "dynamic" point of view through (1.15), involving the 1-homogeneous action functional (1.13). The same approach can be applied to arbitrary action functionals.

Definition 5.1 (Weighted Wasserstein distances) Let $\gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ be a fixed reference measure and $\phi : (0,+\infty)\times\mathbb{R}^d \to [0,+\infty)$ a function satisfying Conditions (3.2a,b,c). The (ϕ,γ)-Wasserstein (pseudo-) distance between $\mu_0,\mu_1 \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ is defined as

$$\mathcal{W}^p_{\phi,\gamma}(\mu_0,\mu_1) := \inf\Big\{\int_0^1 \Phi(\mu_t,\boldsymbol{\nu}_t|\gamma)\,\mathrm{d}t : \quad (\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,1;\mu_0\to\mu_1)\Big\}. \tag{5.1}$$

We denote by $\mathcal{M}_{\phi,\gamma}[\mu_0]$ the set of all the measures $\mu \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ which are at finite $\mathcal{W}_{\phi,\gamma}$-distance from μ_0.

Remark 5.2 Let us recall the notation $W_{p,\alpha;\gamma}$ of (1.23) in the case $\phi_{p,\alpha}(\rho,\mathbf{w}) = \rho^\alpha|\mathbf{w}/\rho^\alpha|^p$. When $\alpha = 0$ we find the dual homogeneous Sobolev (pseudo-)distance (1.7) and in the case $\alpha = 1$ and $\mathrm{supp}(\gamma) = \mathbb{R}^d$ we get the usual Wasserstein distance:

$$\|\mu_0-\mu_1\|_{\dot W^{-1,p}_\gamma} = W_{p,0;\gamma}(\mu_0,\mu_1), \qquad W_p(\mu_0,\mu_1) = W_{p,1;\gamma}(\mu_0,\mu_1).$$

Remark 5.3 Taking into account Lemma 4.3, a linear time rescaling shows that

$$W^p_{\phi,\gamma}(\mu_0,\mu_T) := \inf\Big\{T^{p-1}\int_0^T \Phi(\mu_t,\boldsymbol{\nu}_t|\gamma)\,\mathrm{d}t : (\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,T;\mu_0\to\mu_T)\Big\}. \tag{5.2}$$

Theorem 5.4 (Existence of minimizers) *Whenever the infimum in* (5.1) *is a finite value* $W < +\infty$*, it is attained by a curve* $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,1)$ *such that*

$$\Phi(\mu_t,\boldsymbol{\nu}_t|\gamma) = W \quad \textit{for } \mathscr{L}^1\textit{-a.e. } t \in (0,1). \tag{5.3}$$

The curve $(\mu_t)_{t\in[0,1]}$ *associated to a minimum for* (5.1) *is a constant speed mimimal geodesic for* $\mathcal{W}_{\phi,\gamma}$ *since it satisfies*

$$\mathcal{W}_{\phi,\gamma}(\mu_s,\mu_t) = |t-s|\,\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) \quad \forall s,t \in [0,1]. \tag{5.4}$$

We have also the equivalent characterization

$$\mathcal{W}_{\phi,\gamma}(\sigma,\eta) = \inf\Big\{\int_0^T \Big(\Phi(\mu_t,\boldsymbol{\nu}_t|\gamma)\Big)^{1/p}\mathrm{d}t : (\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,T;\sigma\to\eta)\Big\}. \tag{5.5}$$

Proof When $\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) < +\infty$, Corollary 4.10 immediately yields the existence of a minimizing curve $(\mu,\boldsymbol{\nu})$. Just for the proof of (5.5), let us denote by $\overline{\mathcal{W}}_{\phi,\gamma}(\sigma,\eta)$ the infimum of the right-hand side of (5.5). Hölder inequality immediately shows that $\mathcal{W}_{\phi,\gamma}(\sigma,\eta) \ge \overline{\mathcal{W}}_{\phi,\gamma}(\sigma,\eta)$. In order to prove the opposite inequality, we argue as in [3, Lemma 1.1.4], defining for $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,T;\sigma\to\eta)$

$$\mathsf{s}_\varepsilon(t) := \int_0^t \Big(\varepsilon + \Phi(\mu_r,\boldsymbol{\nu}_r|\gamma)\Big)^{1/p}\mathrm{d}r, \quad t \in [0,T]; \tag{5.6}$$

s_ε is strictly increasing with $\mathsf{s}'_\varepsilon \geq \varepsilon$, $\mathsf{s}_\varepsilon(0,T) = (0,S_\varepsilon)$ with $S_\varepsilon := \mathsf{s}_\varepsilon(T)$, so that its inverse map $\mathsf{t}_\varepsilon : [0,S_\varepsilon] \to [0,T]$ is well defined and Lipschitz continuous, with

$$\mathsf{t}'_\varepsilon \circ \mathsf{s}_\varepsilon = \Big(\varepsilon + \Phi(\mu_t, \boldsymbol{\nu}_t)\Big)^{-1/p} \quad \text{a.e. in } (0,T). \tag{5.7}$$

If $\hat\mu^\varepsilon = \mu \circ \mathsf{t}_\varepsilon, \hat{\boldsymbol{\nu}}^\varepsilon := \mathsf{t}'_\varepsilon \boldsymbol{\nu} \circ \mathsf{t}_\varepsilon$, we know that $(\hat\mu^\varepsilon, \hat{\boldsymbol{\nu}}^\varepsilon) \in \mathcal{CE}(0,S_\varepsilon;\sigma \to \eta)$ so that

$$\mathcal{W}^p_{\phi,\gamma}(\sigma,\eta) \leq S_\varepsilon^{p-1} \int_0^{S_\varepsilon} \Phi(\hat\mu^\varepsilon_s, \hat{\boldsymbol{\nu}}^\varepsilon_s|\gamma)\,\mathrm{d}s = S_\varepsilon^{p-1} \int_0^T \frac{\Phi(\mu^\varepsilon_t, \boldsymbol{\nu}^\varepsilon_t|\gamma)}{\varepsilon + \Phi(\mu^\varepsilon_t, \boldsymbol{\nu}^\varepsilon_t|\gamma)} \Big(\varepsilon + \Phi(\mu_t, \boldsymbol{\nu}_t)\Big)^{1/p}\mathrm{d}t,$$

the latter integral being less than S_ε^p. Passing to the limit as $\varepsilon \downarrow 0$, we get

$$\mathcal{W}_{\phi,\gamma}(\sigma,\eta) \leq \int_0^T \Phi(\mu_t, \boldsymbol{\nu}_t|\gamma)^{1/p}\,\mathrm{d}t \quad \forall (\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,T;\sigma \to \eta), \tag{5.8}$$

and therefore $\mathcal{W}_{\phi,\gamma}(\sigma,\eta) \leq \overline{\mathcal{W}}_{\phi,\gamma}(\sigma,\eta)$. If $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,1;\mu_0 \to \mu_1)$ is a minimizer of (5.1), then (5.8) yields

$$W^{1/p} = \mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) = \Big(\int_0^1 \Phi(\mu_t, \boldsymbol{\nu}_t|\gamma)\,\mathrm{d}t\Big)^{1/p} = \int_0^1 \Phi(\mu_t, \boldsymbol{\nu}_t|\gamma)^{1/p}\,\mathrm{d}t,$$

so that (5.3) holds. □

5.1 Topological properties

Theorem 5.5 (Distance and weak convergence) *The functional $\mathcal{W}_{\phi,\gamma}$ is a (pseudo)-distance on $\mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ which induces a stronger topology than the weak* one. Bounded sets with respect to $\mathcal{W}_{\phi,\gamma}$ are weakly* relatively compact.*

Proof It is immediate to check that $\mathcal{W}_{\phi,\gamma}$ is symmetric (since $\phi(\rho,-\mathbf{w}) = \phi(\rho,\mathbf{w})$) and $\mathcal{W}_{\phi,\gamma}(\sigma,\eta) = 0 \Rightarrow \sigma \equiv \eta$. The triangular inequality follows as well from the characterization (5.5) and the gluing Lemma 4.4.

From (4.27) (keeping the same notation (4.23)) and (5.5) we immediately get for every $\mu_0, \mu_1 \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ and nonnegative $\zeta \in C^1_{\mathrm{c}}(\mathbb{R}^d)$ with $\|\zeta\|_{L^p(\gamma)} > 0$

$$\Big|\omega\big(\mu_1(\zeta^p)/G_p(\zeta)\big) - \omega\big(\mu_0(\zeta^p)/G_p(\zeta)\big)\Big| \leq \frac{pD(\zeta)}{\|\zeta\|_{L^p(\gamma)}} \mathcal{W}_{\phi,\gamma}(\sigma,\eta),$$

which shows the last assertion, since ω is strictly increasing and the set

$$\big\{\zeta^p : \zeta \in C^1_{\mathrm{c}}(\mathbb{R}^d), \quad \zeta \geq 0, \quad \|\zeta\|_{L^p(\gamma)} > 0\big\}$$

is dense in the space of nonnegative continuous functions with compact support (endowed with the uniform topology). □

Theorem 5.6 (Lower semicontinuity) *The map $(\mu_0,\mu_1) \mapsto \mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1)$ is lower semicontinuous with respect to weak* convergence in $\mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$. More generally, suppose that $\gamma^n \rightharpoonup^* \gamma$ in $\mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$, ϕ^n is monotonically increasing w.r.t. n and pointwise converging to ϕ, and $\mu^n_0 \rightharpoonup^* \mu_0, \mu^n_1 \rightharpoonup^* \mu_1$ in $\mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ as $n \uparrow +\infty$. Then*

$$\liminf_{n\uparrow+\infty} \mathcal{W}_{\phi^n,\gamma^n}(\mu^n_0,\mu^n_1) \geq \mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1). \tag{5.9}$$

Proof It is not restrictive to assume that $\mathcal{W}_{\phi^n,\gamma^n}(\mu_0^n,\mu_1^n) < S < +\infty$, so that we can find a sequence $(\mu^n,\boldsymbol{\nu}^n) \in \mathcal{CE}_{\phi^n,\gamma^n}(0,1;\mu_0^n \to \mu_1^n)$ such that

$$\Phi^m(\mu_t^n,\boldsymbol{\nu}_t^n|\gamma^n) \le S \quad \text{a.e. in } (0,1), \quad \forall m \le n \in \mathbb{N}, \tag{5.10}$$

where Φ^m denotes the integral functional associated to ϕ^m. We can apply Theorem 4.10 and we can extract a suitable subsequence (still denoted bu $\mu^n,\boldsymbol{\nu}^n$) and a limit curve $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,1;\mu_0 \to \mu_1)$ such that (4.50) holds. We eventually have

$$\mathcal{W}^p_{\phi^m,\gamma}(\mu_0,\mu_1) \le \int_0^1 \Phi^m(\mu_t,\boldsymbol{\nu}_t|\gamma)\,\mathrm{d}t \le S. \tag{5.11}$$

Passing to the limit w.r.t. $m \uparrow +\infty$ we conclude. □

Theorem 5.7 (Completeness) *For every $\sigma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ the space $\mathcal{M}_{\phi,\gamma}[\sigma]$ endowed with the distance $\mathcal{W}_{\phi,\gamma}$ is complete.*

Proof Let $(\mu_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathcal{M}_{\phi,\gamma}[\sigma]$ w.r.t. the distance $\mathcal{W}_{\phi,\gamma}$; in particular, (μ_n) is bounded so that we can extract a suitable convergence subsequence μ_{n_k} weakly* converging to μ_∞ in $\mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$. Thanks to Theorem 5.6 we easily get $\mathcal{W}_{\phi,\gamma}(\mu_m,\mu_\infty) \le \liminf_{k\to\infty}\mathcal{W}_{\phi,\gamma}(\mu_m,\mu_{n_k})$, and therefore, taking into account the Cauchy condition, $\limsup_{m\to\infty}\mathcal{W}_{\phi,\gamma}(\mu_m,\mu_\infty) \le \limsup_{n,m\to\infty}\mathcal{W}_{\phi,\gamma}(\mu_m,\mu_n) = 0$ so that μ_n converges to μ_∞. □

Let us now consider the case of measures with finite mass (just to fix the constant, probability measures in $\mathcal{P}(\mathbb{R}^d)$). We introduce the parameter

$$\kappa := \begin{cases} \dfrac{p}{\theta-1} = \dfrac{q}{1-\alpha} & \text{if } \phi \text{ is } (\alpha\text{-}\theta)\text{-homogeneous},\\ \frac{p}{p-1} = q & \text{otherwise.} \end{cases} \tag{5.12}$$

Theorem 5.8 (Distance and total mass) *Let us assume that $\tilde{\mathsf{m}}_{-\kappa}(\gamma) < +\infty$ and let us suppose that $\sigma \in \mathcal{P}(\mathbb{R}^d)$. Then $\mathcal{M}_{\phi,\gamma}[\sigma] \subset \mathcal{P}(\mathbb{R}^d)$, the weighted Wasserstein distance $\mathcal{W}_{\phi,\gamma}$ is stronger than the narrow convergence in $\mathcal{P}(\mathbb{R}^d)$, and $\mathcal{P}(\mathbb{R}^d)$ endowed with the (pseudo-) distance $\mathcal{W}_{\phi,\gamma}$ is a complete (pseudo-)metric space.*

Proof If $\eta \in \mathcal{M}_{\phi,\gamma}[\sigma]$ then Theorem 4.9 (in the θ-homogeneous case) or 4.8 (in the general case) yields $\eta(\mathbb{R}^d) = \sigma(\mathbb{R}^d) = 1$, so that $\mathcal{M}_{\phi,\gamma}[\sigma] \subset \mathcal{P}(\mathbb{R}^d)$. Since the narrow topology coincide with the weak* one in $\mathcal{P}(\mathbb{R}^d)$, Theorem 5.5 proves the second statement. The completeness of $\mathcal{P}(\mathbb{R}^d)$ with respect to the (pseudo) distance $\mathcal{W}_{\phi,\gamma}$ follows by Theorem 5.7. □

We can also prove some useful moment estimates.

Theorem 5.9 (Moment estimates) *Let us assume that $\tilde{\mathsf{m}}_r(\gamma) < +\infty$ for some $r \in \mathbb{R}$ and let us set*

$$\bar{\delta} := \begin{cases} \frac{1}{\theta}p + (1-\frac{1}{\theta})r = \frac{p}{\theta}(1+r/\kappa) & \textit{if } \phi \textit{ is } \theta\textit{-homogeneous},\\ 1+r/q \le p & \textit{otherwise.} \end{cases} \tag{5.13}$$

If $\tilde{\mathsf{m}}_\delta(\sigma) < +\infty$ for some $\delta \le \bar{\delta}$, and $\eta \in \mathcal{M}_{\phi,\gamma}[\sigma]$, then $\tilde{\mathsf{m}}_\delta(\eta)$ is finite and there exists a constant C only depending on ϕ,δ such that

$$\begin{cases} \tilde{\mathsf{m}}_\delta(\eta) \le \mathsf{C}\Big(\tilde{\mathsf{m}}_\delta(\sigma) + \tilde{\mathsf{m}}_r(\gamma) + \mathcal{W}^p_{\phi,\gamma}(\sigma,\eta)\Big)\\ \tilde{\mathsf{m}}_\delta(\eta) \le \mathsf{C}\Big(\tilde{\mathsf{m}}_\delta(\sigma) + \tilde{\mathsf{m}}_r(\gamma)^{1-1/\theta}W^{p/\theta}_{p,\alpha}(\sigma,\eta)\Big). \end{cases} \tag{5.14}$$

Moreover, when $\delta \ge 1$, the topology induced by $\mathcal{W}_{\phi,\gamma}$ in $\mathcal{M}_{\phi,\gamma}[\sigma]$ is stronger than the one induced by the Wasserstein distance W_δ.

Proof Let us first consider the general case: applying (4.33) we easily obtain (5.14). In order to prove the assertion about the convergence of the moments induced by $\mathcal{W}_{\phi,\gamma}$ (which is equivalent to the convergence in W_δ when $\delta \geq 1$), a simple modification of (4.42) yields

$$\int_{|x|\geq 2^n} |x|^\delta \,\mathrm{d}\eta \leq \mathsf{C}_3\Big(\int_{|x|\geq 2^{n-1}} |x|^\delta \,\mathrm{d}\sigma + \tilde{\mathsf{m}}_r(\gamma)^{1/q}\mathcal{W}_{\phi,\gamma}(\sigma,\eta) + W^p_{\phi,\gamma}(\sigma,\eta)\Big), \tag{5.15}$$

which shows that every sequence η_n converging to σ has δ-moments equi-integrable and therefore it is relatively compact with respect to the δ-Wasserstein distance when $\delta \geq 1$.

The θ-homogeneous case follows by the same argument and Theorem 4.9. □

Remark 5.10 There are interesting particular cases covered by the previous result:

1. When $\gamma(\mathbb{R}^d) < +\infty$ then $\mathcal{W}_{\phi,\gamma}$ is always stronger than the 1-Wasserstein distance W_1; in the θ-homogeneous case, $W_{p,\alpha;\gamma}$ also controls the $W_{p/\theta}$ distance.
2. When $\mathsf{m}_p(\gamma) < +\infty$, then $\mathcal{W}_{\phi,\gamma}$ is always stronger than W_p.
3. When ϕ is θ-homogeneous with $\theta > 1$ and γ is a probability measure with finite moments of arbitrary orders (this is the case of a log-concave probability measure), then all the measures $\sigma \in \mathcal{M}_{\phi,\gamma}[\gamma]$ have finite moments of arbitrary orders and the convergence with respect to $\mathcal{W}_{\phi,\gamma}$ yields the convergence in $\mathcal{P}_\delta(\mathbb{R}^d)$ for every $\delta > 0$.

5.2 **Geometric properties**

Theorem 5.11 (Convexity of the distance and uniqueness of geodesics) $\mathcal{W}^p_{\phi,\gamma}(\cdot,\cdot)$ *is convex, i.e. for every* $\mu_i^j \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$, $i,j = 0,1$, *and* $\tau \in [0,1]$, *if* $\mu_i^\tau = (1-\tau)\mu_i^0 + \tau\mu_i^1$,

$$\mathcal{W}^p_{\phi,\gamma}(\mu_0^\tau,\mu_1^\tau) \leq (1-\tau)\mathcal{W}^p_{\phi,\gamma}(\mu_0^0,\mu_1^0) + \tau\mathcal{W}^p_{\phi,\gamma}(\mu_0^1,\mu_1^1). \tag{5.16}$$

If ϕ is strictly convex and $\tilde\phi$ has a sublinear growth w.r.t. ρ (i.e. $\tilde\varphi_\infty \equiv 0$), then for every $\mu_0,\mu_1 \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ with $\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) < +\infty$ there exists a unique *mimimizer $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,1)$ of* (5.1).

Proof Let $(\mu^j,\boldsymbol{\nu}^j) \in \mathcal{CE}_{\phi,\gamma}(0,1;\mu_0^j \to \mu_1^j)$ be two minimizers of (5.1), $j = 0,1$. For $\tau \in [0,1]$ we set $\mu_t^\tau := (1-\tau)\mu_t^0 + \tau\mu_t^1$, $\boldsymbol{\nu}_t^\tau := (1-\tau)\boldsymbol{\nu}_t^0 + \tau\boldsymbol{\nu}_t^1$. Since $(\mu^\tau,\boldsymbol{\nu}^\tau) \in \mathcal{CE}(0,1;\mu_0^\tau \to \mu_1^\tau)$, the convexity of ϕ yields

$$\begin{aligned}\mathcal{W}^p_{\phi,\gamma}(\mu_0^\tau,\mu_1^\tau) &\leq \int_0^1 \Phi(\mu_t^\tau,\boldsymbol{\nu}_t^\tau|\gamma)\,\mathrm{d}t \leq \int_0^1 \Big((1-\tau)\Phi(\mu_t^0,\boldsymbol{\nu}_t^0|\gamma) + \tau\Phi(\mu_t^1,\boldsymbol{\nu}_t^1|\gamma)\Big)\,\mathrm{d}t \\ &= (1-\tau)\mathcal{W}^p_{\phi,\gamma}(\mu_0^0,\mu_1^0) + \tau\mathcal{W}^p_{\phi,\gamma}(\mu_0^1,\mu_1^1).\end{aligned}$$

Let us now suppose that ϕ is strictly convex and sublinear. Setting, as usual, $\mu_t^\tau = \rho_t^\tau\gamma + (\mu_t^\tau)^\perp$, $\boldsymbol{\nu}_t^\tau = \mathbf{w}_t^\tau\gamma$, we have for a.e. $t \in (0,1)$

$$\Phi(\mu_t^\tau,\boldsymbol{\nu}_t^\tau|\gamma) \leq (1-\tau)\int_{\mathbb{R}^d} \phi(\rho_t^0,\mathbf{w}_t^0)\,\mathrm{d}\gamma + \tau\int_{\mathbb{R}^d} \phi(\rho_t^1,\mathbf{w}_t^1)\,\mathrm{d}\gamma \tag{5.17}$$

and the inequality is strict unless $\rho_t^0 \equiv \rho_t^1$ and $\mathbf{w}_t^0 \equiv \mathbf{w}_t^1$ for γ-a.e. $x \in \mathbb{R}^d$. If $\mu_0^0 = \mu_0^1$ and $\mu_1^0 = \mu_1^1$, two minimizers should satisfy

$$\rho_t^0(x) = \rho_t^1(x),\ \mathbf{w}_t^0(x) = \mathbf{w}_t^1(x) \quad \gamma\text{-a.e.}, \quad \boldsymbol{\nu}_t^0 = \boldsymbol{\nu}_t^1 \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,1).$$

Since $(\mu^i, \boldsymbol{\nu}^i)$ are solutions of the continuity equation, taking the difference we obtain

$$\partial_t\big((\mu_t^0)^\perp - (\mu_t^1)^\perp\big) = \partial_t(\mu_t^0 - \mu_t^1) = -\nabla\cdot(\boldsymbol{\nu}_t^0 - \boldsymbol{\nu}_t^1) = 0 \quad \text{in } \mathbb{R}^d \times (0,1).$$

The difference $(\mu_t^0)^\perp - (\mu_t^1)^\perp$ is then independent of time and vanishes at $t = 0$, so that $\mu_t^0 = \mu_t^1$ for every $t \in [0,1]$. □

Theorem 5.12 (Subadditivity) *For every $\mu_i^j \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$, $i, j = 0, 1$, we have*

$$\mathcal{W}_{\phi,\gamma}(\mu_0^0 + \mu_0^1, \mu_1^0 + \mu_1^1) \le \mathcal{W}_{\phi,\gamma}(\mu_0^0, \mu_1^0) + \mathcal{W}_{\phi,\gamma}(\mu_0^1, \mu_1^1). \tag{5.18}$$

In particular

$$\mathcal{W}_{\phi,\gamma}(\mu_0 + \sigma, \mu_1 + \sigma) \le \mathcal{W}_{\phi,\gamma}(\mu_0, \mu_1) \quad \forall \sigma \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d). \tag{5.19}$$

Proof Let $(\mu^j, \boldsymbol{\nu}^j) \in \mathcal{CE}_{\phi,\gamma}(0,1; \mu_0^j \to \mu_1^j)$ be as in the proof of the previous Theorem. Since $(\mu^0 + \mu^1, \boldsymbol{\nu}^0 + \boldsymbol{\nu}^1) \in \mathcal{CE}(0,1; \mu_0^0 + \mu_0^1 \to \mu_1^0 + \mu_1^1)$, we get

$$\begin{aligned}
&\mathcal{W}_{\phi,\gamma}(\mu_0^0 + \mu_0^1, \mu_1^0 + \mu_1^1) \le \int_0^1 \Big(\Phi(\mu_t^0 + \mu_t^1, \boldsymbol{\nu}_t^0 + \boldsymbol{\nu}_t^1|\gamma)\Big)^{1/p}\,\mathrm{d}t \\
&\le \int_0^1 \Big[\Big(\Phi(\mu_t^0 + \mu_t^1, \boldsymbol{\nu}_t^0|\gamma)\Big)^{1/p} + \Big(\Phi(\mu_t^0 + \mu_t^1, \boldsymbol{\nu}_t^1|\gamma)\Big)^{1/p}\Big]\,\mathrm{d}t \\
&\le \int_0^1 \Big[\Big(\Phi(\mu_t^0, \boldsymbol{\nu}_t^0|\gamma)\Big)^{1/p} + \Big(\Phi(\mu_t^1, \boldsymbol{\nu}_t^1|\gamma)\Big)^{1/p}\Big]\,\mathrm{d}t = \mathcal{W}_{\phi,\gamma}(\mu_0^0, \mu_1^0) + \mathcal{W}_{\phi,\gamma}(\mu_0^1, \mu_1^1). \quad \square
\end{aligned}$$

Proposition 5.13 (Rescaling) *For every $\mu_0, \mu_1 \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ and $\lambda > 0$ we have*

$$W^p_{\phi,\lambda\gamma}(\lambda\mu_0, \lambda\mu_1) = \lambda \mathcal{W}^p_{\phi,\gamma}(\mu_0, \mu_1), \tag{5.20}$$

$$\begin{cases} W^p_{\phi,\gamma}(\lambda\mu_0, \lambda\mu_1) \le \lambda^p \mathcal{W}^p_{\phi,\gamma}(\mu_0, \mu_1) & \text{if } \lambda \ge 1 \\ W^p_{\phi,\gamma}(\lambda\mu_0, \lambda\mu_1) \le \lambda \mathcal{W}^p_{\phi,\gamma}(\mu_0, \mu_1) & \text{if } \lambda \le 1. \end{cases} \tag{5.21}$$

Proof (5.20) follows from the corresponding property $\Phi(\lambda\mu, \lambda\boldsymbol{\nu}|\lambda\gamma) = \lambda\Phi(\mu, \boldsymbol{\nu}|\gamma)$. Analogously, the monotonicity and homogeneity properties of ϕ yield

$$\phi(\lambda\rho, \lambda\mathbf{w}) \le \phi(\rho, \lambda\mathbf{w}) = \lambda^p \phi(\rho, \mathbf{w}) \quad \text{if } \lambda > 1;$$

the convexity of ϕ and the fact that $\phi(0,0) = 0$ yield

$$\phi(\lambda\rho, \lambda\mathbf{w}) \le \lambda\phi(\rho, \mathbf{w}) \quad \text{if } \lambda < 1.$$

(5.21) follows immediately by the previous inequalities. □

Proposition 5.14 (Monotonicity) *If $\gamma_1 \ge \gamma_2$ and $\phi_1 \le \phi_2$, then for every $\mu_0, \mu_1 \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ we have*

$$\mathcal{W}_{\phi_1,\gamma_1}(\mu_0, \mu_1) \le \mathcal{W}_{\phi_2,\gamma_2}(\mu_0, \mu_1). \tag{5.22}$$

Theorem 5.15 (Convolution) *Let $k \in C^\infty_c(\mathbb{R}^d)$ be a nonnegative convolution kernel with $\int_{\mathbb{R}^d} k(x)\,\mathrm{d}x = 1$ and let $k_\varepsilon(x) := \varepsilon^{-d} k(x/\varepsilon)$. For every $\mu_0, \mu_1 \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ we have*

$$\mathcal{W}_{\phi,\gamma * k_\varepsilon}(\mu_0 * k_\varepsilon, \mu_1 * k_\varepsilon) \le \mathcal{W}_{\phi,\gamma}(\mu_0, \mu_1) \quad \forall \varepsilon > 0; \tag{5.23}$$

$$\lim_{\varepsilon \downarrow 0} \mathcal{W}_{\phi,\gamma * k_\varepsilon}(\mu_0 * k_\varepsilon, \mu_1 * k_\varepsilon) = \mathcal{W}_{\phi,\gamma}(\mu_0, \mu_1). \tag{5.24}$$

Proof Let $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,1;\mu_0 \to \mu_1)$ be an optimal connecting curve as in Theorem 5.11 and let us set $\mu_t^\varepsilon = \mu_t * k_\varepsilon, \boldsymbol{\nu}_t^\varepsilon := \boldsymbol{\nu}_t * k_\varepsilon$. Since $(\mu^\varepsilon,\boldsymbol{\nu}^\varepsilon) \in \mathcal{CE}(0,1;\mu_0^\varepsilon \to \mu_1^\varepsilon)$, (5.23) then follows by (2.12) whereas (5.24) is a consequence of Theorem 5.6 □

Remark 5.16 (Smooth approximations) For a given curve $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,1)$ the convolution technique of the previous Theorem exhibits an approximations $(\mu^\varepsilon,\boldsymbol{\nu}^\varepsilon)$ in $\mathcal{CE}_{\phi,\gamma^\varepsilon}(0,1)$, $\gamma^\varepsilon := \gamma * k_\varepsilon$, which enjoy some useful properties:

1. $\mu^\varepsilon = \rho^\varepsilon \mathscr{L}^d$, $\boldsymbol{\nu}^\varepsilon = \mathbf{w}^\varepsilon \mathscr{L}^d$ with $\rho^\varepsilon, \mathbf{w}^\varepsilon \in C^\infty(\mathbb{R}^d)$; if $\mu_0(\mathbb{R}^d) < +\infty$ and $\tilde{\mathsf{m}}_{-\kappa}(\gamma) < +\infty$ (recall Theorem 5.8), then ρ^ε is also uniformly bounded.
2. If $\operatorname{supp}(k) \subset \overline{B_1}$ then $\rho^\varepsilon, \mathbf{w}^\varepsilon$ are supported in $G^\varepsilon := \{x \in \mathbb{R}^d : \operatorname{dist}(x,G) \le \varepsilon\}$, $G = \operatorname{supp}(\gamma)$.
3. $\rho^\varepsilon, \mathbf{w}^\varepsilon$ are classical solution of the continuity equation
$$\partial_t \rho^\varepsilon + \nabla \cdot \mathbf{w}^\varepsilon = 0 \quad \text{in } \mathbb{R}^d \times (0,1).$$
4. If $(\mu,\boldsymbol{\nu})$ is also a geodesic, $\int_{\mathbb{R}^d} \phi(\rho_t^\varepsilon, \mathbf{w}_t^\varepsilon)\,\mathrm{d}\gamma^\varepsilon \le \boldsymbol{\Phi}(\rho_t,\boldsymbol{\nu}_t|\gamma) = \mathcal{W}^p_{\phi,\gamma}(\mu_0,\mu_1)$ for every $t \in [0,1]$.

5.3 Absolutely continuous curves and geodesics

We now study absolutely continuous curves with respect to $\mathcal{W}_{\phi,\gamma}$ and their length. Let us first recall (see e.g. [3, Chap. 1]) that a curve $t \mapsto \mu_t \in \mathcal{M}_{\rm loc}(\mathbb{R}^d)$, $t \in [0,T]$, is absolutely continuous w.r.t. $\mathcal{W}_{\phi,\gamma}$ if there exists a function $m \in L^1(0,T)$ such that

$$\mathcal{W}_{\phi,\gamma}(\mu_{t_1},\mu_{t_0}) \le \int_{t_0}^{t_1} m(t)\,\mathrm{d}t \quad \forall\, 0 \le t_0 < t_1 \le T. \tag{5.25}$$

The curve has finite p-energy if moreover $m \in L^p(0,T)$. The metric derivative $|\mu'|$ of an absolutely continuous curve is defined as

$$|\mu_t'| := \lim_{h\to 0} \frac{\mathcal{W}_{\phi,\gamma}(\mu_{t+h},\mu_t)}{|h|}, \tag{5.26}$$

and it is possible to prove that $|\mu_t'|$ exists and satisfies $|\mu_t'| \le m(t)$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$. The length of μ is then defined as the integral of $|\mu'|$ in the interval $(0,T)$.

Theorem 5.17 (Absolutely continuous curves and their metric velocity) *A curve $t \mapsto \mu_t$, $t \in [0,T]$, is absolutely continuous with respect to $\mathcal{W}_{\phi,\gamma}$ if and only if there exists a Borel family of measures $(\boldsymbol{\nu}_t)_{t\in(0,T)}$ in $\mathcal{M}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)$ such that $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,T)$ and*

$$\int_0^T \Big(\boldsymbol{\Phi}(\mu_t,\boldsymbol{\nu}_t|\gamma)\Big)^{1/p}\,\mathrm{d}t < +\infty. \tag{5.27}$$

In this case we have

$$|\mu_t'|^p \le \boldsymbol{\Phi}(\mu_t,\boldsymbol{\nu}_t|\gamma) \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T), \tag{5.28}$$

and there exists a unique Borel family $\tilde{\boldsymbol{\nu}}_t$ such that

$$|\mu_t'|^p = \boldsymbol{\Phi}(\mu_t,\tilde{\boldsymbol{\nu}}_t|\gamma) \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T). \tag{5.29}$$

Proof One implication is trivial: if $(\mu,\boldsymbol{\nu})\in\mathcal{CE}_{\phi,\gamma}(0,T)$ and (5.27) holds, then (5.5) yields

$$\mathcal{W}_{\phi,\gamma}(\mu_{t_1},\mu_{t_0})\le\int_{t_0}^{t_1}\Big(\Phi(\mu_t,\boldsymbol{\nu}_t|\gamma)\Big)^{1/p}\,\mathrm{d}t, \tag{5.30}$$

so that μ is absolutely continuous and (5.28) holds.

Conversely, let us assume that μ is an absolutely continuous curve with length L. A standard reparametrization results [3, Lemma 1.1.4] shows that it is not restrictive to assume that μ is a Lipschitz map. We fix an integer $N>0$, a step size $\tau:=2^{-N}T$, and a family of geodesics $(\mu^{k,N},\boldsymbol{\nu}^{k,N})\in\mathcal{CE}_{\phi,\gamma}((k-1)\tau,k\tau;\mu_{(k-1)\tau}\to\mu_{k\tau})$, $k=1,\cdots,2^N$, such that

$$\tau\Phi(\mu_t^{k,N},\boldsymbol{\nu}_t^{k,N}|\gamma)=\tau^{1-p}W^p_{\phi,\gamma}(\mu_{(k-1)\tau},\mu_{k\tau})\le\int_{(k-1)\tau}^{k\tau}|\mu_t'|^p\,\mathrm{d}t. \tag{5.31}$$

Let $(\mu^N,\boldsymbol{\nu}^N)\in\mathcal{CE}_{\phi,\gamma}(0,T)$ be the curve obtained by gluing together all the geodesics $(\mu^{k,N},\boldsymbol{\nu}^{k,N})$. Applying Corollary 4.10, we can find a subsequence $(\mu^{N_h},\boldsymbol{\nu}^{N_h})$ and a couple $(\mu,\boldsymbol{\nu})\in\mathcal{CE}_{\phi,\gamma}(0,T)$ such that $\mu_t^{N_h}\rightharpoonup^*\tilde\mu_t$ for every $t\in[0,T]$ and $\boldsymbol{\nu}^{N_h}\rightharpoonup^*\boldsymbol{\nu}$ in $\mathcal{M}_{\rm loc}(\mathbb{R}^d\times(0,T);\mathbb{R}^d)$. It is immediate to check that $\mu_t\equiv\tilde\mu_t$ for every $t\in[0,T]$ and

$$\int_0^T|\mu_t'|^p\,\mathrm{d}t\ge\liminf_{h\uparrow+\infty}\int_0^T\Phi(\mu_t^{N_h},\boldsymbol{\nu}_t^{N_h}|\gamma)\,\mathrm{d}t\ge\int_0^T\Phi(\mu_t,\boldsymbol{\nu}_t|\gamma)\,\mathrm{d}t\ge\int_0^T|\tilde\mu_t'|^p\,\mathrm{d}t,$$

which concludes the proof. □

Corollary 5.18 (Geodesics) *For every $\mu\in\mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ the space $\mathcal{M}_{\phi,\gamma}[\mu]$ is a geodesic space, i.e. every couple $\mu_0,\mu_1\in\mathcal{M}_{\phi,\gamma}[\mu]$ can be connected by a (minimal, constant speed) geodesic $t\in[0,1]\mapsto\mu_t\in\mathcal{M}_{\phi,\gamma}[\mu]$ such that*

$$\mathcal{W}_{\phi,\gamma}(\mu_s,\mu_t)=|t-s|\,\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1)\quad\forall s,t\in[0,1]. \tag{5.32}$$

All the (minimal, constant speed) geodesics satisfies the continuity equation (4.1) *for a Borel family of vector valued measures $(\boldsymbol{\nu}_t)_{t\in(0,1)}$ such that*

$$\Phi(\mu_t,\boldsymbol{\nu}_t|\gamma)=\mathcal{W}^p_{\phi,\gamma}(\mu_0,\mu_1)\quad\textit{for a.e. } t\in(0,1). \tag{5.33}$$

If ϕ is strictly convex and sublinear, geodesics are unique.

Remark 5.19 (A formal differential characterization of geodesics) Arguing as in [27, Chap. 3], it would not be difficult to show that a geodesic $\mu_t=\rho_t\mathscr{L}^d$ with respect to $W_{2,\alpha;\mathscr{L}^d}$ should satisfy the system of nonlinear PDE's in $\mathbb{R}^d\times(0,1)$

$$\begin{cases}\partial_t\rho+\nabla\cdot(\rho^\alpha\nabla\psi)=0,\\ \partial_t\psi+\dfrac{\alpha}{2}\rho^{\alpha-1}\big|\nabla\psi\big|^2=0,\end{cases}$$

for some potential ψ. Unlike the Wasserstein case, however, the two equations are coupled, and it is not possible to solve the second Hamilton-Jacobi equation in ψ independently of the first scalar cosnervation law. In the present paper, we do not explore this direction.

We can give a more precise description of the vector measure $\tilde{\boldsymbol{\nu}}$ satisfying the optimality condition (5.29). For every measure $\mu \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ we set

$$\begin{aligned}\mathrm{Tan}_{\phi,\gamma}(\mu) := \Big\{ \boldsymbol{\nu} \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d) : \Phi(\mu,\boldsymbol{\nu}|\gamma) < +\infty, \\ \Phi(\mu,\boldsymbol{\nu}|\gamma) \le \Phi(\mu,\boldsymbol{\nu}+\boldsymbol{\eta}|\gamma) \quad \forall \boldsymbol{\eta} \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d) : \nabla\cdot\boldsymbol{\eta} = 0 \Big\}.\end{aligned} \tag{5.34}$$

Observe that for every $\boldsymbol{\nu} \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)$ such that $\Phi(\mu,\boldsymbol{\nu}|\gamma) < +\infty$ there exists a unique $\tilde{\boldsymbol{\nu}} := \Pi(\boldsymbol{\nu}) \in \mathrm{Tan}_{\phi,\gamma}(\mu)$ such that $\nabla\cdot(\tilde{\boldsymbol{\nu}}-\boldsymbol{\nu}) = 0$. In fact, the set $K(\boldsymbol{\nu}) := \big\{\boldsymbol{\nu}' \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d) : \nabla(\boldsymbol{\nu}'-\boldsymbol{\nu}) = 0\big\}$ is weakly* closed and, by the estimate (3.47) the sublevels of the functional $\boldsymbol{\nu}' \mapsto \Phi(\mu,\boldsymbol{\nu}'|\gamma)$ are weakly* relatively compact. Therefore, a minimizer $\tilde{\boldsymbol{\nu}}$ exists and it is also unique, being $\Phi(\mu,\cdot|\gamma)$ strictly convex.

Corollary 5.20 *Let $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}_{\phi,\gamma}(0,T)$ so that μ is absolutely continuous w.r.t. $\mathcal{W}_{\phi,\gamma}$. The vector measure $\boldsymbol{\nu}$ satisfies the optimality condition* (5.29) *if and only if $\boldsymbol{\nu}_t \in \mathrm{Tan}_{\phi,\gamma}(\mu_t)$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$.*

Let us consider the particular case of Example 3.5 in the case of a differentiable norm $\|\cdot\|$ with associated duality map $j_1 = \mathrm{D}\|\cdot\|$. We denote by $j_p(\mathbf{w}) = \|\mathbf{w}\|^{p-2} j_1(\mathbf{w})$ the p-duality map, i.e. the differential of $\frac{1}{p}\|\cdot\|^p$ and we suppose that the concave function $h:[0,+\infty)\to[0,+\infty)$ satisfies

$$\lim_{r\downarrow 0} h(r) = \lim_{r\uparrow+\infty} r^{-1}h(r) = 0. \tag{5.35}$$

For every nonnegative Radon measure $\mu \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ whose support is a subset of $\mathrm{supp}(\gamma)$, we define the Radon measure $h(\mu|\gamma)$ by

$$h(\mu|\gamma) := h(\rho)\cdot\gamma \quad \text{where } \rho := \frac{\mathrm{d}\mu}{\mathrm{d}\gamma}. \tag{5.36}$$

Observe that $h(\mu|\gamma) \ll \gamma$ even if μ is singular w.r.t. γ.

Theorem 5.21 *Let $\mu \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ and ϕ as in* (3.23) *with h satisfying* (5.35)*. A vector measure $\boldsymbol{\nu}$ satisfies $\Phi(\mu,\boldsymbol{\nu}|\gamma) < +\infty$ iff $\boldsymbol{\nu} = \mathbf{v}\,h(\mu|\gamma)$ for some vector field $\mathbf{v} \in L^p_{h(\mu|\gamma)}(\mathbb{R}^d;\mathbb{R}^d)$. Moreover, $\boldsymbol{\nu} \in \mathrm{Tan}_{\phi,\gamma}(\mu)$ if and only if the vector field $\mathbf{v}$ satisfies*

$$j_p(\mathbf{v}) \in \overline{\big\{\nabla\zeta : \zeta \in C^\infty_c(\mathbb{R}^d)\big\}}^{L^p_{h(\mu|\gamma)}(\mathbb{R}^d;\mathbb{R}^d)}. \tag{5.37}$$

Proof Being h sublinear, the functional Φ admits the representation

$$\Phi(\mu,\boldsymbol{\nu}|\gamma) = \int_{\mathbb{R}^d} \phi(\rho,\mathbf{w})\,\mathrm{d}\gamma = \int_{\mathbb{R}^d} h(\rho)\|\mathbf{w}/h(\rho)\|^p\,\mathrm{d}\gamma = \int_{\mathbb{R}^d} \|\mathbf{v}\|^p\,\mathrm{d}h(\mu|\gamma), \tag{5.38}$$

where $\mu = \rho\gamma + \mu^\perp$ and $\boldsymbol{\nu} = \mathbf{w}\gamma = h(\rho)\mathbf{v}\gamma$. The condition $\boldsymbol{\nu} = \mathbf{v}h(\mu|\gamma) \in \mathrm{Tan}_{\phi,\gamma}(\mu)$ is then equivalent to

$$\int_{\mathbb{R}^d} \|\mathbf{v}\|^p\,\mathrm{d}h(\mu|\gamma) \le \int_{\mathbb{R}^d} \|\mathbf{v}+\mathbf{z}\|^p\,\mathrm{d}h(\mu|\gamma) \qquad \forall \mathbf{z} \in L^p(h(\mu|\gamma)) : \nabla\cdot\big(\mathbf{z}\,h(\mu|\gamma)\big) = 0.$$

Thanks to the convexity of $||\cdot||^p$, the previous condition is equivalent to

$$\int_{\mathbb{R}^d} j_p(\mathbf{v})\cdot\mathbf{z}\,\mathrm{d}h(\mu|\gamma) = 0 \quad \forall \mathbf{z} \in L^p_{h(\mu|\gamma)}(\mathbb{R}^d;\mathbb{R}^d) : \quad \int_{\mathbb{R}^d} \mathbf{z}\cdot\nabla\zeta\,\mathrm{d}h(\mu|\gamma) = 0, \tag{5.39}$$

i.e. $j_p(\mathbf{v})$ belongs to the closure of $\{\nabla\zeta : \zeta \in C^\infty_c(\mathbb{R}^d)\}$ in $L^p_{h(\mu|\gamma)}(\mathbb{R}^d;\mathbb{R}^d)$. □

5.4 **Comparison with Wasserstein and $\dot W^{-1,p}$ distances.**

Theorem 5.22 *If $\gamma(\mathbb{R}^d) < +\infty$ then for every $\mu_0, \mu_1 \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ and $\alpha < 1$ we have*

$$W_{p/\theta}(\mu_0,\mu_1) \le W_{p/\theta,1;\gamma}(\mu_0,\mu_1) \le \gamma(\mathbb{R}^d)^{1/\kappa} W_{p,\alpha;\gamma}(\mu_0,\mu_1), \tag{5.40}$$

where, as usual, $\theta = (1-\alpha)p + \alpha$.

Proof Let $(\mu, \boldsymbol{\nu}) \in \mathcal{CE}_{\phi_{p,\alpha},\gamma}(0,1;\mu_0 \to \mu_1)$ be an optimal curve, so that

$$W^p_{p,\alpha;\gamma}(\mu_0,\mu_1) = \int_0^1 \Phi_{p,\alpha}(\mu_t, \boldsymbol{\nu}_t)\,\mathrm{d}t = \int_0^1 \int_{\mathbb{R}^d} (\rho_t)^{\theta-p}\,|\mathbf{w}_t|^p\,\mathrm{d}\gamma\mathrm{d}t, \tag{5.41}$$

where $\mu_t := \rho_t\gamma + \mu_t^\perp$, $\boldsymbol{\nu}_t = \mathbf{w}_t\gamma \ll \gamma$. Hölder inequality yields

$$W^{p/\theta}_{p/\theta,1;\gamma}(\mu_0,\mu_1) \le \int_0^1 \int_{\mathbb{R}^d} (\rho_t)^{1-p/\theta}\,|\mathbf{w}_t|^{p/\theta}\,\mathrm{d}\gamma\mathrm{d}t \le \gamma(\mathbb{R}^d)^{1-1/\theta} W^{p/\theta}_{p,\alpha;\gamma}(\mu_0,\mu_1). \quad \square$$

Theorem 5.23 *Let us suppose that $\tilde{\mathsf{m}}_{-k}(\gamma) < +\infty$, $\kappa = p/(\theta-1) = q/(1-\alpha)$, let $\mu_0, \mu_1 \in \mathcal{P}(\mathbb{R}^d)$, and let $\kappa^* = \kappa/(\kappa-1)$ be the Hölder's conjugate exponent of κ. Then*

$$\|\mu_0 - \mu_1\|_{\dot W_\gamma^{-1,\kappa^*}} = W_{\kappa^*,0;\gamma}(\mu_0,\mu_1) \le W_{p,\alpha;\gamma}(\mu_0,\mu_1). \tag{5.42}$$

Proof We keep the same notation of the previous Theorem, setting

$$\tau = p/r := 1 + p - \theta, \quad \tau^* := \frac{\tau}{\tau - 1} = 1 + \frac{1}{p-\theta}, \quad x = (\tau^*)^{-1} = \frac{p-\theta}{1+p-\theta}.$$

Observing that $\mu_t \in \mathcal{P}(\mathbb{R}^d)$ thanks to Theorem 4.9, we obtain

$$\begin{aligned} W^r_{r,0;\gamma}(\mu_0,\mu_1) &\le \int_0^1 \int_{\mathbb{R}^d} |\mathbf{w}_t|^r\,\mathrm{d}\gamma\mathrm{d}t = \int_0^1 \int_{\mathbb{R}^d} \big(\rho_t\big)^x \big(\rho_t\big)^{-x} |\mathbf{w}_t|^r\,\mathrm{d}\gamma\mathrm{d}t \\ &\le \int_0^1 \Big(\int_{\mathbb{R}^d} \rho^{-x\tau} |\mathbf{w}_t|^{r\tau}\,\mathrm{d}\gamma\Big)^{1/\tau} \mathrm{d}t = \Big(\int_0^1 \int_{\mathbb{R}^d} \rho^{\theta-p} |\mathbf{w}_t|^p\,\mathrm{d}\gamma\mathrm{d}t\Big)^{1/\tau} = W^r_{p,\alpha;\gamma}(\mu_0,\mu_1). \quad \square \end{aligned}$$

Theorem 5.24 (Comparison with W_p) *Assume that $\gamma \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ is a bounded perturbation of a log-concave measure (e.g. $\gamma = f e^{-V}\mathscr{L}^d$, where V is a convex function and f nonnegative and bounded). If $\mu_i = s_i\gamma \in \mathcal{P}(\mathbb{R}^d)$ with $s_i \in L^\infty(\gamma)$ and $\mathsf{m}_p(\mu_i) \le L < +\infty$ then $\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) < +\infty$ and there exists a constant C only depending on L, ϕ, and γ such that*

$$\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) \le \mathsf{C}\, W_p(\mu_0,\mu_1). \tag{5.43}$$

Proof It is not restrictive to assume that γ is log-concave. We can then consider the optimal plan $\Sigma \in \mathcal{M}^+(\mathbb{R}^d \times \mathbb{R}^d)$ induced by the p-Wasserstein distance (1.1) between μ_0 and μ_1 and the interpolant μ_t defined as

$$\mu_t(A) = \Sigma\big(\{(x_0,x_1) \in \mathbb{R}^d \times \mathbb{R}^d : (1-t)x_0 + t x_1 \in A\}\big) \quad \forall A \in \mathcal{B}(\mathbb{R}^d). \tag{5.44}$$

It is possible to prove (see e.g. [3, Theorems 7.2.2, 8.3.1, 9.4.12]) that μ_t is the geodesic interpolant between μ_0 and μ_1, it satisfies the continuity equation

$$\partial_t \mu_t + \nabla \cdot \boldsymbol{\nu}_t = 0 \quad \text{in } \mathbb{R}^d \times (0,1)$$

with respect to a vector valued measure $\boldsymbol{\nu}_t = \mathbf{v}_t\mu_t \ll \mu_t$ where the vector field $\mathbf{v}_t$ satisfies

$$\int_0^1 \Phi_p(\mu_t, \boldsymbol{\nu}_t)\,\mathrm{d}t = \int_0^1 \int_{\mathbb{R}^d} |\mathbf{v}_t(x)|^p\,\mathrm{d}\mu_t(x)\,\mathrm{d}t = W_p^p(\mu_0,\mu_1),$$

and finally $\mu_t = s_t\gamma$ with $\|s_t\|_{L^\infty(\gamma)} \le L := \max\big(\|s_0\|_{L^\infty(\gamma)}, \|s_1\|_{L^\infty(\gamma)}\big)$. Observe that, being $s_t(x) \le L$ for γ-a.e. $x \in \mathbb{R}^d$ and $\phi(0,0) = 0$, Theorem 3.1 yields

$$\phi(s_t, s_t\mathbf{v}_t) \le \frac{s_t}{L}\phi(L, L\mathbf{v}_t) \le \mathsf{C}_L s_t|\mathbf{v}_t|^p \quad \gamma\text{-a.e.},$$

so that

$$\Phi(\mu_t, \boldsymbol{\nu}_t) = \int_{\mathbb{R}^d} \phi(s_t, s_t\mathbf{v}_t)\,\mathrm{d}\gamma(x) \le \mathsf{C}_L \int_{\mathbb{R}^d} |\mathbf{v}_t|^p s_t\,\mathrm{d}\gamma = \mathsf{C}_L \int_{\mathbb{R}^d} |\mathbf{v}_t|^p\,\mathrm{d}\mu_t,$$

and therefore

$$\mathcal{W}^p_{\phi,\gamma}(\mu_0,\mu_1) \le \int_0^1 \Phi(\mu_t, \boldsymbol{\nu}_t)\,\mathrm{d}t \le \mathsf{C}_L \int_0^1 \int_{\mathbb{R}^d} |\mathbf{v}_t|^p\,\mathrm{d}\mu_t\,\mathrm{d}t = \mathsf{C}_L W_p^p(\mu_0,\mu_1). \quad \square$$

Corollary 5.25 *If $\mu_i = s_i\mathscr{L}^d \in \mathcal{P}(\mathbb{R}^d)$ have L^∞-densities with compact support (or, more generally, finite p-momentum), then $\mathcal{W}_{\phi,\mathscr{L}^d}(\mu_0,\mu_1) < +\infty$.*

Theorem 5.26 *If $\mu_i = s_i\gamma$ with $s_i \ge L > 0$ γ-a.e. in $\mathbb{R}^d$, then there exists a constant C depending on L and ϕ such that*

$$\mathcal{W}_{\phi,\gamma}(\mu_0,\mu_1) \le C_L \|\mu_0 - \mu_1\|_{\dot{W}_\gamma^{-1,p}}. \tag{5.45}$$

Proof Let us first observe that if $\|\mu_0 - \mu_1\|_{\dot{W}_\gamma^{-1,p}} < +\infty$ then there exists $\mathbf{w} \in L^p_\gamma(\mathbb{R}^d;\mathbb{R}^d)$ such that

$$-\nabla\cdot\boldsymbol{\nu} = \mu_1 - \mu_0, \quad \boldsymbol{\nu} := \mathbf{w}\gamma, \quad \int_{\mathbb{R}^d} |\mathbf{w}|^p\,\mathrm{d}\gamma = \|\mu_0 - \mu_1\|^p_{\dot{W}_\gamma^{-1,p}}. \tag{5.46}$$

In fact, in the Banach space $X := L^q_\gamma(\mathbb{R}^d;\mathbb{R}^d)$ we can consider the linear space $Y := \big\{\mathrm{D}\zeta : \zeta \in C^1_c(\mathbb{R}^d)\big\}$ and the linear functional

$$\langle \ell, \mathbf{y}\rangle := \int_{\mathbb{R}^d} \zeta\,\mathrm{d}(\mu_1 - \mu_0) \quad \text{if } \mathbf{y} = \mathrm{D}\zeta \quad \text{for some } \zeta \in C^1_c(\mathbb{R}^d).$$

ℓ is well defined and satisfies $\big|\langle \ell, \mathbf{y}\rangle\big| \le \|\mu_0 - \mu_1\|_{\dot{W}_\gamma^{-1,p}} \|\mathbf{y}\|_{L^q_\gamma(\mathbb{R}^d;\mathbb{R}^d)}$ for every $\mathbf{y} \in Y$. Hahn-Banach Theorem and Riesz representation Theorem yield the existence of $\mathbf{w} \in L^p(\mathbb{R}^d;\mathbb{R}^d)$ such that $\langle \ell, \mathbf{y}\rangle = \int_{\mathbb{R}^d} \mathbf{w}\cdot\mathbf{y}\,d\gamma$, which yields (5.46). Setting $\mu_t = (1-t)\mu_0 + t\mu_1$, it is then immediate to check that $(\mu_t, \boldsymbol{\nu}) \in \mathcal{CE}(0,1;\mu_0 \to \mu_1)$; we can then compute

$$\mathcal{W}^p_{\phi,\gamma}(\mu_0,\mu_1) \le \int_0^1 \int_{\mathbb{R}^d} \phi((1-t)s_0 + ts_1, \mathbf{w})\,\mathrm{d}\gamma\mathrm{d}t \le \int_{\mathbb{R}^d} \phi(L, \mathbf{w})\,\mathrm{d}\gamma \le C_L \int_{\mathbb{R}^d} |\mathbf{w}|^p\,\mathrm{d}\gamma,$$

where we used the fact that $(1-t)s_0 + ts_1 \ge L$ γ-almost everywhere and the map $\rho \mapsto \phi(\rho, \mathbf{w})$ is nonincreasing. $\square$

5.5 The case $\gamma = \mathscr{L}^d$ and the Heat equation as gradient flow

One of the most interesting cases corresponds to the choice

$$\gamma := \mathscr{L}^d, \quad h_\alpha(\rho) := \rho^\alpha,\ 0 < \alpha < 1, \qquad \phi_{p,\alpha}(\rho,\mathbf{w}) := \rho^\alpha |\mathbf{w}/\rho^\alpha|^p. \tag{5.47}$$

In this case the expression of the weighted Wasserstein distance becomes

$$W^p_{p,\alpha;\mathscr{L}^d}(\mu_0,\mu_1) := \min\Big\{ \int_0^1 \int_{\mathbb{R}^d} \rho_t^\alpha |\mathbf{v}_t|^p \,\mathrm{d}x\,\mathrm{d}t :\ \partial_t\mu + \nabla\cdot(\rho^\alpha \mathbf{v}) = 0 \text{ in } \mathbb{R}^d\times(0,1)$$
$$\mu_t = \rho_t\mathscr{L}^d + \mu_t^\perp, \qquad \mu_{|_{t=0}} = \mu_0,\ \mu_{|_{t=1}} = \mu_1 \Big\}.$$

The metric $W_{p,\alpha;\mathscr{L}^d}$ restricted to $\mathcal{P}(\mathbb{R}^d)$ is complete if $d < \kappa = \frac{p}{\theta-1} = \frac{q}{1-\alpha}$.

Remark 5.27 ($\mathcal{P}(\mathbb{R}^d)$ is not complete w.r.t. $W_{p,\alpha;\mathscr{L}^d}$ if $d > \kappa$) The above condition is almost sharp; here is a simple counterexample in the case $d > \kappa$. We consider an initial probability measure with compact support $\mu_0 = \rho_0\mathscr{L}^d$, $\rho_0 \in L^\infty(\mathbb{R}^d)$, and, for $t \ge 0$, the family

$$\mu_t := \rho_t\mathscr{L}^d, \quad \rho_t(x) := e^{-dt}\rho_0(e^{-t}x), \quad \boldsymbol{\nu}_t := x\mu_t = x\rho_t(x)\mathscr{L}^d. \tag{5.48}$$

It is easy to check that $(\mu,\boldsymbol{\nu}) \in \mathcal{CE}(0,+\infty)$, $\mu_t(\mathbb{R}^d) = 1$. Evaluating the functional $\Phi_t := \Phi_{p,\alpha}(\mu_t,\boldsymbol{\nu}_t|\mathscr{L}^d)$ we get

$$\Phi_t = \int_{\mathbb{R}^d} \rho_t^{\theta-p}|\rho_t x|^p \,\mathrm{d}x = \int_{\mathbb{R}^d} e^{-d\theta t}\rho_0^\theta(e^{-dt}x)|x|^p\,\mathrm{d}x$$
$$= e^{dt-d\theta t+pt}\int_{\mathbb{R}^d} e^{-dt}\rho_0^\theta(e^{-dt}x)|e^{-dt}x|^p\,\mathrm{d}x = e^{(d(1-\theta)+p)t}\int_{\mathbb{R}^d}\rho_0^\theta(y)|y|^p\,\mathrm{d}y$$

so that

$$\Phi_t^{1/p} = \mathsf{c}\, e^{(1-d/\kappa)t}, \qquad \int_0^{+\infty}\Phi_t^{1/p}\,\mathrm{d}t = \mathsf{c}\frac{\kappa}{d-\kappa} < +\infty \quad \text{if } d > \kappa.$$

If $d > \kappa$ we obtain a curve $t \mapsto \mu_t \in \mathcal{P}(\mathbb{R}^d)$ of finite length w.r.t. $W_{p,\alpha;\mathscr{L}^d}$ (in particular $(\mu_n)_{n\in\mathbb{N}}$ is a Cauchy sequence) such that $\lim_{t\uparrow+\infty}\mu_t = 0$ in the weak* topology.

In the remaining part of this section, we want to study the properties of $W_{p,\alpha;\mathscr{L}^d}$ with respect to the heat flow. We thus introduce

$$g(x) = g_1(x) = \frac{1}{(4\pi)^{d/2}}e^{-|x|^2/4}, \qquad g_t(x) := \frac{1}{(4\pi t)^{d/2}}e^{-|x|^2/4t} = t^{-d/2}g_1(x/\sqrt{t}),$$

and for every $\mu \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ with $\tilde{\mathsf{m}}_\delta(\mu) < +\infty$ for some $\delta \le 0$, we set

$$\mathcal{S}_t[\mu] = \mu * g_t = u_t\mathscr{L}^d, \quad u_t(x) = S_t[\mu](x) := \int_{\mathbb{R}^d} g_t(x-y)\,\mathrm{d}\mu(y). \tag{5.49}$$

It is well known that $u \in C^\infty(\mathbb{R}^d\times(0,+\infty))$ and

$$\partial_t u - \Delta u = 0 \quad \text{in } \mathbb{R}^d\times(0,+\infty), \qquad \mathcal{S}_t[\mu] \rightharpoonup^* \mu \quad \text{as } t\downarrow 0. \tag{5.50}$$

Theorem 5.28 (Contraction property) *Let $\mu^0,\mu^1 \in \mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ with $\tilde{\mathsf{m}}_\delta(\mu^i) < +\infty$ and $\mathcal{W}_{\phi,\mathscr{L}^d}(\mu^1,\mu^2) < +\infty$. If $\mu^i_t := \mathcal{S}_t[\mu^i]$ are the corresponding solutions of the heat flow, then*

$$\mathcal{W}_{\phi,\mathscr{L}^d}(\mu^1_t,\mu^2_t) \le \mathcal{W}_{\phi,\mathscr{L}^d}(\mu^1,\mu^2) \quad \forall t > 0. \tag{5.51}$$

Proof It sufficient to approximate the Gaussian kernel g by a family of C^∞ kernels k^n with compact support and then apply Theorem 5.15, observing that $k^n * \mathscr{L}^d = \mathscr{L}^d$. □

We consider now the particular case of the $W_{2,\alpha;\mathscr{L}^d}$ weighted distance with $\alpha > 1-2/d$. Let us first introduce the convex density function (recall (1.9))

$$\psi_\alpha(\rho) := \frac{1}{(2-\alpha)(1-\alpha)}\rho^{2-\alpha}, \quad \text{such that} \quad \psi''_\alpha(\rho) = \frac{1}{h(\rho)} = \rho^{-\alpha}, \tag{5.52}$$

and the corresponding entropy functional

$$\Psi_\alpha(\mu) = \Psi_\alpha(\mu|\mathscr{L}^d) := \int_{\mathbb{R}^d} \psi_\alpha(\rho)\,\mathrm{d}x, \quad \text{if } \mu = \rho\mathscr{L}^d \ll \mathscr{L}^d. \tag{5.53}$$

We also introduce the set $\mathcal{Q} := \big\{\mu \in \mathcal{P}(\mathbb{R}^d) : \Psi(\mu) < +\infty\big\}$.

Theorem 5.29 *If $\mu \in \mathcal{P}(\mathbb{R}^d)$ then $\mu_t = \mathcal{S}_t[\mu] = u_t\mathscr{L}^d \in \mathcal{Q}$ for every $t > 0$, the map $t \mapsto \Psi_\alpha(\mu_t)$ is nonincreasing, and it satisfies the energy identity*

$$\Psi_\alpha(\mu_t) + \int_s^t \Phi_{2,\alpha}(u_r, \nabla u_r)\,\mathrm{d}r = \Psi_\alpha(\mu_s) \quad \forall\, 0 < s \le t < +\infty; \tag{5.54}$$

when $\mu \in \mathcal{Q}$ then the previous identity holds even for $s = 0$. Moreover, μ_t satisfies the Evolution Variational Inequality

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t} W^2_{2,\alpha;\mathscr{L}^d}(\mu_t, \sigma) + \Psi_\alpha(\mu_t) \le \Psi_\alpha(\sigma) \quad \forall t \ge 0,\ \forall \sigma \in \mathcal{Q}. \tag{5.55}$$

Proof Since $\psi''_\alpha(u) = u^{-\alpha}$, a direct computation shows

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} \psi_\alpha(u_t)\,\mathrm{d}x = -\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} \nabla u_t \cdot \nabla \psi'_\alpha(u_t)\,\mathrm{d}x = \int_{\mathbb{R}^d} |\nabla u_t|^2 u_t^{-\alpha}\,\mathrm{d}x = \Phi_{2,\alpha}(u_t, \nabla u_t).$$

Concerning (5.55), we use the technique introduced by [12, § 2]: we consider a geodesic $(\sigma_s, \boldsymbol{\nu}_s)_{s\in[0,1]} \in \mathcal{CE}(0,1;\sigma \to \mu)$, which satisfies $\sigma_s(\mathbb{R}^d) = 1$ by Theorem 4.9. We set

$$\sigma^\varepsilon_{s,t} = u^\varepsilon_{s,t}\mathscr{L}^d := \mathcal{S}_{\varepsilon+st}[\sigma_s], \quad \tilde{\boldsymbol{\nu}}^\varepsilon_{s,t} = \tilde{\mathbf{w}}^\varepsilon_{s,t}\mathscr{L}^d := \mathcal{S}_{\varepsilon+st}[\boldsymbol{\nu}_s], \quad \mathbf{w}^\varepsilon_{s,t} := \tilde{\mathbf{w}}^\varepsilon_{s,t} - t\nabla u^\varepsilon_{s,t}.$$

It is not difficult to check that

$$\partial_s u^\varepsilon_{s,t} + \nabla \cdot \mathbf{w}^\varepsilon_{s,t} = 0 \quad \text{in } \mathbb{R}^d \times (0,1), \tag{5.56}$$

so that

$$W^2_{2,\alpha;\mathscr{L}^d}(\mu_{\varepsilon+t}, \sigma) \le \int_0^1 A^\varepsilon_{s,t}\,\mathrm{d}s, \quad A^\varepsilon_{s,t} := \int_{\mathbb{R}^d} \big(u^\varepsilon_{s,t}\big)^{-\alpha} |\mathbf{w}^\varepsilon_{s,t}|^2\,\mathrm{d}x = \Phi_{2,\alpha}(\sigma^\varepsilon_{s,t}, \boldsymbol{\nu}^\varepsilon_{s,t}|\mathscr{L}^d).$$

We thus evaluate

$$\begin{aligned}
A^\varepsilon_{s,t} &= \int_{\mathbb{R}^d} \big(u^\varepsilon_{s,t}\big)^{-\alpha}\Big(-2t\nabla u^\varepsilon_{s,t}\cdot \mathbf{w}^\varepsilon_{s,t} + |\tilde{\mathbf{w}}^\varepsilon_{s,t}|^2 - t^2|\nabla u^\varepsilon_{s,t}|^2\Big)\,\mathrm{d}x \\
&\le -2t\int_{\mathbb{R}^d} \big(u^\varepsilon_{s,t}\big)^{-\alpha}\nabla u^\varepsilon_{s,t}\cdot \mathbf{w}^\varepsilon_{s,t}\,\mathrm{d}x + \int_{\mathbb{R}^d} \big(u^\varepsilon_{s,t}\big)^{-\alpha}|\tilde{\mathbf{w}}^\varepsilon_{s,t}|^2\,\mathrm{d}x \\
&\le -2t\,\partial_s \Psi(\sigma^\varepsilon_{s,t}) + \Phi_{2,\alpha}(\sigma_s, \boldsymbol{\nu}_s|\mathscr{L}^d),
\end{aligned} \tag{5.57}$$

where we used the facts

$$\partial_s \int_{\mathbb{R}^d} \psi_\alpha(u^\varepsilon_{s,t})\,\mathrm{d}x \overset{(5.56)}{=} \int_{\mathbb{R}^d} \nabla\psi'_\alpha(u^\varepsilon_{s,t})\cdot \mathbf{w}^\varepsilon_{s,t}\,\mathrm{d}x \overset{(5.52)}{=} \int_{\mathbb{R}^d} \big(u^\varepsilon_{s,t}\big)^{-\alpha}\nabla u^\varepsilon_{s,t}\cdot \mathbf{w}^\varepsilon_{s,t}\,\mathrm{d}x,$$

$$\int_{\mathbb{R}^d} \big(u^\varepsilon_{s,t}\big)^{-\alpha}|\tilde{\mathbf{w}}^\varepsilon_{s,t}|^2\,\mathrm{d}x = \Phi_{2,\alpha}(\sigma_s * g_{\varepsilon+st}, \boldsymbol{\nu}_s * g_{\varepsilon+st}|\mathscr{L}^d) \le \Phi_{2,\alpha}(\sigma_s, \boldsymbol{\nu}_s|\mathscr{L}^d)$$

thanks to the convolution contraction property of Theorem 2.3. Integrating (5.57) with respect to s from 0 to 1 and recalling that $(\sigma_s, \boldsymbol{\nu}_s)_{s\in[0,1]}$ is a minimal geodesic and that $\sigma^\varepsilon_{1,t} = \mu_{\varepsilon+t}$ and $\sigma^\varepsilon_{0,t} = \sigma$, we get

$$\int_0^1 A^\varepsilon_{s,t}\,\mathrm{d}s + 2t\,\Psi_\alpha(\mu_{\varepsilon+t}) \le 2t\,\Psi_\alpha(\sigma) + W^2_{2,\alpha;\mathscr{L}^d}(\mu,\sigma). \tag{5.58}$$

We deduce that

$$\tfrac{1}{2}W^2_{2,\alpha}(\mu_{\varepsilon+t},\sigma) + t\,\Psi(\mu_{\varepsilon+t}) \le t\,\Psi(\sigma) + \tfrac{1}{2}W^2_{2,\alpha}(\mu,\sigma). \tag{5.59}$$

Passing to the limit as $\varepsilon \downarrow 0$ and then as $t \downarrow 0$ after dividing the inequality by t we get (5.55) at $t = 0$. Recalling the semigroup property of the heat equation, we obtain (5.55) for every time $t \ge 0$. □

(5.55) is the metric formulation of the gradient flow of the (geodesically convex) functional Ψ_α in the metric space $(\mathcal{Q}, W_{2,\alpha;\mathscr{L}^d})$, see [3, Chap. 4]. Applying [12, Theorem 3.2] we eventually obtain:

Corollary 5.30 (Geodesic convexity of Ψ_α) *Let $\alpha > 1 - 2/d$, $\mu_i = \rho_i\mathscr{L}^d \in \mathcal{P}(\mathbb{R}^d)$ with $W_{2,\alpha;\mathscr{L}^d}(\mu_0,\mu_1) < +\infty$ and $\int_{\mathbb{R}^d}\rho_i^{2-\alpha}\,\mathrm{d}x < +\infty$, and let μ_t, $t \in [0,1]$, be the minimal speed geodesic connecting μ_0 to μ_1 w.r.t. $W_{2,\alpha;\mathscr{L}^d}$. Then for every $t \in [0,1]$ $\mu_t = \rho_t\mathscr{L}^d \ll \mathscr{L}^d$*

$$\int_{\mathbb{R}^d}\rho_t^{2-\alpha}\,\mathrm{d}x \le (1-t)\int_{\mathbb{R}^d}\rho_0^{2-\alpha}\,\mathrm{d}x + t\int_{\mathbb{R}^d}\rho_1^{2-\alpha}\,\mathrm{d}x. \tag{5.60}$$

Acknowledgements. J.D. and B.N. have been partially supported by the ANR project IFO. B.N. has also been partially supported by the ANR project OTARIE.

References

1. L. AMBROSIO AND G. BUTTAZZO, *Weak lower semicontinuous envelope of functionals defined on a space of measures*, Ann. Mat. Pura Appl. (4), 150 (1988), pp. 311–339.
2. L. AMBROSIO, N. FUSCO, AND D. PALLARA, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, Clarendon Press, Oxford, 2000.
3. L. AMBROSIO, N. GIGLI, AND G. SAVARÉ, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005.
4. L. AMBROSIO AND F. SANTAMBROGIO, *Necessary optimality conditions for geodesics in weighted Wasserstein spaces*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl., 18 (2007), pp. 23–37.
5. L. AMBROSIO, G. SAVARÉ, AND L. ZAMBOTTI, *Existence and stability for Fokker-Planck equations with log-concave reference measure*, ArXiv Mathematics e-prints, (2007).
6. A. ARNOLD, P. MARKOWICH, G. TOSCANI, AND A. UNTERREITER, *On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations*, Comm. Partial Differential Equations, 26 (2001), pp. 43–100.
7. J.-D. BENAMOU AND Y. BRENIER, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math., 84 (2000), pp. 375–393.

8. A. BRANCOLINI, G. BUTTAZZO, AND F. SANTAMBROGIO, *Path functionals over Wasserstein spaces*, J. Eur. Math. Soc. (JEMS), 8 (2006), pp. 415–434.
9. Y. BRENIER, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math., 44 (1991), pp. 375–417.
10. G. BUTTAZZO, C. JIMENEZ, AND E. OUDET, *An optimization problem for mass transportation with congested dynamics*, Preprint, available online at http://cvgmt.sns.it, (2007).
11. J. CARRILLO, S. LISINI, AND G. SAVARÉ, in preparation, (2008).
12. S. DANERI AND G. SAVARÉ, *Eulerian calculus for the displacement convexity in the Wasserstein distance*, Preprint, arXiv:0801.2455v1, ((2008)).
13. E. DE GIORGI, *New problems on minimizing movements*, in Boundary Value Problems for PDE and Applications, C. Baiocchi and J. L. Lions, eds., Masson, 1993, pp. 81–98.
14. C. DELLACHERIE AND P.-A. MEYER, *Probabilities and potential*, vol. 29 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam, 1978.
15. J. DOLBEAULT, B. NAZARET, AND G. SAVARÉ, in preparation, (2008).
16. L. C. EVANS, *Partial differential equations and Monge-Kantorovich mass transfer*, in Current developments in mathematics, 1997 (Cambridge, MA), Int. Press, Boston, MA, 1999, pp. 65–126.
17. L. C. EVANS AND W. GANGBO, *Differential equations methods for the Monge-Kantorovich mass transfer problem*, Mem. Amer. Math. Soc., 137 (1999), pp. viii+66.
18. W. GANGBO AND R. J. MCCANN, *The geometry of optimal transportation*, Acta Math., 177 (1996), pp. 113–161.
19. R. JORDAN, D. KINDERLEHRER, AND F. OTTO, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal., 29 (1998), pp. 1–17 (electronic).
20. P. MARÉCHAL, *On a functional operation generating convex functions. I. Duality*, J. Optim. Theory Appl., 126 (2005), pp. 175–189.
21. ———, *On a functional operation generating convex functions. II. Algebraic properties*, J. Optim. Theory Appl., 126 (2005), pp. 357–366.
22. J. NASH, *C^1 isometric imbeddings.*, Ann. of Math., (1954).
23. ———, *The imbedding problem for Riemannian manifolds.*, Ann. of Math., (1956).
24. F. OTTO, *Doubly degenerate diffusion equations as steepest descent*, Manuscript, (1996).
25. ———, *Evolution of microstructure in unstable porous media flow: a relaxational approach*, Comm. Pure Appl. Math., 52 (1999), pp. 873–915.
26. ———, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations, 26 (2001), pp. 101–174.
27. F. OTTO AND C. VILLANI, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal., 173 (2000), pp. 361–400.
28. S. T. RACHEV AND L. RÜSCHENDORF, *Mass transportation problems. Vol. I*, Probability and its Applications, Springer-Verlag, New York, 1998. Theory.
29. R. ROCCKAFELLAR, *A general correspondence between dual minimax problems and convex problems*, Pacific J. of Math., 25 (1968), pp. 597–611.
30. C. VILLANI, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.
31. C. VILLANI, *Optimal Transport, Old and New*, Springer Verlag, To appear.