Stability of flows associated to gradient vector fields and convergence of iterated transport maps *

Luigi Ambrosio †
Scuola Normale Superiore
Pisa

Stefano Lisini ‡
Dipartimento di Matematica
Università di Pavia

Giuseppe Savaré §
Dipartimento di Matematica
Università di Pavia

November 9, 2005

Abstract

In this paper we address the problem of stability of flows associated to a sequence of vector fields under minimal regularity requirements on the limit vector field, that is supposed to be a gradient.

We apply this stability result to show the convergence of iterated compositions of optimal transport maps arising in the implicit time discretization (with respect to the Wasserstein distance) of nonlinear evolution equations of a diffusion type.

Finally, we use these convergence results to study the gradient flow of a particular class of polyconvex functionals recently considered by Gangbo, Evans ans Savin. We solve some open problems raised in their paper and obtain existence and uniqueness of solutions under weaker regularity requirements and with no upper bound on the jacobian determinant of the initial datum.

Contents

*This work was partially supported by grants of M.I.U.R. and of IMATI-CNR, Pavia, Italy.

†l.ambrosio@sns.it
‡stefano.lisini@unipv.it
§giuseppe.savare@unipv.it

1 Introduction

Let us assume that we are given a functional ϕ defined on $\mathscr{P}_2(\mathscr{V})$, the space of probability measures in a open set $\mathscr{V} \subset \mathbb{R}^d$ with finite quadratic moments, and let us consider the variational formulation of the (Euler) implicit time discretization of the gradient flow of ϕ with respect to the Kantorovich-Rubinstein-Wasserstein metric W_2: namely, given a time step $\tau > 0$ and an initial datum $\bar{\mu}$, we consider the sequence (μ^k) obtained by the recursive minimization of

$$\mu \mapsto \frac{1}{2\tau} W_2^2(\mu, \mu^{k-1}) + \phi(\mu), \quad k = 1, 2, \cdots, \tag{1.1}$$

with the initial condition $\mu^0 = \bar{\mu}$ (in this introduction we disregard, to keep the exposition as much simple as possible, the existence issue), and the related piecewise constant interpolant $\overline{M}_{\tau,t} := \mu^{[t/\tau]}$, $t > 0$, ($[t/\tau]$ denotes the integer part of t/τ) of the values μ^k on a uniform grid $\{0, \tau, 2\tau, \cdots, k\tau, \cdots\}$ of step size τ. The convergence of $\overline{M}_\tau$ to a continuous solution as $\tau \downarrow 0$ has attracted a lot of attention in recent years, see the references, starting from [28], mentioned in more detail after the statement of Theorem 4.7.

Under quite general assumptions on ϕ, it is possible to prove the existence of the limit $\mu_t = \lim_{\tau\downarrow 0} \overline{M}_{\tau,t}$ in $\mathscr{P}_2(\mathscr{V})$ for every time $t \geq 0$: the evolution of the limit curve is governed by a velocity vector field $\boldsymbol{v}_t \in L^2(\mu_t; \mathbb{R}^d)$ linked to μ_t itself through a nonlinear relation depending on the particular form of the functional ϕ, which we (formally, at this level) denote by with $-\boldsymbol{v}_t \in \partial\phi(\mu_t)$. It turns out that $\boldsymbol{v}_t$ is *tangent*, according to Otto's calculus, i.e. it is orthogonal in $L^2(\mu_t; \mathbb{R}^d)$ to vector fields $\boldsymbol{w}$ such that $\boldsymbol{w}\mu_t$ is divergence-free (or, equivalently, $\boldsymbol{v}_t$ belongs to the closure in $L^2(\mu_t; \mathbb{R}^d)$ of gradients of smooth maps) and the resulting evolution system reads as

$$\frac{\partial}{\partial t}\mu_t + D_x \cdot (\boldsymbol{v}_t \mu_t) = 0 \quad \text{in } \mathscr{D}'\big((0,T) \times \mathscr{V}\big), \quad -\boldsymbol{v}_t \in \partial\phi(\mu_t). \tag{1.2}$$

Here we consider the convergence of the following discrete quantities: assuming that all the measures μ^k are absolutely continuous with respect to the Lebesgue one $\mathscr{L}^d$, we denote by $\boldsymbol{t}^k$ the optimal transport map between μ^k and μ^{k+1} and we introduce the *iterated transport map*

$$\boldsymbol{T}^k := \boldsymbol{t}^{k-1} \circ \boldsymbol{t}^{k-2} \circ \cdots \circ \boldsymbol{t}^1 \circ \boldsymbol{t}^0,$$

mapping $\bar{\mu} = \mu^0$ to μ^k. We want to study the convergence of the maps $\boldsymbol{T}_{\tau,t} := \boldsymbol{T}^{[t/\tau]}$ as $\tau \downarrow 0$. A simple formal argument shows that their limit should be $\boldsymbol{X}_t$, where $\boldsymbol{X}_t$ is the flow associated to the vectorfield $\boldsymbol{v}_t$ arising in (1.2), i.e.

$$\begin{cases} \dfrac{d}{dt}\boldsymbol{X}_t(x) = \boldsymbol{v}_t\left(\boldsymbol{X}_t(x)\right), \\ \boldsymbol{X}_0(x) = x. \end{cases} \tag{1.3}$$

In order to make this intuition precise, and to show the convergence result, we use several auxiliary results, all of them with an independent interest: the first one, proved in §3, is a general stability result for flows associated to vectorfields in the same spirit of the results proved in [20], [3] and based in particular on the YOUNG measure technique in the space of (absolutely) continuous maps adopted in [3] (see also the Lecture Notes [2]). The main new feature here, compared to the previous results, is that we use the information that the limit vectorfield $\boldsymbol{v}$ is a *tangent* vectorfield, while *no* regularity is required either on the approximating vectorfields or on the approximating flows.

In §4 we recall some results from [6] (see in particular Theorem 4.7) relative to the convergence of the discrete EULER solutions $\overline{M}_\tau$ to the continuous one μ, which show that the stability theorem can be applied in many cases, depending on the regularity of the initial data. In particular we show in Proposition 4.8 that the convergence scheme works for the gradient flow of the *internal energy* functional

$$\phi(\mu) := \int_{\mathscr{V}} \psi(\beta)\, dy, \quad \text{with} \quad \mu = \beta \mathscr{L}^d \llcorner \mathscr{V}, \tag{1.4}$$

even without MCCANN's displacement convexity assumption, under a suitable boundedness assumptions on β. In this case it is well known that the gradient flow of ϕ corresponds to a non-linear diffusion equation with homogeneous Neumann boundary conditions, see (4.70). In order to achieve more general results we study also the *backward* problem, namely the convergence of $(\boldsymbol{T}_\tau)^{-1}$, which seems to be less dependent from the regularity of the initial data.

The convergence of the discrete transport flow $\boldsymbol{T}_\tau$ may seem an academic question, but this is not the case. It appears in a recent and very enlightening paper by EVANS, GANGBO and SAVIN [22], where the authors study the gradient flow with respect to the L^2 metric of the polyconvex functional

$$I(\boldsymbol{u}) := \int_{\mathscr{U}} \Phi(\det D\boldsymbol{u})\, dx,$$

and build a solution by purely differential methods. In that paper the authors raise the problem of the convergence of the (variational formulation of the) EULER's implicit scheme, resulting in the recursive minimization (analogous to (1.1)) of the functional

$$\boldsymbol{u} \mapsto \frac{1}{2\tau}\|\boldsymbol{u} - \boldsymbol{u}^{k-1}\|^2_{L^2(\mathscr{U};\mathbb{R}^d)} + I(\boldsymbol{u}), \quad k = 1, 2, \cdots \tag{1.5}$$

among the smooth diffeomorphisms mapping a reference domain $\mathscr{U} \subset \mathbb{R}^d$ onto a given target $\mathscr{V} \subset \mathbb{R}^d$. It turns out that (1.5) and (1.1) are strictly related, due to the fact that a change of variables gives

$$I(\boldsymbol{u}) = \int_{\mathscr{V}} \frac{\Phi(\det D\boldsymbol{u})}{\det D\boldsymbol{u}} \circ \boldsymbol{u}^{-1}\, dy = \int_{\mathscr{V}} \psi\left(\frac{1}{\det D\boldsymbol{u}} \circ \boldsymbol{u}^{-1}\right) dy$$

with $\psi(s) = s\Phi(1/s)$. Notice that the scalar quantity $[1/\det D\boldsymbol{u}] \circ \boldsymbol{u}^{-1}$ can be interpreted as the density $\beta[\boldsymbol{u}]$ of the push forward of $\mathscr{L}^d \llcorner \mathscr{U}$ under the map $\boldsymbol{u}$, so that $I(\boldsymbol{u})$ reduces to the "scalar" functional ϕ in (1.4).

Denoting by $\boldsymbol{u}_{\tau,t} := \boldsymbol{u}^{[t/\tau]}$ the interpolant of the discrete solution $(\boldsymbol{u}^k)$ to (1.5), GANGBO, EVANS and SAVIN show that $\boldsymbol{u}_\tau$ is indeed linked to the initial datum $\bar{\boldsymbol{u}}$ by an iterated composition of optimal transport maps, whose convergence as $\tau \downarrow 0$ can be studied by the methods developed in the present paper, see in particular Theorem 5.7.

We also extend the notion of solution of the gradient flow of I to the case when $\det D\boldsymbol{u}$ is possibly unbounded from above, thus allowing for degeneracies in $\beta = \beta[\boldsymbol{u}]$. At the same time, we allow for general domains $\mathscr{V}$, not necessarily bounded or coinciding with the support of the initial datum $\bar\beta = \beta[\bar{\boldsymbol{u}}]$, so that when this support is strictly contained in $\mathscr{V}$, the domain $\mathscr{V}$ plays the rôle of an ostacle. Under mild regularity assumptions on $\bar\beta$ we show in Theorem 5.4 that still a unique solution can be built by purely differential methods; nevertheless, we show in Theorem 5.7 that this solution is still the limit of the discrete EULER ones.

Concluding this introduction, we notice that this is a nice model problem where the strengths and the weaknesses of the differential and of the variational methods can be compared. In this perspective, it is worth mentioning that in DE GIORGI's variational approach to gradient flows (even in metric spaces), a fundamental rôle is played by the so called *descending slopes*, namely

$$|\partial I|(\boldsymbol{u}) := \limsup_{\|\boldsymbol{v}-\boldsymbol{u}\|_2 \to 0} \frac{[I(\boldsymbol{u}) - I(\boldsymbol{v})]^+}{\|\boldsymbol{u}-\boldsymbol{v}\|}, \quad |\partial\phi|(\beta) := \limsup_{W_2(\rho,\beta)\to 0} \frac{[\phi(\beta)-\phi(\rho)]^+}{W_2(\beta,\rho)}.$$

and by the "upper gradient" properties of their lower semicontinuous envelopes (see [6, Chap. 2], Def. 4.5, and Thm. 4.7). As we will discuss in Remark 5.6, one can obtain that

$$|\partial I|(\boldsymbol{u}) = |\partial\phi|(\beta) \quad \text{if } \beta = \beta[\boldsymbol{u}], \quad I(\boldsymbol{u}) = \phi(\beta) < +\infty, \tag{1.6}$$

and, at least when $\boldsymbol{u}$ is regular, $\mathscr{V}$ is convex, and the map $s \mapsto \Phi(s^d)$ is convex and nonincreasing, one can check that

$$|\partial I|(\boldsymbol{u}) = \big\|\mathrm{div}\,\big(\Phi'(\det D\boldsymbol{u})\,\mathrm{cof}\,D\boldsymbol{u}\big)\big\|_2. \tag{1.7}$$

Introducing the function F defined on $d\times d$-matrices

$$F(\mathsf{A}) := \Phi(\det\mathsf{A}) \qquad \text{so that} \qquad I(\boldsymbol{u}) = \int_{\mathscr{U}} F(D\boldsymbol{u})\,dx, \tag{1.8}$$

(1.7) takes the more familiar form

$$|\partial I|(\boldsymbol{u}) = \|\,\mathrm{div}\,DF(D\boldsymbol{u})\|_2, \tag{1.9}$$

which corresponds to the L^2-norm of the function which represents the first variation of I. It is considerably easier to obtain (1.9) and its lower semicontinuity when F *is convex*, which in particular forces DF to be monotone. In the present case (1.8), the lower semicontinuous envelope of $|\partial I|$ can be expressed through the "ad hoc" representation formula (1.6).

For general polyconvex (but not convex) functionals, even in a smooth and coercive setting, it would be interesting to find necessary/sufficient conditions for the validity of equality in (1.9), or to compute explicitly the slope functional.

2 Notation and preliminary results

We start by recalling some basic facts in Measure Theory. Let X, Y be *Polish* spaces, i.e. topological spaces whose topology is induced by a complete and separable metric (open subsets of complete and separable metric spaces, with the topology induced by the metric, are still Polish). We endow a Polish space X with the corresponding Borel σ-algebra and denote by $\mathscr{P}(X)$ (resp. $\mathscr{M}(X)$, $\mathscr{M}_+(X)$) the family of Borel probability (resp. real, nonnegative real) measures in X. $\vec{\mu} \leftrightarrow (\mu^\alpha)_{\alpha=1}^d$ will denote a $\mathbb{R}^d$-valued *vector measure* in $\big[\mathscr{M}(X)\big]^d$, identified with a d-tuple of real measures $\mu^\alpha \in \mathscr{M}(X)$.
We will denote by $\boldsymbol{i} : X \to X$ the identity map.

Definition 2.1 (Push-forward) *Let $\vec{\mu}$ be a $\mathbb{R}^d$-valued measure in X with finite total variation and let $f : X \to Y$ be a Borel map. The push-forward $f_\#\vec{\mu}$ is the $\mathbb{R}^d$-valued measure in Y defined by $f_\#\vec{\mu}(B) = \vec{\mu}(f^{-1}(B))$ for any Borel set $B \subset Y$.*

It is easy to check that $f_\#\vec{\mu}$ has finite total variation as well and that $|f_\#\vec{\mu}| \le f_\#|\vec{\mu}|$. An elementary approximation by simple functions shows the chain rule

$$\int_Y g\, df_\#\vec{\mu} = \int_X g \circ f\, d\vec{\mu} \tag{2.1}$$

for any bounded Borel function (or even either nonnegative or nonpositive, and $\overline{\mathbb{R}}$-valued, in the case $d = 1$ and $\vec{\mu} = \mu \in \mathscr{P}(X)$) $g : Y \to \mathbb{R}$.

Definition 2.2 (Narrow convergence and compactness) *Narrow (sequential) convergence in $\mathscr{P}(X)$ is the convergence induced by the duality with $C_b(X)$, the space of continuous and bounded functions in X. By Prokhorov theorem, a family $\mathscr{F}$ in $\mathscr{P}(X)$ is sequentially relatively compact with respect to the narrow convergence if and only if it is tight, i.e. for any $\varepsilon > 0$ there exists a compact set $K \subset X$ such that $\mu(X \setminus K) < \varepsilon$ for any $\mu \in \mathscr{F}$.*

In this paper we use only the "easy" implication in Prokhorov theorem, namely that any tight family is sequentially relatively compact. It is immediate to check that a sufficient condition for tightness of a family $\mathscr{F}$ of probability measures is the existence of a *coercive* functional $\Psi : X \to [0, +\infty]$ (i.e. a functional such that its sublevel sets $\{\Psi \le t\}$, $t \in \mathbb{R}^+$, are relatively compact in X) such that

$$\int_X \Psi(x) d\mu(x) \le 1 \qquad \forall \mu \in \mathscr{F}.$$

If $\mu \in \mathscr{P}(X)$, recall that a Y-valued sequence (v_h) of Borel maps between X and Y is said to converge in μ-measure to v if

$$\lim_{h\to\infty} \mu\left(\{d_Y(v_h, v) > \delta\}\right) = 0 \qquad \forall \delta > 0.$$

This is equivalent to the $L^1(\mu)$ convergence to 0 of the maps $1 \wedge d_Y(v_h, v)$. It is also well known that if $Y = \mathbb{R}$ and $|v_h|^p$ is equi-integrable, then $v_h \to v$ in μ-measure if and only if $v_h \to v$ in $L^p(\mu)$.

Lemma 2.3 (Convergence in measure and narrow convergence) *Let v_h, $v : X \to Y$ be Borel maps and let $\mu \in \mathscr{P}(X)$. Then $v_h \to v$ in μ-measure iff*

$$(\boldsymbol{i}, v_h)_\# \mu \text{ converges to } (\boldsymbol{i}, v)_\# \mu \text{ narrowly in } \mathscr{P}(X \times Y).$$

Proof. If $v_h \to v$ in μ-measure then $\varphi(x, v_h(x))$ converges in $L^1(\mu)$ to $\varphi(x, v(x))$, and therefore thanks to (2.1) we immediately obtain the convergence of the push-forward. Conversely, let $\delta > 0$ and, for any $\varepsilon > 0$, let $w \in C_b(X; Y)$ such that $\mu(\{v \neq w\}) \leq \varepsilon$. We define

$$\varphi(x, y) := 1 \wedge \frac{d_Y(y, w(x))}{\delta} \in C_b(X \times Y)$$

and notice that

$$\int_{X \times Y} \varphi \, d(\boldsymbol{i}, v_h)_\# \mu \geq \mu(\{d_Y(w, v_h) > \delta\}), \quad \int_{X \times Y} \varphi \, d(\boldsymbol{i}, v)_\# \mu \leq \mu(\{w \neq v\}).$$

Taking into account the narrow convergence of the push-forward we obtain that

$$\limsup_{h \to \infty} \mu(\{d_Y(v, v_h) > \delta\}) \leq \limsup_{h \to \infty} \mu(\{d_Y(w, v_h) > \delta\}) + \mu(\{w \neq v\}) \leq 2\mu(\{w \neq v\}) \leq 2\varepsilon$$

and since ε is arbitrary the proof is achieved. □

We recall also the criterion for strong convergence in $L^1(\mu)$ (see for instance Exercise 1.19 of [5])

$$\liminf_{h \to \infty} G_h \geq G \geq 0, \quad \limsup_{h \to \infty} \int_X G_h \, d\mu \leq \int_X G \, d\mu < +\infty \quad \Longrightarrow \quad \lim_{h \to \infty} \int_X |G_h - G| \, d\mu = 0. \tag{2.2}$$

Lemma 2.4 *Let $f : X \to Y$ be a Borel map, $\mu \in \mathscr{P}(X)$, and let $\boldsymbol{v} \in L^p(\mu; \mathbb{R}^d)$ for some $p \in (1, +\infty)$. Then, setting $\nu = f_\# \mu$, we have $f_\#(\boldsymbol{v}\mu) = \boldsymbol{w}\nu$ for some $\boldsymbol{w} \in L^p(\nu; \mathbb{R}^d)$ with*

$$\|\boldsymbol{w}\|_{L^p(\nu; \mathbb{R}^d)} \leq \|\boldsymbol{v}\|_{L^p(\mu; \mathbb{R}^d)}. \tag{2.3}$$

In case of equality we have $\boldsymbol{v} = \boldsymbol{w} \circ f$ μ-a.e. in X.

Proof. Let q be the dual exponent of p, $\boldsymbol{\nu} := f_\#(\boldsymbol{v}\mu)$, and $\varphi \in L^\infty(Y; \mathbb{R}^d)$; denoting by ν^α, $\alpha = 1, \cdots, d$, the components of $\boldsymbol{\nu}$ we have

$$\left| \sum_{\alpha=1}^d \int_Y \varphi^\alpha \, d\nu^\alpha \right| = \left| \sum_{\alpha=1}^d \int_X (\varphi^\alpha \circ f) \, \boldsymbol{v}^\alpha \, d\mu \right| \leq \|\varphi \circ f\|_{L^q(\mu; \mathbb{R}^d)} \|\boldsymbol{v}\|_{L^p(\mu; \mathbb{R}^d)} = \|\varphi\|_{L^q(\nu; \mathbb{R}^d)} \|\boldsymbol{v}\|_{L^p(\mu; \mathbb{R}^d)}.$$

Since φ is arbitrary this proves (2.3) and, as a consequence, the same identities above hold when $\varphi \in L^q(\nu; \mathbb{R}^d)$. In case of equality it suffices to choose $\varphi = |\boldsymbol{w}|^{p-2}\boldsymbol{w}$ to obtain that $\boldsymbol{v}$ coincides with $|\varphi \circ f|^{q-2}(\varphi \circ f) = \boldsymbol{w} \circ f$ μ-a.e. in X. □

We conclude this section by recalling a few basic facts from the theory of optimal transportation (see for instance [24], [41], [21], [6] for much more on this fascinating subject). Let $\mathscr{V}$ be an open set in $\mathbb{R}^d$ and for $p \in (1, +\infty)$ let us denote by $\mathscr{P}_p(\mathscr{V})$ the collection of all probability measures of $\mathscr{P}(\mathscr{V})$ with finite p-moment, i.e.

$$\mu \in \mathscr{P}_p(\mathscr{V}) \quad \Longleftrightarrow \quad \mu \in \mathscr{P}(\mathscr{V}), \quad \mathsf{m}_p(\mu) := \int_{\mathscr{V}} |x|^p \, d\mu(x) < +\infty. \tag{2.4}$$

We denote by W_p the p-Kantorovich-Rubinstein-Wasserstein distance in $\mathscr{P}(\mathscr{V})$, defined by

$$W_p^p(\mu, \nu) := \min \left\{ \int_{\mathscr{V} \times \mathscr{V}} |x - y|^p \, d\gamma : \ \gamma \in \mathscr{P}(\mathscr{V}), \ (\pi_1)_\# \gamma = \mu, \ (\pi_2)_\# \gamma = \nu \right\} \tag{2.5}$$

(here π_i, $i = 1, 2$, denote the canonical projections on the factors). It is not hard to show (see for instance [41], [6, Prop. 7.1.5]) that the convergence induced by the distance W_p is equivalent to the narrow convergence and the convergence of the p-moment, i.e.

$$\lim_{n \to \infty} W_p(\mu_n, \mu) = 0 \iff \begin{cases} \mu_n \text{ narrowly converge to } \mu, \\ \lim_{n \to \infty} \mathsf{m}_p(\mu_n) = \mathsf{m}_p(\mu). \end{cases} \tag{2.6}$$

Notice that *when $\mathscr{V}$ is bounded* $\mathscr{P}_p(\mathscr{V}) = \mathscr{P}(\mathscr{V})$ and the convergence induced by the distance W_p is precisely the narrow convergence.

In the case when $\mu \in \mathscr{P}_p^r(\mathscr{V})$, the subset of $\mathscr{P}_p(\mathscr{V})$ made of absolutely continuous measures with respect to Lebesgue measure $\mathscr{L}^d$, it can be shown [10, 24] that the minimum problem (2.5) has a unique solution γ, and γ is induced by a transport map $\boldsymbol{t}$:

$$\gamma = (\boldsymbol{i}, \boldsymbol{t})_\# \mu.$$

In particular $\boldsymbol{t}$ is the unique solution of Monge's optimal transport problem

$$\min \left\{ \int_{\mathscr{V}} |\boldsymbol{r} - \boldsymbol{i}|^p \, d\mu : \ \boldsymbol{r}_\# \mu = \nu \right\},$$

of which (2.5) is a relaxed version. Finally, we recall that if also $\nu \in \mathscr{P}^r(\mathscr{V})$, then

$$\boldsymbol{s} \circ \boldsymbol{t} = \boldsymbol{i} \quad \mu\text{-a.e.} \quad \text{and} \quad \boldsymbol{t} \circ \boldsymbol{s} = \boldsymbol{i} \quad \nu\text{-a.e.},$$

where $\boldsymbol{s}$ is the optimal transport map between ν and μ. Moreover, it has been proved in [24] that the optimal transport map $\boldsymbol{t}$ is differentiable at μ-a.e. point in $\mathscr{V}$ and $\det D\boldsymbol{t}(x) > 0$ for μ-a.e. $x \in \mathscr{V}$ As a consequence the change of variables formula [6, Lemma 5.5.3] can be applied to give

$$\mu = \rho \mathscr{L}^d \llcorner \mathscr{V}, \quad \nu = \sigma \mathscr{L}^d \llcorner \mathscr{V} \qquad \Longrightarrow \qquad \sigma(y) = \frac{\rho(\boldsymbol{s}(y))}{\det \tilde{D}\boldsymbol{t}(\boldsymbol{s}(y))} \quad \text{for } \nu\text{-a.e. } y \in \mathscr{V}. \tag{2.7}$$

In the case $p = 2$ the properties above can be rephrased in terms of the classical differentiability of convex analysis, as the gradients of a convex function is differentiable $\mathscr{L}^d$-a.e. in the interior of the domain of the function.

3 The main stability result

Time dependent measures, transport equation, and velocity vector fields

Throughout this section we fix a positive time $T > 0$, a summability exponent $p \in (1, +\infty)$ and an open set $\mathscr{V} \subset \mathbb{R}^d$.

Definition 3.1 *We denote by $\mathscr{F}_p(\mathscr{V})$ the family of time-dependent measures $\mu_t \in \mathscr{P}_p(\mathscr{V})$, with $t \in [0,T]$, such that $t \mapsto \mu_t$ is narrowly continuous and there exists a Borel velocity field $\boldsymbol{v}(t,x) = \boldsymbol{v}_t(x)$ such that*

$$\int_0^T \int_{\mathscr{V}} |\boldsymbol{v}_t(x)|^p \, d\mu_t(x) \, dt < +\infty, \tag{3.1}$$

and

$$\frac{\partial}{\partial t}\mu_t + D_x \cdot (\boldsymbol{v}_t \mu_t) = 0 \quad \textit{in } \mathcal{D}'((0,T) \times \mathscr{V}). \tag{3.2}$$

Notice that the narrow continuity assumption is made just for simplicity in the definition: indeed, it can be easily shown (see for instance Lemma 8.1.2 of [6]) that the existence of a velocity field $\boldsymbol{v}$ satisfying (3.1,3.2) implies the narrow continuity, possibly redefining μ_t for a $\mathscr{L}^1$-negligible set of times. Notice also that, since we assuming that all μ_t's are probability measures concentrated on $\mathscr{V}$, there is no transfer of mass across $\partial\mathscr{V}$ and (3.2) still holds in $\mathcal{D}'((0,T) \times \mathbb{R}^d)$. Finally, notice that a simple approximation argument shows that (3.2) and (3.1) ensure the validity of the implication

$$\mu_t \in \mathscr{P}(\mathscr{V}), \quad \mu_0 \in \mathscr{P}_p(\mathscr{V}) \qquad \Longrightarrow \qquad \mu_t \in \mathscr{P}_p(\mathscr{V}) \ \ \forall t \in [0,T].$$

So, the assumption that $\mu_t \in \mathscr{P}_p(\mathscr{V})$ could be replaced by $\mu_0 \in \mathscr{P}_p(\mathscr{V})$.

Remark 3.2 (Space-time representation) We can obviously identify the families $\mu_t, \boldsymbol{v}_t\mu_t$ with the associated nonnegative real measure $\mu \in \mathscr{M}_+((0,T) \times \mathscr{V})$ and the vector measure $\boldsymbol{v}\mu \in \big[\mathscr{M}((0,T) \times \mathscr{V})\big]^d$ defined as

$$\mu := \int_0^T \mu_t \, d\mathscr{L}^1(t), \quad \boldsymbol{v}\mu := \int_0^T \boldsymbol{v}_t\mu_t \, d\mathscr{L}^1(t), \tag{3.3}$$

so that for every bounded Borel real function φ defined in $(0,T) \times \mathscr{V}$

$$\begin{aligned} \iint_{(0,T)\times\mathscr{V}} \varphi(t,x) \, d\mu(t,x) &:= \int_0^T \int_{\mathscr{V}} \varphi(t,x) \, d\mu_t(x) \, dt, \\ \iint_{(0,T)\times\mathscr{V}} \varphi(t,x) \, d(\boldsymbol{v}\mu)(t,x) &:= \int_0^T \int_{\mathscr{V}} \varphi(t,x)\boldsymbol{v}(t,x) \, d\mu_t(x) \, dt. \end{aligned} \tag{3.4}$$

Thus (3.1) simply means $\boldsymbol{v} \in L^p(\mu; \mathbb{R}^d)$.

Given $\mu_t \in \mathscr{F}_p(\mathscr{V})$, there are in principle many vector fields $\boldsymbol{v}_t$ satisfying (3.1) and (3.2), and all of them will be called *admissible*. It will be useful in the sequel to define a convergence in $\mathscr{F}_p(\mathscr{V})$ which takes into account also the behaviour of the admissible velocity fields.

Definition 3.3 (Convergence in $\mathscr{F}_p(\mathscr{V})$) *We say that $\mu_t^n \in \mathscr{F}_p(\mathscr{V})$ converge to $\mu_t \in \mathscr{F}_p(\mathscr{V})$ if $\mu_t^n \to \mu_t$ narrowly for any $t \in [0,T]$. If $\boldsymbol{v}_t^n$, $\boldsymbol{v}_t$ are admissible velocity fields corresponding to μ_t^n, μ_t respectively, with*

$$\sup_n \int_0^T \int_{\mathscr{V}} |\boldsymbol{v}_t^n|^p \, d\mu_t^n \, dt < +\infty,$$

then we say that $\boldsymbol{v}_t^n$ converge to $\boldsymbol{v}_t$ if (recalling the notation (3.3))

$$\lim_{n\to\infty} \boldsymbol{v}^n \mu^n = \boldsymbol{v}\mu \qquad \text{in } \mathcal{D}'((0,T)\times\mathscr{V}). \tag{3.5}$$

The following result, proved in Theorem 8.3.1 and Proposition 8.4.5 of [6], provides a characterization of the "optimal" velocity field among all the admissible ones. See also Chapter 8 of [6] for a more detailed explaination of why this result can be used to make Otto's calculus (see [35], [36]) in $\mathscr{P}_p(\mathscr{V})$ rigorous, and to characterize the infinitesimal behaviour of the Wasserstein distance.

Theorem 3.4 (Minimal velocity field) *For any $\mu_t \in \mathscr{F}_p(\mathscr{V})$ there exists a unique, up to $\mathscr{L}^1$-negligible sets, admissible vector field $\boldsymbol{v}_t$ with the property*

$$\boldsymbol{v}_t \in \mathrm{Tan}_{\mu_t}\mathscr{P}_p(\mathscr{V}) := \overline{\{j_q(\nabla\varphi) : \ \varphi \in C_c^\infty(\mathbb{R}^d)\}}^{L^p(\mu_t;\mathbb{R}^d)} \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in [0,T], \tag{3.6}$$

where $q \in (1,\infty)$ is the dual exponent of p and $j_q(x) = |x|^{q-2}x$. This vector field satisfies also

$$\int_{\mathscr{V}} |\boldsymbol{v}_t|^p \, d\mu_t \le \int_{\mathscr{V}} |\tilde{\boldsymbol{v}}_t|^p \, d\mu_t \qquad \text{for } \mathscr{L}^1\text{-a.e. } t \in [0,T] \tag{3.7}$$

for any other admissible velocity field $\tilde{\boldsymbol{v}}_t$.

We call the velocity field given by Theorem 3.4 *tangent* velocity field. So, checking that a vector field is tangent amounts to check that the continuity equation (3.2) holds and that $j_p(\boldsymbol{v}_t)$ is approximable, in $L^q(\mu_t;\mathbb{R}^d)$, by gradients. Finally, we recall that Proposition 8.3.1 of [6] also shows that $\mathscr{F}_p(\mathscr{V})$ coincides with the class of absolutely continuous curves with values in $\mathscr{P}_p(\mathscr{V})$, when the target is endowed with the p-th Wasserstein metric, and that the L^p norm of the tangent velocity field can be characterized by the rate of change of the p-th Wasserstein metric along the curve:

$$\|\boldsymbol{v}_t\|_{L^p(\mu_t;\mathbb{R}^d)} = \lim_{h\to 0} \frac{W_p(\mu_{t+h},\mu_t)}{|h|} \qquad \text{for } \mathscr{L}^1\text{-a.e. } t \in [0,T].$$

Flows and their stability

Given a reference time $s \in [0,T]$, a measure $\bar{\mu}_s \in \mathscr{P}_p(\mathscr{V})$, and a Borel field $\boldsymbol{v} : (t,x) \in (0,T)\times\mathscr{V} \to \boldsymbol{v}_t(x) \in \mathbb{R}^d$, a canonical way to build a solution of (3.2) is to find a *flow* associated to $\boldsymbol{v}_t$, i.e. a map $\boldsymbol{X}(t,s,x) : [0,T] \times [0,T] \times \mathscr{V} \to \mathscr{V}$ such that $t \mapsto \boldsymbol{X}(t,s,x)$ is absolutely continuous in $[0,T]$ and it is an integral solution of the ODE

$$\begin{cases} \dfrac{d}{dt}\boldsymbol{X}(t,s,x) = \boldsymbol{v}_t\left(\boldsymbol{X}(t,s,x)\right) \\ \boldsymbol{X}(s,s,x) = x \end{cases} \tag{3.8}$$

for $\bar{\mu}_s$-a.e. x, with

$$\int_{\mathscr{V}} \int_0^T \Big|\frac{d}{dt}\boldsymbol{X}(t,s,x)\Big|^p \,dt\,d\bar{\mu}_s(x) < +\infty. \tag{3.9}$$

When $s = 0$ we will speak of *forward flows* and we will often omit to indicate the explicit occurrence of s in $\boldsymbol{X}(t,0,x)$, writing either $\boldsymbol{X}(t,x)$ or $\boldsymbol{X}_t(x)$; analogously, the case $s = T$ corresponds to *backward flows*. Using test functions of the form $\chi(t)\varphi(x)$, it is then immediate to check that the narrowly continuous family of measures

$$\mu_t = \boldsymbol{X}(t,s,\cdot)_\#\bar{\mu}_s \tag{3.10}$$

solves (3.1,3.2) with the condition $\mu_s = \bar{\mu}_s$ and belongs to $\mathscr{F}_p(\mathscr{V})$ if $\bar{\mu}_s \in \mathscr{P}_p(\mathscr{V})$; in this case we say that $\boldsymbol{X}$ is a flow associated with $(\mu_t, \boldsymbol{v}_t)$. A similar situation occurs when one already knows a solution μ_t of (3.2) and looks for a representation formula like (3.10) [6, Chap. 8].

Without assuming any regularity on $\boldsymbol{v}_t$ we do not know anything about the existence and the uniqueness of a flow associated to $(\mu_t, \boldsymbol{v}_t)$. Postponing to the next section a more detailed discussion of this aspect, we recall that the simplest condition [6, Prop. 8.1.8 and Thm. 8.2.1] which ensures both the existence and the uniqueness of the *forward* flow representing $(\mu_t, \boldsymbol{v}_t)$ is that the ODE (3.8) *admits a unique $\overline{\mathscr{V}}$-valued solution* for $\bar{\mu}_0$-a.e. initial datum $x \in \mathscr{V}$. This property is surely verified if $\boldsymbol{v}_t$ satisfies the classical (local) Lipschitz condition: denoting by $\mathrm{Lip}(\boldsymbol{w}, B)$ the Lipschitz constant of a map $\boldsymbol{w}$ on B, it means that

$$\forall\,(t_0,x_0) \in \Big((0,T)\times\overline{\mathscr{V}}\Big) \cup \Big(\{0\}\times\mathscr{V}\Big) \quad \exists \varepsilon,\tau > 0: \quad \int_{t_0}^{t_0+\tau} \mathrm{Lip}(\boldsymbol{v}_t, B_\varepsilon(x_0)\cap\overline{\mathscr{V}})\,dt < +\infty. \tag{3.11}$$

We are mainly concerning with the *stability properties* of flows: supposing that $\boldsymbol{X}^n$ and $\boldsymbol{X}$ are *forward* flows associated with $(\mu^n, \boldsymbol{v}^n)$ and $(\mu, \boldsymbol{v})$ respectively, we look for conditions on the measures and on their velocity vectorfields ensuring the convergence of $\boldsymbol{X}^n$ to $\boldsymbol{X}$.

The next main convergence result shows that if

(a) $\boldsymbol{v}$ *is the tangential velocity field of* μ and

(b) (3.8) admits at most one solution $\boldsymbol{X}$ for $\bar{\mu}_0$-a.e. $x \in \mathscr{V}$,

then the convergence of μ_t^n in $\mathscr{F}_p(\mathscr{V})$ and of the $L^p(\mu^n)$-norms of $\boldsymbol{v}^n$ are sufficient. Its proof uses, in the same spirit of [3], [2], narrow convergence in the space of continuous maps as a technical tool for proving convergence in measure of the flows.

Theorem 3.5 (Main convergence result) *Assume that we are given:*

(i) flows $\mu_t^n \in \mathscr{F}_p(\mathscr{V})$ converging to $\mu_t \in \mathscr{F}_p(\mathscr{V})$ as in Definition 3.3;

(ii) velocity fields $\boldsymbol{v}_t^n \in L^p(\mu_t^n; \mathbb{R}^d)$ admissible relative to μ_t^n;

(iii) forward flows $\boldsymbol{X}^n$ associated to $(\mu_t^n, \boldsymbol{v}_t^n)$, i.e. satisfying the ODE system (3.8) *for $\boldsymbol{v}_t^n$ with $s = 0$ and the transport condition* (3.10) *$\mu_t^n = \boldsymbol{X}^n(t, \cdot)_\# \mu_0^n$.*

Let $\boldsymbol{v}_t \in \operatorname{Tan}_{\mu_t} \mathscr{P}_p(\mathscr{V})$ be the tangent velocity field to μ_t (e.g. $\boldsymbol{v}_t$ satisfies (3.2) *and* (3.6)*) and assume that*

$$\mu_0^n = \boldsymbol{d}^n_\# \mu_0 \quad \textit{with} \quad \lim_{n \to \infty} \int_{\mathscr{V}} |\boldsymbol{d}^n - \boldsymbol{i}|^p \, d\mu_0 = 0, \tag{3.12}$$

$$\limsup_{n \to \infty} \int_0^T \int_{\mathscr{V}} |\boldsymbol{v}_t^n|^p \, d\mu_t^n \, dt \le \int_0^T \int_{\mathscr{V}} |\boldsymbol{v}_t|^p \, d\mu_t \, dt, \tag{3.13}$$

and that

$$\textit{the ODE (3.8) admits } \text{at most one } \overline{\mathscr{V}}\text{-valued solution } \textit{for } \mu_0\textit{-a.e. } x \in \mathscr{V}. \tag{3.14}$$

Then there exists a unique forward flow $\boldsymbol{X} : [0, T] \times \mathscr{V} \to \overline{\mathscr{V}}$ associated to $(\mu_t, \boldsymbol{v}_t)$ and the flows $\boldsymbol{X}^n$ converge in L^p to $\boldsymbol{X}$, precisely

$$\lim_{n \to \infty} \int_{\mathscr{V}} \sup_{[0,T]} \Big| \boldsymbol{X}^n(\cdot, \boldsymbol{d}^n(x)) - \boldsymbol{X}(\cdot, x) \Big|^p \, d\mu_0(x) = 0. \tag{3.15}$$

Proof. We denote by $\Gamma = C^0([0, T]; \overline{\mathscr{V}})$ the complete and separable metric space of continuous maps from $[0, T]$ to $\overline{\mathscr{V}}$, whose generic element will be denoted by γ, and we denote by $\boldsymbol{e}_t : \Gamma \to \overline{\mathscr{V}}$ the evaluation maps $\boldsymbol{e}_t(\gamma) = \gamma(t)$.

Since $\boldsymbol{X}^n(t, x) \in \mathscr{V}$ for every $x \in \mathscr{V}$, we can define probability measures $\boldsymbol{\eta}^n$ in Γ by

$$\boldsymbol{\eta}^n := (\boldsymbol{X}^n(\cdot, x))_\# \, \mu_0^n, \tag{3.16}$$

where $x \mapsto \boldsymbol{X}^n(\cdot, x)$ is the natural map from $\mathscr{V}$ to Γ. Since $\boldsymbol{e}_t \circ (\boldsymbol{X}^n(\cdot, x)) = \boldsymbol{X}^n(t, x)$ we immediately obtain that

$$(\boldsymbol{e}_t)_\# \boldsymbol{\eta}^n = \boldsymbol{X}^n(t, \cdot)_\# \mu_0^n = \tilde{\mu}_t^n, \tag{3.17}$$

where we denoted by $\tilde{\mu}_t^n$ the trivial extension of μ_t^n to $\overline{\mathscr{V}}$ obtained by setting $\tilde{\mu}_t^n(\partial \mathscr{V}) = 0$.

Step 1. (Tightness of $\boldsymbol{\eta}^n$) We claim that the family $\{\boldsymbol{\eta}^n\}$ is tight as $n \to \infty$, that any limit point $\boldsymbol{\eta}$ is concentrated on $\overline{\mathscr{V}}$*-valued and absolutely continuous* maps in $[0, T]$ and that

$$(\boldsymbol{e}_t)_\# \boldsymbol{\eta} = \tilde{\mu}_t, \tag{3.18}$$

$$\int_\Gamma \int_0^T |\dot{\gamma}(t)|^p \, dt \, d\boldsymbol{\eta}(\gamma) \le \int_0^T \int_{\mathscr{V}} |\boldsymbol{v}_t|^p \, d\mu_t \, dt, \tag{3.19}$$

$$\gamma(t) \in \mathscr{V} \ \ \mathscr{L}^1\text{-a.e. in } [0, T], \text{ for } \boldsymbol{\eta}\text{-a.e. } \gamma. \tag{3.20}$$

Indeed, by Ascoli-Arzelà theorem, the functional

$$\Psi(\gamma) := \begin{cases} |\gamma(0)|^p + \int_0^T |\dot{\gamma}(t)|^p \, dt & \text{if } \gamma \text{ is absolutely continuous} \\ +\infty & \text{otherwise} \end{cases}$$

is coercive in Γ, and since $\dot\gamma = \boldsymbol{v}^n(\gamma)$ for $\boldsymbol{\eta}^n$-a.e. $\gamma\in\Gamma$ and the p-moment of μ_0^n converges to the p-moment of μ_0 by (3.12), we get

$$\begin{aligned}\limsup_{n\to\infty}\int_\Gamma \Psi(\gamma)\,d\boldsymbol{\eta}^n &= \limsup_{n\to\infty}\int_\Gamma |\gamma(0)|^p\,d\boldsymbol{\eta}^n + \int_0^T\int_\Gamma |\boldsymbol{v}_t^n(\gamma(t))|^p\,d\boldsymbol{\eta}^n\,dt\\ &\le \limsup_{n\to\infty}\int_{\mathscr{V}} |x|^p\,d\mu_0^n(x) + \int_0^T\int_{\mathscr{V}} |\boldsymbol{v}_t^n|^p\,d\mu_t^n\,dt < +\infty,\end{aligned}$$

thus proving the tightness of the family. Moreover, the lower semicontinuity of Ψ gives that $\int\Psi\,d\boldsymbol{\eta}$ is finite, so that $\boldsymbol{\eta}$ is concentrated on the absolutely continuous maps. An analogous argument proves (3.19), while (3.18) can be achieved passing to the limit in (3.17).

Finally, by introducing the upper semicontinuous function

$$j_{\mathscr{V}}(x) := \begin{cases} 0 & \text{if } x\in\mathscr{V},\\ 1 & \text{if } x\in\partial\mathscr{V},\end{cases}$$

(3.18) and Fubini's theorem give

$$0 = \int_0^T\int_{\mathscr{V}} j_{\mathscr{V}}(x)\,d\tilde\mu_t(x)\,dt = \int_0^T\int_\Gamma j_{\mathscr{V}}(\gamma(t))\,d\boldsymbol{\eta}(\gamma)\,dt = \int_\Gamma\int_0^T j_{\mathscr{V}}(\gamma(t))\,dt\,d\boldsymbol{\eta}(\gamma) \tag{3.21}$$

which is equivalent to (3.20).

Step 2. (Representation of $\boldsymbol{v}_t$) Let $\boldsymbol{\eta}$ be a limit point as in Step 1, along some sequence $n_i\to\infty$. By the previous claim we know that for $\boldsymbol{\eta}$-a.e. $\gamma\in\Gamma$ the vector field $\boldsymbol{z}_t(\gamma) := \dot\gamma(t)$ is well defined up to a $\mathscr{L}^1$-negligible subset of $(0,T)$. By Fubini's theorem and (3.19), it coincides with a Borel vector field $\boldsymbol{z}\in L^p(\mathscr{L}^1\times\boldsymbol{\eta};\mathbb{R}^d)$ with

$$\int_0^T\int_\Gamma |\boldsymbol{z}_t(\gamma)|^p\,d\boldsymbol{\eta}(\gamma)dt = \int_\Gamma\int_0^T |\dot\gamma(t)|^p\,dt\,d\boldsymbol{\eta}(\gamma) \le \int_0^T\int_{\mathscr{V}} |\boldsymbol{v}_t|^p\,d\mu_t\,dt. \tag{3.22}$$

in particular there exists a Borel set $\mathcal{T}\subset(0,T)$ of full measure such that $\boldsymbol{z}_t\in L^p(\boldsymbol{\eta})$ for every $t\in\mathcal{T}$.

Define now $\boldsymbol{\nu}_t := (\boldsymbol{e}_t)_\#(\boldsymbol{z}_t\cdot\boldsymbol{\eta})$, and notice that Lemma 2.4 gives $\boldsymbol{\nu}_t$ is well defined and absolutely continuous with respect to $\mu_t = (\boldsymbol{e}_t)_\#\boldsymbol{\eta}$ for any $t\in\mathcal{T}$. Then, writing $\nu_t = \boldsymbol{w}_t\mu_t$, Lemma 2.4 again and (3.22) give

$$\int_0^T\int_{\mathscr{V}} |\boldsymbol{w}_t|^p\,d\mu_t\,dt \le \int_0^T\int_\Gamma |\boldsymbol{z}_t(\gamma)|^p\,d\boldsymbol{\eta}\,dt \le \int_0^T\int_{\mathscr{V}} |\boldsymbol{v}_t|^p\,d\mu_t\,dt. \tag{3.23}$$

Now, let us show that $\boldsymbol{w}_t$ is an admissible velocity field relative to μ_t: indeed, for any test function $\varphi\in C_c^\infty(\mathscr{V})$ we have

$$\begin{aligned}\frac{d}{dt}\int_{\mathscr{V}}\varphi\,d\mu_t &= \frac{d}{dt}\int_\Gamma \varphi(\gamma(t))\,d\boldsymbol{\eta} = \int_\Gamma \langle\nabla\varphi(\gamma(t)),\dot\gamma(t)\rangle\,d\boldsymbol{\eta}\\ &= \int_\Gamma\langle\nabla\varphi(\gamma(t)),\boldsymbol{z}_t(\gamma)\rangle\,d\boldsymbol{\eta} = \int_{\mathscr{V}}\langle\nabla\varphi,\boldsymbol{w}_t\rangle\,d\mu_t.\end{aligned} \tag{3.24}$$

As a consequence (3.23) and the minimality of $\boldsymbol{v}_t$ yield $\boldsymbol{v}_t = \boldsymbol{w}_t$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$. In addition, from the equality

$$\int_\Gamma |\dot\gamma(t)|^p \, d\boldsymbol{\eta} = \int_{\mathscr{V}} |\boldsymbol{w}_t|^p \, d(\boldsymbol{e}_t)_\# \boldsymbol{\eta}$$

and Lemma 2.4 we infer

$$\boldsymbol{v}_t(\gamma(t)) = \dot\gamma(t) \quad \boldsymbol{\eta}\text{-a.e. in } \Gamma \tag{3.25}$$

for $\mathscr{L}^1$-a.e. $t \in [0,T]$. Then, (3.25) and Fubini's theorem give

$$\boldsymbol{v}_t(\gamma(t)) = \dot\gamma(t) \quad \mathscr{L}^1\text{-a.e. in } (0,T) \tag{3.26}$$

for $\boldsymbol{\eta}$-a.e. γ. In other words, $\boldsymbol{\eta}$ is concentrated on the absolutely continuous solutions of the ODE relative to the vector field $\boldsymbol{v}_t$.

Step 3. (Conclusion) By assumption, we know that there is at most one solution $\gamma \in \Gamma$ of the ODE $\dot\gamma = \boldsymbol{v}_t(\gamma)$ with an initial condition $\gamma(0) = x \in \mathscr{V}$. Taking into account (3.20) we obtain that the disintegration $\{\boldsymbol{\eta}_x\}_{x\in\mathscr{V}}$ induced by the map $\boldsymbol{e}_0$ is a Dirac mass, concentrated on the unique $\overline{\mathscr{V}}$-valued solution $\boldsymbol{X}(\cdot,x)$ of the ODE starting from x at $t=0$. As a consequence

$$\boldsymbol{\eta} = (\boldsymbol{X}(\cdot,x))_\# \, \mu_0. \tag{3.27}$$

Since $\boldsymbol{\eta}$ does not depend on the subsequence (n_i), the tightness of $(\boldsymbol{\eta}^n)$ gives

$$(\boldsymbol{X}^n(\cdot,x))_\# \, \mu_0^n \ \text{ converge to } \ (\boldsymbol{X}(\cdot,x))_\# \, \mu_0 \ \text{ narrowly as } n \to \infty.$$

Since

$$(\boldsymbol{X}^n(\cdot,x))_\# \, \mu_0^n = (\boldsymbol{X}^n(\cdot,\boldsymbol{d}^n(x)))_\# \, \mu_0$$

Lemma 2.3 yields that the sequence of functions

$$g^n(x) := \sup_{t\in[0,T]} |\boldsymbol{X}^n(\cdot,\boldsymbol{d}^n(x)) - \boldsymbol{X}(\cdot,x)| \quad \text{converges to 0 in } \mu_0\text{-measure.} \tag{3.28}$$

In order to prove the L^p-convergence (3.15) we can assume with no loss of generality that $\boldsymbol{X}^n(\cdot,x) \to \boldsymbol{X}(\cdot,x)$ uniformly in $[0,T]$ for μ_0-a.e. x, and we need to show that $|g^n|^p$ is equi-integrable in $L^1(\mu_0)$. To this aim, taking into account the fact that strongly converging sequences are equi-integrable and (2.2), we exhibit a sequence of functions $G^n \geq |g^n|^p$ such that

$$\liminf_{n\to\infty} G^n(x) \geq G(x) \quad \text{for } \mu_0\text{-a.e. } x \in \mathscr{V} \qquad \text{and} \qquad \limsup_{n\to\infty} \int_{\mathscr{V}} G^n(x)\, d\mu_0(x) \leq \int_{\mathscr{V}} G(x)\, d\mu_0(x).$$

We can choose

$$G^n(x) := 3^{p-1}\Big(\sup_{[0,T]} |\boldsymbol{X}(\cdot,x)|^p + |\boldsymbol{d}^n(x)|^p + T^{p-1} h^n(x)\Big), \quad \text{where} \quad h^n(x) := \int_0^T \Big|\frac{d}{dt}\boldsymbol{X}^n(t,x)\Big|^p \, dt.$$

Since $\boldsymbol{d}^n \to \boldsymbol{i}$ in $L^p(\mu_0;\mathbb{R}^d)$ and standard lower semicontinuity result yield

$$\liminf_{n\to\infty} h^n(x) \geq h(x) := \int_0^T \Big|\frac{d}{dt}\boldsymbol{X}(t,x)\Big|^p \, dt \quad \text{for } \mu_0\text{-a.e. } x \in \mathscr{V} \tag{3.29}$$

the proof is achieved if

$$\limsup_{n\to\infty}\int_{\mathscr{V}} h^n(x)\,d\mu_0(x) \le \int_{\mathscr{V}} h(x)\,d\mu_0(x). \tag{3.30}$$

For, we can calculate

$$\begin{aligned}\int_{\mathscr{V}} h^n(x)\,d\mu_0(x) &= \int_{\mathscr{V}}\int_0^T \Big|\frac{d}{dt}\boldsymbol{X}^n(t,x)\Big|^p\,dt\,d\mu_0(x) = \int_{\mathscr{V}}\int_0^T \big|\boldsymbol{v}_t^n(\boldsymbol{X}^n(t,x))\big|^p\,dt\,d\mu_0(x)\\ &= \int_0^T\int_{\mathscr{V}} \big|\boldsymbol{v}_t^n(y)\big|^p\,d\mu_t(y)\,dt\end{aligned}$$

and by (3.13) we get

$$\begin{aligned}\limsup_{n\to\infty}\int_{\mathscr{V}} h^n(x)\,d\mu_0(x) &\le \int_0^T\int_{\mathscr{V}} \big|\boldsymbol{v}_t(y)\big|^p\,d\mu_t(y)\,dt = \int_{\mathscr{V}}\int_0^T \Big|\frac{d}{dt}\boldsymbol{X}(t,x)\Big|^p\,dt\,d\mu_0(x)\\ &= \int_{\mathscr{V}} h(x)\,d\mu_0(x).\end{aligned}$$

□

On the regularity of the limit vectorfield

We notice that in Theorem 3.5 no regularity is imposed on the approximating velocity fields $\boldsymbol{v}_t^n$ and that assumption (3.14) stating μ_0-a.e. uniqueness of solutions of the ODE (3.8) can be guaranteed by weaker assumptions than (3.11): first, we need only uniqueness of *forward* characteristics, and this requires only a *one-sided* Lipschitz condition (see Remark 3.6); second, we don't really need uniqueness of forward characteristics in a *pointwise* sense, but rather that any probability measure $\boldsymbol{\eta}$ in Γ concentrated on absolutely continuous solutions of the ODE $\dot\gamma = \boldsymbol{v}_t(\gamma)$ is representable as in (3.27) for some "natural" flow $\boldsymbol{X}$. Several situations where this happens are described in Remark 3.7. Remark 3.8 shows that, when $\mu = \beta\mathscr{L}^d \llcorner \mathscr{V}$ and β is continuous, it is sufficient to control the local regularity of $\boldsymbol{v}$ in the positivity set of β and the integrability of the positive part of its divergence.

Remark 3.6 (One-sided Lipschitz condition) All conclusions of Theorem 3.5 remain valid if $\boldsymbol{v}_t$ (or, more precisely, at least one of the functions in its equivalence class modulo μ_t-negligible sets) satisfies the one-sided Lipschitz condition

$$\langle \boldsymbol{v}_t(x) - \boldsymbol{v}_t(y), x-y\rangle \le \omega_t|x-y|^2 \qquad \forall x,\, y\in\mathscr{V} \tag{3.31}$$

with $\omega \in L^1_{\rm loc}([0,T))$. Indeed, it is well known that this condition ensures pointwise uniqueness for the forward Cauchy problem associated to $\boldsymbol{v}_t$, and therefore uniqueness of the measure $\boldsymbol{\eta}$ built during the proof of Theorem 3.5. See also [38], [8], [9] for other well-posedness results in this context.

Remark 3.7 (Sobolev or BV regularity) As the proof clearly shows, assumption (3.14) in Theorem 3.5 could be replaced by the following more technical (but also much more general) one:

$$
\begin{aligned}
&\textit{any probability measure } \boldsymbol{\eta} \textit{ in } \Gamma = C^0([0,T];\overline{\mathscr{V}}) \textit{ concentrated on absolutely continuous}\\
&\qquad \textit{solutions of the ODE} \quad \dot\gamma = \boldsymbol{v}_t(\gamma) \quad \textit{and satisfying} \quad (\boldsymbol{e}_t)_\# \boldsymbol{\eta} = \mu_t \\
&\qquad\qquad \textit{is representable as} \quad \boldsymbol{\eta} = (\boldsymbol{X}(\cdot,x))_\# \mu_0 \quad \textit{for some flow } \boldsymbol{X}.
\end{aligned}
\tag{3.32}
$$

Let us consider for instance the case when $\mathscr{V}$ is bounded and

$$
\boldsymbol{v} \in L^1_{\rm loc}\Big([0,T); W^{1,1}_{\rm loc}(\mathscr{V};\mathbb{R}^d)\Big), \quad [D_x \cdot \boldsymbol{v}]^- \in L^1_{\rm loc}\left([0,T); L^\infty(\mathscr{V})\right).
$$

Under these assumptions it has been proved in [20] (the assumption $\nabla_x \cdot \boldsymbol{v} \in L^1(L^\infty)$ made in that paper can be weakened, requiring only a bound on the negative part, arguing as in [3]) that the continuity equation (3.2) has at most one solution in the class of μ_t's of the form $\mu_t = \beta_t \mathscr{L}^d$ with $\beta \in L^\infty_{\rm loc}([0,T); L^\infty(\mathscr{V}))$, for any initial condition $\mu_0 = \beta_0 \mathscr{L}^d$, $\beta_0 \in L^\infty(\mathscr{V})$. As explained in [2], it is a general fact that the well-posedness of the continuity equation in the class of μ_t's above (precisely, the validity of a comparison principle) implies the validity of (3.32) for μ_t's in the same class; in this particular case $\boldsymbol{X}$ is the so-called DiPerna–Lions flow associated to $\boldsymbol{v}$ (see [20] and also [3] for a different characterization of it).
It was shown in [3] (the original $L^1(L^\infty)$ estimate on the negative part of the divergence has been improved to an $L^1(L^1)$ one in [2]) that also a BV regularity on $\boldsymbol{v}$ can be considered, together with the absolute continuity of the distributional divergence:

$$
\boldsymbol{v} \in L^1_{\rm loc}\Big([0,T); BV_{\rm loc}(\mathscr{V};\mathbb{R}^d)\Big), \quad [D_x \cdot \boldsymbol{v}]^- \in L^1_{\rm loc}\left([0,T); L^1(\mathscr{V})\right).
$$

In this case there is uniqueness of bounded solutions of (3.2) and again (3.32) holds in the class of solutions $\mu_t = \beta_t \mathscr{L}^d$ of (3.2) with β bounded. Other classes of vectorfields to which (3.32) applies are considered in [27], [26], [4], [30], [31].

Remark 3.8 (Continuous densities) Let P_0 be an open subset of $\mathscr{V}$ with $\mu_0(\mathscr{V} \setminus P_0) = 0$. When

$$
\mu_t = \beta_t \mathscr{L}^d \llcorner \mathscr{V} \quad \text{and the map } (t,x) \mapsto \beta_t(x) \text{ is continuous in } \Big((0,T)\times\overline{\mathscr{V}}\Big) \cup \Big(\{0\}\times P_0\Big) \tag{3.33}
$$

we can localize condition (3.11) to the (open) positivity set of β. Thus we introduce the sets

$$
P := \Big\{(t,x) \in (0,T)\times\overline{\mathscr{V}} : \beta_t(x) > 0\Big\} \tag{3.34}
$$

and we assume that

$$
\beta \in C^1(P),\ \boldsymbol{v},\ D_x \boldsymbol{v} \in C^0(P), \tag{3.35}
$$

$$
\forall\, x_0 \in P_0 \text{ we have } \beta_0(x_0) > 0 \text{ and } \exists\, \varepsilon > 0 \text{ such that } \int_0^\varepsilon \sup_{x \in B_\varepsilon(x_0)} \|D_x \boldsymbol{v}_t(x)\|\, dt < +\infty, \tag{3.36}
$$

so that solutions of the ODE (3.8) are locally unique in P and the Cauchy problem for $x \in P_0$ admits a unique maximal forward $\mathscr{V}$-valued solution. If

$$\int_\varepsilon^T \int_{\{x:\beta_t(x)>0\}} \big[D_x \cdot \boldsymbol{v}_t(x)\big]^+ \, d\mu_t(x)\, dt < +\infty \quad \forall\, \varepsilon > 0, \tag{3.37}$$

then the same conclusion of Theorem 3.5 hold. This fact is a direct consequence of the following "confinement" Lemma.

Lemma 3.9 *Let us suppose that* (3.33)*,* (3.35)*,* (3.36)*,* (3.37) *hold; then for μ_0-a.e. initial datum $x \in \mathscr{V}$ the graph in $(0,T) \times \overline{\mathscr{V}}$ of each maximal forward solution of the ODE* (3.8) *(for $s = 0$) belongs to the open set P; in particular the maximal solution is unique by* (3.35)*,* (3.36)*.*

Proof. For every $x \in P_0$ let $[0, \tau(x))$ be the open domain of existence and uniqueness of the maximal solution of the system (3.8) restricted to P; classical results on perturbation of ordinary differential equations show that the map $x \mapsto \tau(x)$ is lower semicontinuous in P_0, so that the set

$$D := \Big\{(t,x) \in [0,T) \times P_0 : t < \tau(x)\Big\} \quad \text{is open in } [0,T) \times \mathscr{V},$$

$\boldsymbol{X}$ is of class C^1 in D, and

$$r \in (0,T), \quad \limsup_{t \uparrow r} |\boldsymbol{X}_t(x)| + \frac{1}{\beta_t(\boldsymbol{X}_t(x))} < +\infty \qquad \Longrightarrow \qquad \tau(x) > r. \tag{3.38}$$

Moreover, by integrating $\int_0^{\tau(x)} |\frac{d}{dt} X_t(x)|\, dt$ with respect to μ_0 (see for instance [6, Proposition 8.1.8] for details), one obtains that for μ_0-a.e. x the map $\boldsymbol{X}_\cdot(x)$ is bounded in $(0, \tau(x))$, so that (3.38) implies that $\tau(x)$ can be strictly less than T only if $\beta_t(X_t(x))$ approaches 0 as $t \uparrow \tau(x)$.

Let us denote by ϖ_t the divergence $D_x \cdot \boldsymbol{v}_t$ of the vectorfield $\boldsymbol{v}_t$; since β is of class C^1 in P it is a classical solution of the continuity equation in P and therefore a simple computation gives

$$\frac{\partial}{\partial t} \log(\beta_t) + \boldsymbol{v}_t \cdot D_x \log \beta_t = -\varpi_t \quad \text{in } P, \tag{3.39}$$

so that

$$\frac{d}{dt} \log\Big(\beta_t(\boldsymbol{X}(t,x))\Big) = -\varpi(t, \boldsymbol{X}(t,x)) \qquad t \in (0, \tau(x)), \tag{3.40}$$

and

$$\beta_t(\boldsymbol{X}_t(x)) \det D_x \boldsymbol{X}_t(x) = \beta_0(x), \quad \forall t \in (0, \tau(x)). \tag{3.41}$$

We introduce the decreasing family of open sets

$$E_\varepsilon := \Big\{x \in P_0 : \tau(x) > \varepsilon\Big\}$$

whose union is P_0 and for every $\varepsilon > 0$ we get

$$\int_{E_\varepsilon} \sup_{(\varepsilon,\tau(x))} \log\Big(\beta_\varepsilon(\boldsymbol{X}_\varepsilon(x))/\beta_t(\boldsymbol{X}_t(x))\Big)\, d\mu_0(x) \le \int_{E_\varepsilon}\int_\varepsilon^{\tau(x)} \varpi_t^+(\boldsymbol{X}_t(x))\, dt\, d\mu_0(x)$$
$$= \int_\varepsilon^T \int_{E_t} \varpi_t^+(\boldsymbol{X}_t(x))\, \beta_t(\boldsymbol{X}_t(x)) \det D_x\boldsymbol{X}_t(x)\, dx\, dt = \int_\varepsilon^T \int_{\boldsymbol{X}_t(E_t)} \varpi_t^+(y)\beta_t(y)\, dy\, dt$$
$$\le \int_\varepsilon^T \int_{\mathscr{V}} \varpi_t^+(x)\, d\mu_t(x)\, dt < +\infty,$$

by (3.37). It follows that $\beta_t(\boldsymbol{X}_t(x))$ is bounded away from 0 on $(\varepsilon,\tau(x))$ for μ_0-a.e. $x \in E_\varepsilon$ and therefore $\tau(x) = T$ for μ_0-a.e. $x \in E_\varepsilon$. Taking a sequence $\varepsilon_n \to 0$ and recalling that the union of E_{ε_n} is P_0, and that $\mu_0(\mathscr{V} \setminus P_0) = 0$, we conclude that $\tau(x) = T$ for μ_0-a.e. $x \in \mathscr{V}$. □

Remark 3.10 Recalling (3.39), condition (3.37) surely holds if

$$\int_\varepsilon^T \int_{\{x:\beta_t(x)>0\}} \big(\partial_t\beta_t\big)^- + \big(\boldsymbol{v}_t \cdot D_x\beta_t\big)^-\, dx\, dt < +\infty. \tag{3.42}$$

Remark 3.11 Under the same assumptions of Lemma 3.9, setting $V_\varepsilon := \big\{x \in P_0 : \tau(x) > T-\varepsilon\big\}$ we get a family of open subsets $V_\varepsilon \subset P_0$ such that

$$\mu_0(\mathscr{V} \setminus V_\varepsilon) = 0 \ \text{ and the restriction of } \boldsymbol{X} \text{ to } [0, T-\varepsilon] \times V_\varepsilon \text{ is of class } C^1. \tag{3.43}$$

Starting from (3.39) and arguing as in Lemma 3.9, it is easy to check that if (3.37) is replaced by the stronger condition

$$\int_\varepsilon^T \sup_{x:\beta_t(x)>0} [D_x \cdot \boldsymbol{v}_t(x)]^+\, dt < +\infty \qquad \forall \varepsilon > 0, \tag{3.44}$$

then maximal solutions are unique for *every* $x_0 \in P_0$ and $\boldsymbol{X}$ is of class C^1 in $[0,T] \times P_0$.

4 Gradient flows and convergence of iterated transport maps

Subdifferentials, slopes, and Gradient flows in $\mathscr{P}_2(\mathscr{V})$

In this section we assume that $p = 2$ and that $\mathscr{V}$ is an open subset of $\mathbb{R}^d$ with $\mathscr{L}^d(\partial\mathscr{V}) = 0$. In order to be formally consistent with the theory developped in [6], we identify $\mathscr{P}_2(\mathscr{V})$ with the set of measures $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ such that $\mu(\mathbb{R}^d \setminus \mathscr{V}) = 0$ and we consider a proper functional $\phi : \mathscr{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ which is lower semicontinuous with respect to the *narrow* convergence of $\mathscr{P}(\mathbb{R}^d)$ on the bounded sets of $\mathscr{P}_2(\mathbb{R}^d)$, i.e.

$$\mu_n \to \mu \quad \text{narrowly}, \quad \sup_n \int_{\mathbb{R}^d} |x|^2\, d\mu_n(x) < +\infty \quad \Rightarrow \quad \liminf_{n\to\infty} \phi(\mu_n) \ge \phi(\mu); \tag{4.1}$$

we also assume that

$$\inf_{\mathscr{P}_2(\mathbb{R}^d)} \phi > -\infty \text{ and } \phi(\mu) = +\infty \text{ for any } \mu \in \mathscr{P}_2(\mathbb{R}^d) \setminus \mathscr{P}_2^r(\mathscr{V}). \tag{4.2}$$

$\mathrm{Dom}\,(\phi) \subset \mathscr{P}_2^r(\mathscr{V})$ denotes the domain of finiteness of the functional.

Remark 4.1 The second of the assumptions (4.2) on ϕ, namely that ϕ can be finite only on absolutely continuous measures, has been made only to simplify the exposition, also because it is fulfilled in Example 4.4 below and in the context described in the next section.

In order to define in a precise way the concept of "gradient flow" of ϕ in $\mathscr{P}_2(\mathscr{V})$, we introduce the notion of *strong* and *limiting* subdifferential (see [6, 10.1.1 and 11.1.5], and more generally Chapter 10 of [6] for a systematic development of the subdifferential calculus in spaces of probability measures).

Definition 4.2 (Strong and limiting subdifferentials in $\mathscr{P}_2(\mathscr{V})$) *We say that $\boldsymbol{\xi} \in L^2(\mu;\mathbb{R}^d)$ belongs to the* strong subdifferential $\partial_s\phi(\mu)$ *of ϕ at μ if $\mu \in \mathrm{Dom}\,(\phi)$ and*

$$\phi(\boldsymbol{t}_{\#}\mu) - \phi(\mu) \geq \int_{\mathscr{V}} \langle \boldsymbol{\xi}(x), \boldsymbol{t}(x) - x\rangle \, d\mu(x) + o\big(\|\boldsymbol{t} - \boldsymbol{i}\|_{L^2(\mu;\mathbb{R}^d)}\big). \tag{4.3}$$

We say that $\boldsymbol{\xi} \in L^2(\mu;\mathbb{R}^d)$ belongs to the limiting subdifferential $\partial_\ell\phi(\mu)$ *of ϕ at μ if $\mu \in \mathrm{Dom}\,(\phi)$ and there exist sequences $\boldsymbol{\xi}_k \in \partial_s\phi(\mu_k)$ such that*

$$\mu_k \to \mu \quad \textit{narrowly in } \mathscr{P}(\mathscr{V}), \qquad \boldsymbol{\xi}_k\mu_k \to \boldsymbol{\xi}\mu \quad \textit{in the sense of distributions in } \mathcal{D}'(\mathscr{V}), \tag{4.4}$$

$$\sup_k \Big(\phi(\mu_k), \int_{\mathscr{V}} \big(|x|^2 + |\boldsymbol{\xi}_k(x)|^2\big)\, d\mu_k(x)\Big) < +\infty. \tag{4.5}$$

We also set

$$|\partial_\ell\phi|(\mu) := \inf\Big\{\|\boldsymbol{\xi}\|_{L^2(\mu;\mathbb{R}^d)} : \boldsymbol{\xi} \in \partial_\ell\phi(\mu)\Big\}, \qquad \textit{with the convention} \qquad \inf\emptyset = +\infty. \tag{4.6}$$

Thanks to (4.2) and (4.5), (4.4) is also equivalent to the apparently stronger condition

$$\mu_k \to \mu \quad \text{narrowly in } \mathscr{P}(\mathbb{R}^d), \qquad \boldsymbol{\xi}_k\mu_k \to \boldsymbol{\xi}\mu \quad \text{in the sense of distributions of } \mathcal{D}'(\mathbb{R}^d). \tag{4.5'}$$

Definition 4.3 (Gradient flow) *We say that $\mu_t \in \mathscr{F}_2(\mathscr{V})$ is a gradient flow relative to ϕ with initial datum $\bar\mu \in \mathrm{Dom}\,(\phi)$ if $\phi(\mu_t) \leq \phi(\bar\mu)$, $\mu_t \to \bar\mu$ in $\mathscr{P}_2(\mathscr{V})$ as $t \downarrow 0$, and its tangent velocity field $\boldsymbol{v}_t$ satisfies*

$$-\boldsymbol{v}_t \in \partial_\ell\phi(\mu_t) \qquad \textit{for } \mathscr{L}^1\textit{-a.e. } t \in (0,T). \tag{4.7}$$

We recall here the main example we are interested in, referring to [6, Chap. 10 and 11] for other applications.

Example 4.4 (The *internal energy* functional) Let us assume that

$$\psi : [0, +\infty) \to (-\infty, +\infty] \text{ is a } \textit{convex l.s.c. function with superlinear growth at infinity} \tag{$\psi\,1$}$$

which is differentiable in $(0, +\infty)$ and satisfies

$$\begin{cases} \psi(0) = 0, \\ \big(\psi(s)\big)^- \le Cs^\alpha \quad \text{for some } \alpha > 1 - \frac{d}{d+2} \end{cases} \qquad \text{if } \mathscr{L}^d(\mathscr{V}) = +\infty. \tag{$\psi\,2$}$$

We consider the lower semicontinuous functional

$$\phi(\mu) := \begin{cases} \displaystyle\int_{\mathscr{V}} \psi\left(\beta(x)\right) dx & \text{if } \mu = \beta \cdot \mathscr{L}^d \in \mathscr{P}^r(\mathscr{V}), \\ +\infty & \text{if } \mu \in \mathscr{P}_2(\mathscr{V}) \setminus \mathscr{P}_2^r(\mathscr{V}), \end{cases}$$

which obviously satisfies conditions (4.1) and (4.2). If

$$\psi(0) < +\infty, \qquad \mu = \beta \cdot \mathscr{L}^d \in \mathscr{P}_2^r(\mathscr{V}), \quad \text{with} \quad \beta \in L^\infty(\mathscr{V}) \tag{4.8}$$

and we set

$$L_\psi(\beta) := \beta\psi'(\beta) - \psi(\beta), \tag{4.9}$$

then it is possible to show (see [6, Example 11.1.9]; observe that even when $\mathscr{L}^d(\mathscr{V}) = +\infty$ the inequality $0 \le L_\psi(\beta) \le \beta\psi'\big(\|\beta\|_\infty\big) + \big(\psi(\beta)^-\big)$ yields the integrability of L_ψ) that

$$\boldsymbol{\xi} \in \partial_\ell\phi(\mu) \quad \Rightarrow \quad L_\psi(\beta) \in W^{1,1}(\mathscr{V}) \text{ and } \nabla L_\psi(\beta) = \beta\boldsymbol{\xi}, \tag{4.10}$$

so that

$$\text{if} \quad \partial_\ell\phi(\mu) \neq \emptyset \quad \text{then} \quad \partial_\ell\phi(\mu) \quad \text{contains the unique element} \quad \boldsymbol{\xi} = \frac{\nabla L_\psi(\beta)}{\beta} \in L^2(\mu; \mathbb{R}^d). \tag{4.11}$$

The L^2-norm of $\boldsymbol{\xi}$ is a crucial quantity which we call

$$\mathscr{I}(\beta) := \int_{\mathscr{V}} \left|\frac{\nabla L_\psi(\beta)}{\beta}\right|^2 \beta\, dx = \int_{\mathscr{V}} \frac{|\nabla L_\psi(\beta)|^2}{\beta}\, dx. \tag{4.12}$$

Thus, if a curve $\mu_t = \beta_t\mathscr{L}^d \in \mathscr{F}_2(\mathscr{V})$ with $\beta \in L^\infty((0,T) \times \mathscr{V})$ is a gradient flow relative to ϕ starting from $\bar\mu = \bar\beta\mathscr{L}^d \in \mathrm{Dom}\,(\phi)$ with $\bar\beta \in L^\infty(\mathscr{V})$, then

$$L_\psi(\beta) \in L^2(0,T; W^{1,2}(\mathscr{V})), \qquad \int_0^T \mathscr{I}(\beta_t)\, dt < +\infty, \quad \phi(\beta_t) \le \phi(\bar\beta) < +\infty, \tag{4.13}$$

and β_t solves the nonlinear diffusion PDE

$$\partial_t\beta_t - \Delta L_\psi(\beta_t) = 0 \quad \text{in } (0,T) \times \mathscr{V}, \quad \partial_{\boldsymbol{n}} L_\psi(\beta_t) = 0 \quad \text{on } (0,T) \times \partial\mathscr{V}, \tag{4.14}$$

in the following weak sense

$$\frac{d}{dt}\int_{\mathscr{V}} \zeta(x)\beta_t(x)\,dx + \int_{\mathscr{V}} \nabla L_\psi(\beta_t(x))\cdot\nabla\zeta(x)\,dx = 0 \quad \forall\,\zeta\in C_c^\infty(\mathbb{R}^d). \tag{4.15}$$

In particular, integrating by parts against test function ζ with $\partial_{\boldsymbol{n}}\zeta = 0$ on $\partial\mathscr{V}$, we shall also see that β_t is the *unique* [11] weak solution of (4.15) satisfying

$$\frac{d}{dt}\int_{\mathscr{V}} \zeta(x)\beta_t(x)\,dx = \int_{\mathscr{V}} L_\psi(\beta_t(x))\Delta\zeta(x)\,dx \quad \forall\,\zeta\in C_c^\infty(\mathbb{R}^d), \quad \partial_{\boldsymbol{n}}\zeta = 0 \text{ on } \partial\mathscr{V}. \tag{4.16}$$

We introduce the metric counterparts to the notion of strong and limit subdifferentials:

Definition 4.5 (Local and relaxed metric slopes) *The* local metric slope *of a l.s.c. functional $\phi : \mathscr{P}_2(\mathscr{V}) \to (-\infty, +\infty]$ is defined as*

$$|\partial\phi|(\mu) := \limsup_{W(\mu,\nu)\to 0} \frac{\big(\phi(\mu)-\phi(\nu)\big)^+}{W(\mu,\nu)} \qquad \forall\,\mu\in \mathrm{Dom}\,(\phi). \tag{4.17}$$

Its narrow relaxation *is*

$$|\partial^-\phi|(\mu) := \inf\Big\{ \liminf_{n\to\infty} |\partial\phi|(\mu_n) : \mu_n\to\mu \textit{ narrowly}, \quad \sup_n \big\{W(\mu_n,\mu),\phi(\mu_n)\big\} < +\infty\Big\}. \tag{4.18}$$

A simple link between the vector and the metric concepts is discussed in the next lemma:

Lemma 4.6 (Comparison between subdifferentials and slope) *For every $\mu \in \mathrm{Dom}\,(\phi)$ and $\boldsymbol{\xi}\in\partial_s\phi(\mu)$ we have*

$$|\partial_\ell\phi|(\mu) \le |\partial^-\phi|(\mu) \le |\partial\phi|(\mu) \le \|\boldsymbol{\xi}\|_{L^2(\mu;\mathbb{R}^d)}. \tag{4.19}$$

Proof. It is obvious that the relaxed slope $|\partial^-\phi|(\mu)$ cannot be greater than $|\partial\phi|(\mu)$, which is also bounded by the norm of any element in $\partial_s\phi(\mu)$ simply by its very definition (4.3).

In order to prove the first inequality in (4.19) let us choose $\varepsilon > 0$ and measures μ_k such that

$$\mu_k\to\mu \quad \text{narrowly}, \quad \sup_k \big\{W(\mu_k,\mu),\phi(\mu_k)\big\} < +\infty, \quad \lim_{k\to\infty} |\partial\phi|(\mu_k) \le |\partial^-\phi|(\mu)+\varepsilon,$$

as in (4.18). Combining Lemma 10.1.2, 3.1.3, and 3.1.5 of [6], we find measures $\tilde\mu_k$ and vectors $\tilde{\boldsymbol{\xi}}_k \in \partial_s\phi(\tilde\mu_k)$ such that

$$W(\mu_k,\tilde\mu_k) \le k^{-1}, \quad \|\tilde{\boldsymbol{\xi}}_k\|_{L^2(\tilde\mu_k;\mathbb{R}^d)} \le |\partial\phi|(\mu_k)+k^{-1}, \quad \phi(\tilde\mu_k)\le\phi(\mu_k), \tag{4.20}$$

so that (up to extracting a subsequence such that $\boldsymbol{\xi}_k\mu_k$ is converging in the distribution sense; using the L^2 bound on $\boldsymbol{\xi}_k$ it is easy to check that the limit is representable as $\boldsymbol{\xi}\mu$) we find a limiting subdifferential $\boldsymbol{\xi}_\varepsilon \in \partial_\ell\phi(\mu)$ with $\|\boldsymbol{\xi}\| \le |\partial^-\phi|(\mu)+\varepsilon$. Being $\varepsilon > 0$ arbitrary, we get the thesis. □

The "Minimizing Movement" approximation scheme. One of the possible ways to show existence of gradient flows is to prove the convergence (up to subsequence) of the time discretization of (4.7) by means of a variational formulation of the implicit Euler scheme (we refer to [6] for a more general discussion of this approach and an up-to-date bibliography). Specifically, given a time step $\tau > 0$ and $\bar\mu \in \mathrm{Dom}\,(\phi)$, we recursively define a sequence of measures μ^k in such a way that $\mu^0 = \bar\mu$ and

$$\mu^k \text{ minimizes } \mu \mapsto \frac{1}{2\tau} W_2^2(\mu, \mu^{k-1}) + \phi(\mu) \tag{4.21}$$

for any integer $k \geq 1$. Then, we can define a piecewise constant discrete solution $\overline{M}_\tau : [0, +\infty) \to \mathscr{P}_2(\mathscr{V})$ by

$$\overline{M}_{\tau,t} := \mu^k \qquad \text{if } t \in \big((k-1)\tau, k\tau\big]; \tag{4.22}$$

analogously, denoting by $\boldsymbol{t}^k$ the optimal transport map between μ^{k-1} and μ^k, with inverse $\boldsymbol{s}^k$ we can define a piecewise constant (with respect to time) velocity field by

$$\overline{\boldsymbol{V}}_{\tau,t} := \frac{\boldsymbol{i} - \boldsymbol{s}^k}{\tau} \in L^2(\overline{M}_{\tau,t}; \mathbb{R}^d) \quad \text{with} \quad -\overline{\boldsymbol{V}}_{\tau,t} \in \partial_s \phi(\overline{M}_{\tau,t}) \quad \text{for } t \in \big((k-1)\tau, k\tau\big]. \tag{4.23}$$

With this notation the following energy convergence result holds:

Theorem 4.7 (Convergence of discrete approximations and Gradient flows) *Let us assume that ϕ satisfies* (4.1) *and* (4.2) *and that*

$$\partial_\ell \phi(\mu) \quad \textit{contains at most one vector.} \tag{4.24}$$

For every $\bar\mu \in \mathrm{Dom}\,(\phi)$ there exists a vanishing subsequence of time steps $\tau_n \downarrow 0$ and a curve $\mu_t \in \mathscr{F}_2(\mathscr{V})$ such that the discrete solutions $\overline{M}_{\tau_n,t}$, narrowly converge to μ_t as $n \uparrow \infty$ for every $t \in [0,T]$ and $\overline{\boldsymbol{V}}_{\tau_n} \overline{M}_{\tau_n}$ converge to $\overline{\boldsymbol{v}}\mu$ in $\mathcal{D}'((0,T) \times \mathscr{V})$, where $\overline{\boldsymbol{v}}_t$ satisfies

$$\begin{gathered} \partial_t \mu_t + D \cdot (\overline{\boldsymbol{v}}_t \mu_t) = 0 \quad \textit{in } \mathcal{D}'((0,T) \times \mathbb{R}^d), \quad -\overline{\boldsymbol{v}}_t = \partial_\ell \phi(\mu_t) \quad \textit{for } \mathscr{L}^1\textit{-a.e. } t \in (0,T), \\ \int_0^T \|\overline{\boldsymbol{v}}_t\|^2_{L^2(\mu_t;\mathbb{R}^d)}\, dt < +\infty. \end{gathered} \tag{4.25}$$

Moreover, if μ_t satisfies the "upper gradient inequality"

$$\phi(\mu_t) + \int_0^t |\partial^- \phi|(\mu_s) \cdot \|\boldsymbol{v}_s\|_{L^2(\mu_s;\mathbb{R}^d)}\, ds \geq \phi(\bar\mu) \qquad \forall\, t \in [0,T], \tag{4.26}$$

where $\boldsymbol{v}_t$ is the tangent velocity field to μ_t, then

$$\overline{\boldsymbol{v}}_t = \boldsymbol{v}_t \quad \mu_t\textit{-a.e., for } \mathscr{L}^1\textit{-a.e. } t \in (0,T), \tag{4.27}$$

μ_t is a gradient flow relative to ϕ according to (4.7),

$$\lim_{n\to\infty} W_2(\overline{M}_{\tau_n,t}, \mu_t) = 0, \qquad \lim_{n\to+\infty} \phi(\overline{M}_{\tau_n,t}) = \phi(\mu_t) \qquad \forall\, t \in [0,T], \tag{4.28}$$

the discrete velocity fields $\overline{\boldsymbol{V}}_{\tau_n,t}$ satisfy

$$\lim_{n\to\infty}\int_0^T\int_{\mathscr{V}}|\overline{\boldsymbol{V}}_{\tau_n,t}|^2\,d\overline{M}_{\tau_n,t}\,dt=\int_0^T\int_{\mathscr{V}}|\boldsymbol{v}_t|^2\,d\mu_t\,dt,\tag{4.29}$$

the map $t\mapsto\phi(\mu_t)$ is absolutely continuous, and finally

$$\frac{d}{dt}\phi(\mu_t)=-\int_{\mathscr{V}}|\boldsymbol{v}_t(x)|^2\,d\mu_t(x)\qquad\textit{for }\mathscr{L}^1\textit{-a.e. }t\in(0,T).\tag{4.30}$$

The proof combines various *a priori* estimates and a deep variational interpolation argument due to De Giorgi: it allows to derive a discrete energy identity that gives in the limit the (standard) continuous energy identity, the absolute continuity of $t\mapsto\phi(\mu_t)$ and the convergence of all discrete quantities to their continuous counterpart. Related results, are treated in [28] (the seminal paper on this subject), [34], [1]; a comprehensive convergence scheme at the PDE level is also illustrated in §11.1 of [6].

Proof. Theorem 11.1.6 and Corollary 11.1.8 of [6] yield the pointwise narrow convergence of $\overline{M}_{\tau_n}$ to μ, the distributional convergence of $\overline{\boldsymbol{V}}_{\tau_n}\overline{M}_{\tau_n}$ to $\overline{\boldsymbol{v}}\mu$ and (4.25).

If $\boldsymbol{v}_t$ is the velocity vector field associated to the curve μ_t, we have

$$|\mu'|(t)=\lim_{h\to0}\frac{W(\mu_{t+h},\mu_t)}{|h|}=\|\boldsymbol{v}_t\|_{L^2(\mu_t;\mathbb{R}^d)}\le\|\overline{\boldsymbol{v}}_t\|_{L^2(\mu_t;\mathbb{R}^d)}\quad\text{for }\mathscr{L}^1\text{-a.e. }t\in(0,T).\tag{4.31}$$

In order to prove the second part of the Theorem, we are introducing the so called "De Giorgi variational interpolants" $\tilde{M}_{\tau,t}$, which are defined as

$$\tilde{M}_{\tau,t}\quad\text{minimizes}\quad\mu\mapsto\frac{1}{2\sigma}W_2^2(\mu,\mu^{k-1})+\phi(\mu)\quad\text{if }t=(k-1)\tau+\sigma,\quad0<\sigma\le\tau,\tag{4.32}$$

(choosing when $\sigma=\tau$ $\tilde{M}_{\tau,t}=\mu^k$) together with the related optimal transport maps $\tilde{\boldsymbol{s}}_t$ which push $\tilde{M}_{\tau,t}$ on μ^{k-1} for $t=(k-1)\tau+\sigma$, and the velocities

$$\tilde{\boldsymbol{V}}_{\tau,t}:=\frac{\boldsymbol{i}-\tilde{\boldsymbol{s}}_t}{\sigma}\in L^2(\tilde{M}_{\tau,t};\mathbb{R}^d),\qquad-\tilde{\boldsymbol{V}}_{\tau,t}\in\partial_s\phi(\tilde{M}_{\tau,t}).\tag{4.33}$$

The interest of De Giorgi's interpolants relies in the following refined *discrete energy identity* (see Lemma [6, 3.2.2])

$$\frac12\int_0^t\int_{\mathscr{V}}|\overline{\boldsymbol{V}}_{\tau,s}|^2\,d\overline{M}_{\tau,s}\,ds+\frac12\int_0^t\int_{\mathscr{V}}|\tilde{\boldsymbol{V}}_{\tau,s}|^2\,d\tilde{M}_{\tau,s}\,ds+\phi(\overline{M}_{\tau,t})=\phi(\bar{\mu})\quad\text{if }t/\tau\in\mathbb{N}.\tag{4.34}$$

Since $\tilde{M}_{\tau_n,t}$ still narrowly converge to μ_t [6, Cor. 3.3.4], we pass to the limit in the above identity as $\tau_n\to0$ by using Fatou's Lemma, the narrow lower semicontinuity of ϕ, and the very definition of relaxed slope (4.18); by (4.26) we obtain for every $t\in[0,T]$

$$\limsup_{n\to\infty}\frac12\int_0^t\int_{\mathscr{V}}|\overline{\boldsymbol{V}}_{\tau_n,s}|^2\,d\overline{M}_{\tau_n,s}\,ds+\frac12\int_0^t|\partial^-\phi|^2(\mu_s)\,ds+\phi(\mu_t)\le\phi(\bar{\mu})\tag{4.35}$$

$$\le\phi(\mu_t)+\int_0^t|\partial^-\phi|(\mu_s)\cdot\|\boldsymbol{v}_s\|_{L^2(\mu_s;\mathbb{R}^d)}\,ds,\tag{4.36}$$

i.e.

$$\limsup_{n\to\infty} \frac{1}{2}\int_0^t \int_{\mathscr{V}} |\overline{\boldsymbol{V}}_{\tau_n,s}|^2 \, d\overline{M}_{\tau_n,s}\, ds - \frac{1}{2}\int_0^t \int_{\mathscr{V}} |\boldsymbol{v}_s|^2 \, d\mu_s \, ds + \frac{1}{2}\int_0^t \Big(|\partial^-\phi|(\mu_s) - \|\boldsymbol{v}_s\|_{L^2(\mu_s;\mathbb{R}^d)}\Big)^2 ds \le 0.$$

Since general lower semicontinuity results yield (see Theorem 5.4.4 in [6])

$$\liminf_{n\to\infty} \frac{1}{2}\int_0^t \int_{\mathscr{V}} |\overline{\boldsymbol{V}}_{\tau_n,s}|^2 \, d\overline{M}_{\tau_n,s}\, ds \ge \frac{1}{2}\int_0^t \int_{\mathscr{V}} |\overline{\boldsymbol{v}}_s|^2 \, d\mu_s \, ds \ge \frac{1}{2}\int_0^t \int_{\mathscr{V}} |\boldsymbol{v}_s|^2 \, d\mu_s \, ds,$$

taking into account (4.19) we conclude that

$$\int_0^T \|\overline{\boldsymbol{V}}_{\tau_n,t}\|^2_{L^2(M_{\tau_n,t};\mathbb{R}^d)}\, dt \to \int_0^T \|\boldsymbol{v}_t\|^2_{L^2(\mu_t;\mathbb{R}^d)}\, dt,$$

$$\|\boldsymbol{v}_t\|_{L^2(\mu_t;\mathbb{R}^d)} = \|\overline{\boldsymbol{v}}_t\|_{L^2(\mu_t;\mathbb{R}^d)} = |\partial^-\phi|(\mu_t) = |\partial_\ell\phi|(\mu_t) \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T),$$

and, using again (4.35), $\phi(M_{\tau_n,t}) \to \phi(\mu_t)$ for all $[0,T]$ and

$$\phi(\mu_t) = \phi(\bar\mu) - \int_0^t |\partial^-\phi|(\mu_s)\cdot \|\boldsymbol{v}_s\|_{L^2(\mu_s;\mathbb{R}^d)}\, ds \quad \forall t \in [0,T].$$

Hence the map $t \mapsto \phi(\mu_t)$ is absolutely continuous and

$$\frac{d}{dt}\phi(\mu_t) = -\|\boldsymbol{v}_t\|^2_{L^2(\mu_t;\mathbb{R}^d)} \quad \mathscr{L}^1\text{-a.e. in } (0,T).$$

By the minimality of the norm of $\boldsymbol{v}_t$ among all the possible vector fields satisfying the continuity equation (4.25) we also deduce (4.27).

Finally, in order to check the convergence of $\overline{M}_{\tau_n,t}$ to μ_t in $\mathscr{P}_2(\mathbb{R}^d)$ we apply (2.6) and we simply show the convergence of the quadratic moment of $\overline{M}_{\tau_n,t}$. Recall that, if $t \in \big((m-1)\tau, m\tau\big]$

$$\begin{aligned}\int_{\mathbb{R}^d} |x|^2 \, dM_{\tau,t}(x) - \int_{\mathbb{R}^d} |x|^2 \, d\bar\mu(x) &= \sum_{j=1}^m \int_{\mathbb{R}^d} |x|^2 \, d\mu^j(x) - \int_{\mathbb{R}^d} |x|^2 \, d\mu^{j-1}(x) \\ &\le 2\sum_{j=1}^m \int_{\mathbb{R}^d} \langle x - \boldsymbol{s}^j(x), x\rangle \, d\mu^j(x) = 2\int_0^{m\tau} \int_{\mathbb{R}^d} \langle \overline{\boldsymbol{V}}_{\tau,t}(x), x\rangle \, d\overline{M}_{\tau,t}(x),\end{aligned}$$

whereas for the absolutely continuous curve μ_t

$$\int_{\mathbb{R}^d} |x|^2 \, d\mu_t(x) - \int_{\mathbb{R}^d} |x|^2 \, d\bar\mu(x) = 2\int_0^t \int_{\mathbb{R}^d} \langle \boldsymbol{v}_s(x), x\rangle \, d\mu_s(x).$$

Taking into account (4.29) and arguing as in [6, Lemma 5.2.4], we obtain

$$\limsup_{n\to\infty} \int_{\mathbb{R}^d} |x|^2 \, dM_{\tau_n,t} \le \int_{\mathbb{R}^d} |x|^2 \, d\mu_t,$$

which yields the convergence of the quadratic moments. $\square$

(4.24) and the general properties (4.1), (4.2) are not particularly restrictive, as they are easily checked on many examples; the crucial assumption of the previous theorem is in fact the "upper gradient" inequality (4.26).

With respect to the results of [6, Chap.2], here there is a slight improvement: it is sufficient to chek this inequality only on the limit curves arising from the "Minimizing Movement" scheme (instead of proving it on all the curves with finite energy). This simple remark is quite useful to show that bounded solutions to the nonlinear diffusion equation (4.14) we introduced in Example 4.4, satisfy the "upper gradient inequality" (4.26).

Proposition 4.8 (The Gradient flow of the internal energy functional) *Let us assume that the function ψ of Example 4.4 satisfies $(\psi\,1)$, $(\psi\,2)$,*

$$\psi \quad \textit{is smooth in } (0,+\infty) \textit{ with } \psi''>0, \textit{ and } \psi(0)<+\infty. \tag{$\psi\,3$}$$

We also suppose that the initial datum $\bar\mu=\bar\beta\mathscr{L}^d$ satisfies

$$\sup_{x\in\mathscr{V}}\bar\beta(x)<+\infty,\quad \phi(\bar\mu)=\int_{\mathscr{V}}\psi(\bar\beta(x))\,dx<+\infty,\quad \int_{\mathscr{V}}|x|^2\bar\beta(x)\,dx<+\infty. \tag{$\psi\,4$}$$

Then the discrete solutions $M_{\tau,t}$ of the Minimizing Movement scheme converge pointwise to $\mu_t:=\beta_t\mathscr{L}^d$ in $\mathscr{P}_2(\mathscr{V})$ as $\tau\downarrow 0$, β_t is the unique solution in $L^\infty((0,T)\times\mathscr{V})$ of (4.14) *with the integrability conditions* (4.13)*, and β_t satisfies the energy identity*

$$\int_{\mathscr{V}}\psi(\beta_t)\,dx+\int_0^t\int_{\mathscr{V}}\frac{|\nabla L_\psi(\beta_s)|^2}{\beta_s}\,dx\,ds=\int_{\mathscr{V}}\psi(\bar\beta)\,dx\quad\forall\,t\in[0,T]. \tag{4.37}$$

Moreover, the internal energy functional introduced in Example 4.4 satisfies the "upper gradient inequality" along any limit curve $\mu_t=\beta_t\mathscr{L}^d$ and therefore all the convergence properties of Theorem 4.7 hold true.

Proof. Up to perturbing ψ by an additive constant, it is not restrictive to assume that $\psi(0)=0$.

By applying the first part of Theorem 4.7 and the discrete L^∞-estimates of [36], [1], we obtain that any limit curve $\mu_t:=\beta_t\mathscr{L}^d$ of the Minimizing Movement scheme is a uniformly bounded weak solution (according to (4.16)) of (4.14) satisfying (4.13). Since bounded weak solutions are unique [11], we obtain the convergence of the whole sequence $M_{\tau,t}$; standard estimates on nonlinear diffusion equations show that $\|\beta_t\|_{L^\infty(\mathscr{V})}\le M_\infty:=\|\bar\beta\|_{L^\infty(\mathscr{V})}$.

It remains to check the validity of the upper gradient inequality. Let

$$\begin{aligned}\mathscr{S}\quad &:=\quad \bar\beta+\overline{\mathrm{conv}\big\{\beta_t-\bar\beta:t\in[0,T]\big\}}^{L^2(\mathscr{V})}\\ &\subset\quad \Big\{\beta\in L^1(\mathscr{V})\cap L^\infty(\mathscr{V}):\beta\ge 0,\ \|\beta\|_{L^1(\mathscr{V})}=1,\ \|\beta\|_{L^\infty(\mathscr{V})}\le M_\infty\Big\},\end{aligned}$$

and the H^{-1}-like distance on $\mathscr{S}$

$$d(\beta_1,\beta_2):=\sup\Big\{\int_{\mathscr{V}}\zeta(x)\,(\beta_1-\beta_2)\,dx:\zeta\in C^\infty_c(\mathbb{R}^d),\quad \|\nabla\zeta\|_{L^2(\mathscr{V};\mathbb{R}^d)}\le 1\Big\}. \tag{4.38}$$

It is not difficult to check that d is finite on $\mathscr{S}$: let us first observe that (4.15), (4.13), and a standard approximation result in $W^{1,2}(\mathscr{V})$, yield

$$\int_{\mathscr{V}} \zeta\big(\beta_{t_1} - \beta_{t_0}\big)\,dx = -\int_{t_0}^{t_1}\int_{\mathscr{V}} \nabla L_\psi(\beta_t)\cdot\nabla\zeta\,dx\,dt \qquad \forall\,\zeta\in W^{1,2}(\mathscr{V}). \tag{4.39}$$

Choosing now a test function $\zeta \in C_c^\infty(\mathbb{R}^d)$ with $\|\nabla\zeta\|_{L^2(\mathscr{V})} \le 1$ and $0 \le t_0 < t_1 \le T$, we get

$$\int_{\mathscr{V}} \zeta\big(\beta_{t_1} - \beta_{t_0}\big)\,dx \le \int_{t_0}^{t_1} \|\nabla L_\psi(\beta_t)\|_{L^2(\mathscr{V};\mathbb{R}^d)}\,dt \le M_\infty^{1/2}\int_{t_0}^{t_1}\sqrt{\mathscr{I}(\beta_t)}\,dt, \tag{4.40}$$

so that by (4.13)

$$d(\beta_{t_1},\beta_{t_0}) \le M_\infty^{1/2}\int_{t_0}^{t_1}\sqrt{\mathscr{I}(\beta_t)}\,dt, \quad \text{and} \quad \int_0^T \mathscr{I}(\beta_t)\,dt < +\infty. \tag{4.41}$$

Let us now introduce the regularized convex functionals

$$\phi_\varepsilon(\beta) := \int_{\mathscr{V}} \psi_\varepsilon(\beta)\,dx \quad \forall\,\beta\in\mathscr{S}, \qquad \psi_\varepsilon(\beta) := \begin{cases} \psi(\beta) + \varepsilon\psi'(\varepsilon) - \psi(\varepsilon) & \text{if } \beta > \varepsilon,\\ \psi'(\varepsilon)\beta & \text{if } 0 \le \beta \le \varepsilon, \end{cases} \tag{4.42}$$

and the related "Lagrangians"

$$L_{\psi_\varepsilon}(\beta) = \begin{cases} L_\psi(\beta) - L_\psi(\varepsilon) & \text{if } \beta > \varepsilon,\\ 0 & \text{if } 0 \le \beta \le \varepsilon. \end{cases} \tag{4.43}$$

Observe that ϕ_ε are geodesically convex functionals on $\mathscr{S}$, since the usual segments $t \mapsto (1-t)\rho_0 + t\rho_1$, $\rho_0,\rho_1\in\mathscr{S}$, are constant speed geodesics in $\mathscr{S}$. An upper bound for the slope of ϕ_ε

$$|\partial\phi_\varepsilon|_{\mathscr{S}}(\beta) := \limsup_{d(\beta,\rho)\to 0} \frac{\big(\phi_\varepsilon(\beta)-\phi_\varepsilon(\rho)\big)^+}{d(\beta,\rho)}$$

can be readily obtained: first of all, we recall that $\mathscr{I}(\beta) < +\infty$ implies $L_\psi(\beta)\in W^{1,2}(\mathscr{V})$ and $|\nabla L_\psi(\beta)|^2/\beta \in L^1(\mathscr{V})$, hence

$$\mathscr{I}(\beta) < +\infty \quad \Longrightarrow \quad \frac{L_{\psi_\varepsilon}(\beta)}{\beta} = \psi_\varepsilon'(\beta) \in W^{1,2}(\mathscr{V}), \quad \|\nabla\psi_\varepsilon'(\beta)\|^2_{L^2(\mathscr{V};\mathbb{R}^d)} \le \varepsilon^{-1}\mathscr{I}(\beta); \tag{4.44}$$

therefore, assuming that $\mathscr{I}(\beta) < +\infty$ and using the convexity of ψ_ε, we get

$$\phi_\varepsilon(\beta) - \phi_\varepsilon(\rho) \le \int_{\mathscr{V}} \psi_\varepsilon'(\beta)(\beta-\rho)\,dx \le d(\beta,\rho)\|\nabla\psi_\varepsilon'(\beta)\|_{L^2(\mathscr{V};\mathbb{R}^d)}, \tag{4.45}$$

and we can use (4.44) to obtain

$$|\partial\phi_\varepsilon|_{\mathscr{S}}(\beta) \le \varepsilon^{-1/2}\sqrt{\mathscr{I}(\beta)}. \tag{4.46}$$

Applying [6, Theorem 1.2.5] and the estimate (4.41) we find that the map $t \mapsto \phi_\varepsilon(\beta_t)$ is absolutely continuous; since $\psi'_\varepsilon(\beta_t) \in W^{1,2}(\mathscr{V})$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$, combining (4.45) with β_t, β_{t+h} instead of β, ρ and (4.39) with $\zeta := \psi'_\varepsilon(\beta_t)$, we get

$$\phi_\varepsilon(\beta_t) - \phi_\varepsilon(\beta_{t+h}) \le \int_{\mathscr{V}} \psi'_\varepsilon(\beta_t)\big(\beta_t - \beta_{t+h}\big)\,dx = \int_t^{t+h} \int_{\mathscr{V}} \nabla L_\psi(\beta_s) \cdot \nabla \psi'_\varepsilon(\beta_t)\,dx\,ds, \tag{4.47}$$

for $\mathscr{L}^1$-a.e. $t \in (0,T)$. Since Lebesgue differentiation Theorem for functions with values in the Hilbert space $L^2(\mathscr{V};\mathbb{R}^d)$ yields

$$\lim_{h\to 0} \int_{\mathscr{V}} \Big|\frac{1}{h}\int_t^{t+h} \nabla L_\psi(\beta_s)\,ds - \nabla L_\psi(\beta_t)\Big|^2\,dx = 0 \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T),$$

dividing by $h \neq 0$ and taking the limit of (4.47) as h goes to 0 from the right and from the left, we find that the derivative of $\phi_\varepsilon \circ \beta$ is

$$\frac{d}{dt}\phi_\varepsilon(\beta_t) = -\int_{\mathscr{V}} \nabla\psi'_\varepsilon(\beta_t) \cdot \nabla L_\psi(\beta_t)\,dx \qquad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T).$$

Since

$$\beta_t \nabla \psi'_\varepsilon(\beta_t) = \nabla L_{\psi_\varepsilon}(\beta_t) = \nabla L_\psi(\beta_t)\chi_{\varepsilon,t}(x) \quad \text{with} \quad \chi_{\varepsilon,t}(x) := \begin{cases} 1 & \text{if } \beta_t(x) > \varepsilon, \\ 0 & \text{if } 0 \le \beta_t(x) \le \varepsilon, \end{cases} \tag{4.48}$$

integrating in time we eventually find

$$\int_{\mathscr{V}} \psi_\varepsilon(\beta_t)\,dx + \int_0^t \int_{\mathscr{V}} \frac{|\nabla L_\psi(\beta_s)|^2}{\beta_s}\chi_{\varepsilon,s}\,dx\,ds = \int_{\mathscr{V}} \psi_\varepsilon(\bar\beta)\,dx. \tag{4.49}$$

Since $\psi_\varepsilon \downarrow \psi$ and it is easy to check that, e.g., $\psi_1(\beta) \in L^1(\mathscr{V})$, we can pass to the limit as $\varepsilon \downarrow 0$ in (4.49) and we find the energy identity (4.37).

The "upper gradient inequality" (4.26) follows immediately if we show that

$$\bar{\boldsymbol{v}}_t = -\partial_\ell \phi(\mu_t) = -\frac{\nabla L_\psi(\beta_t)}{\beta_t} \in \operatorname{Tan}_{\mu_t}\mathscr{P}_2(\mathscr{V}) \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T), \tag{4.50}$$

i.e., according to (3.6), if for $\mathscr{L}^1$-a.e. $t \in (0,T)$ there exists a family of functions $\zeta_\varepsilon \in C^\infty_c(\mathbb{R}^d)$ such that $\nabla\zeta_\varepsilon \to \bar{\boldsymbol{v}}_t$ in $L^2(\mu_t;\mathbb{R}^d)$ as $\varepsilon \downarrow 0$. Since $\beta_t \in L^\infty(\mathscr{V})$, by standard extension and approximation results, it is sufficient to find an approximating sequence $\zeta_\varepsilon \in W^{1,2}(\mathscr{V})$; disregarding a $\mathscr{L}^1$-negligible subset of $(0,T)$ we can assume that $L_\psi(\beta_t) \in W^{1,2}(\mathscr{V})$ and recalling (4.48) we can choose

$$\zeta_\varepsilon := \psi'_\varepsilon(\beta_t) \in W^{1,2}(\mathscr{V}) \quad \text{so that} \quad \nabla\psi'_\varepsilon(\beta_t) = \frac{\nabla L_\psi(\beta_t)}{\beta_t}\chi_{\varepsilon,t} \to \frac{\nabla L_\psi(\beta_t)}{\beta_t} \tag{4.51}$$

in $L^2(\mu_t;\mathbb{R}^d)$ as $\varepsilon \downarrow 0$. □

Remark 4.9 (The case $\psi(0) = +\infty$) When $\psi(0) = +\infty$, we will also assume that $\bar{\beta} \geq \beta_{\min} > 0$ $\mathscr{L}^d$-a.e. in $\mathscr{V}$, so that the maximum principle yields $\beta_t \geq \beta_{\min}$ in $(0,T) \times \mathscr{V}$. In this case $\mathscr{V}$ has to be bounded and the calculations are even easier than in the previous Proposition: e.g. by modifying ψ in the interval $(0, \beta_{\min})$ we directly obtain (4.37) without performing any preliminary regularization of ψ around 0.

Remark 4.10 (The case when ϕ is displacement convex) When $\mathscr{V}$ is a *convex* subset of $\mathbb{R}^d$, ϕ has compact sublevels in $\mathscr{P}_2^r(\mathbb{R}^d)$ (this is always the case if, e.g., $\mathscr{V}$ is *bounded*), and it is *displacement convex*, i.e. for any μ, $\nu \in \mathscr{P}_2^r(\mathscr{V})$, if we denote by $\boldsymbol{t}_\mu^\nu \in L^2(\mu; \mathscr{V})$ the optimal transport map between μ and ν relative to W_2, the map

$$t \mapsto \phi\left(((1-t)\boldsymbol{t}_\mu^\nu + t\boldsymbol{i})_{\#}\mu\right) \quad \text{is convex in } [0,1], \tag{4.52}$$

then the theory becomes considerably simpler: solutions to the gradient flow equation are in fact unique, the (limiting) subdifferential is characterized by the following system of variational inequalities

$$\boldsymbol{\xi} \in \partial_\ell \phi(\mu) \quad \Longleftrightarrow \quad \phi(\nu) \geq \phi(\mu) + \int_{\mathscr{V}} \langle \boldsymbol{\xi}, \boldsymbol{t}_\mu^\nu - \boldsymbol{i} \rangle \, d\mu \qquad \forall \nu \in \mathscr{P}^r(\mathscr{V}), \tag{4.53}$$

and inequality (4.26) is always satisfied by any absolutely continuous curve in $\mathscr{P}_2(\mathbb{R}^d)$ (see Corollary 2.4.11, §10.1.1 and Chapter 11 of [6]).

Notice also that if the stronger property

$$t \mapsto \phi\left(((1-t)\boldsymbol{t}_\mu^\nu + t\boldsymbol{t}_\mu^\sigma)_{\#}\mu\right) \quad \text{is convex in } [0,1] \text{ for any } \mu,\, \nu,\, \sigma \in \mathscr{P}^r(\mathscr{V})$$

holds, then even *error estimates* for the scheme can be proved, see Theorem 4.0.4 and §11.2 of [6], and as a consequence one can also consider initial data that are in $\overline{\mathrm{Dom}\,(\phi)}$.

In the case of the *internal energy* functional of Example 4.4, the assumption of displacement convexity is equivalent to McCann's condition

$$s \mapsto s^d \psi(s^{-d}) \text{ is convex and non increasing in } (0, +\infty), \tag{4.54}$$

which is more restrictive than convexity if the space dimension d is greater than 1.

Convergence of iterated transport maps

In the final part of the present section we study the convergence as $\tau \downarrow 0$ of the iterated transport maps

$$\boldsymbol{T}^k := \boldsymbol{t}^k \circ \boldsymbol{T}^{k-1} = \boldsymbol{t}^k \circ \boldsymbol{t}^{k-1} \circ \cdots \circ \boldsymbol{t}^1, \tag{4.55}$$

associated to the Minimizing Movement scheme (4.21); recall that we denoted by $\boldsymbol{t}^k$ the (unique) optimal transport map pushing μ^{k-1} on μ^k and by $\boldsymbol{s}^k = \left(\boldsymbol{t}^k\right)^{-1}$ its inverse map, pushing μ^k to μ^{k-1}; in particular $\boldsymbol{T}^k$ maps $\bar{\mu} = \mu^0$ to μ^k.

We can "embed" the discrete sequence $\boldsymbol{T}^k$ into a continuous flow $\boldsymbol{T}_{\tau,t}$ such that $\boldsymbol{T}_{\tau,t} = \boldsymbol{T}^k$ if t is one of the nodes $k\tau$ of the discrete partition. To fix a simple notation, in what follows

$$\text{if } t \in \big[(k-1)\tau, k\tau\big) \quad \text{we decompose it as} \quad t = (k-1)\tau + \sigma\tau, \quad \sigma = \Big\{\frac{t}{\tau}\Big\} \in (0,1], \tag{4.56}$$

and we set

$$\boldsymbol{t}^{k-1,\sigma} := (1-\sigma)\boldsymbol{i} + \sigma \boldsymbol{t}^k, \quad \boldsymbol{t}^{\sigma,k} := \boldsymbol{t}^k \circ \big(\boldsymbol{t}^{k-1,\sigma}\big)^{-1}, \quad \text{so that} \quad \boldsymbol{t}^k = \boldsymbol{t}^{\sigma,k} \circ \boldsymbol{t}^{k-1,\sigma}, \tag{4.57}$$

$$\boldsymbol{T}_{\tau,t} := (1-\sigma)\boldsymbol{T}^{k-1} + \sigma \boldsymbol{T}^k = \boldsymbol{t}^{k-1,\sigma} \circ \boldsymbol{T}^{k-1}, \quad \text{so that} \quad \boldsymbol{T}^k = \boldsymbol{t}^{\sigma,k} \circ \boldsymbol{T}_{\tau,t}. \tag{4.58}$$

The continuous family of maps $\boldsymbol{T}_{\tau,t}$ is naturally associated to a sort of "piecewise linear" interpolant (according to McCann's displacement convexity) *continuous* interpolation $\mu_{\tau,t}$ of the sequence of measures μ^k:

$$\mu_{\tau,t} := \Big(\boldsymbol{t}^{k-1,\sigma}\Big)_{\#}\mu^{k-1} = (\boldsymbol{T}_{\tau,t})_{\#}\,\bar\mu, \quad \text{so that} \quad \mu^k = \Big(\boldsymbol{t}^{\sigma,k}\Big)_{\#}\mu_{\tau,t} \qquad \text{if } t \in \big[(k-1)\tau, k\tau\big). \tag{4.59}$$

Notice that the inverse of the map $\boldsymbol{t}^{k-1,\sigma}$ is well defined up to $\mu_{\tau,t}$-negligible sets and it coincides with the optimal transport map between $\mu_{\tau,t}$ and μ^{k-1}; moreover

$$\boldsymbol{t}^{\sigma,k} = \Big(\boldsymbol{t}^{k-1,\sigma} \circ \boldsymbol{s}^k\Big)^{-1} = \Big(\sigma\boldsymbol{i} + (1-\sigma)\boldsymbol{s}^k\Big)^{-1}. \tag{4.60}$$

We define also the velocity vector field $\boldsymbol{v}_{\tau,t}$ associated to this flow

$$\boldsymbol{v}_{\tau,t} := \overline{\boldsymbol{V}}_{\tau,t} \circ \boldsymbol{t}^{\sigma,k} = \frac{\boldsymbol{i} - \boldsymbol{s}^k}{\tau} \circ \boldsymbol{t}^{\sigma,k} \ \in L^2(\mu_{\tau,t};\mathbb{R}^d) \qquad \text{if } t \in \big[(k-1)\tau, k\tau\big). \tag{4.61}$$

Due to the uniform $C^{0,1/2}$ bound

$$\frac{1}{2}\sum_{k=0}^{\infty} W_2^2(\mu^{k+1},\mu^k) \le \tau \sum_{k=0}^{\infty}\Big(\phi(\mu^k) - \phi(\mu^{k+1})\Big) \le \tau\big(\phi(\bar\mu) - \inf\phi\big), \tag{4.62}$$

the convergence statement of Theorem 4.7 applies not only to the discrete piecewise constant solution $\overline{M}_{\tau,t}$, but also to $\mu_{\tau,t}$. The following lemma shows that $\boldsymbol{v}_{\tau,t}$ is an admissible velocity field relative to $\mu_{\tau,t}$ and provides the corresponding energy estimate.

Lemma 4.11 (Properties of the continuous interpolation) *Let $\mu_{\tau,t}$, $\boldsymbol{v}_{\tau,t}$, $\boldsymbol{T}_{\tau,t}$ be defined as in* (4.59), (4.61), and (4.58) *respectively. Then $\boldsymbol{T}_{\tau,t}$ is a flow relative to $(\mu_t, \boldsymbol{v}_{\tau,t})$ and in particular the continuity equation*

$$\frac{d}{dt}\mu_{\tau,t} + D_x \cdot (\boldsymbol{v}_{\tau,t}\mu_{\tau,t}) = 0 \quad \textit{in } \mathscr{D}'\left((0,T)\times\mathscr{V}\right) \tag{4.63}$$

holds. Moreover we have

$$\int_0^T \int_{\mathscr{V}} |\boldsymbol{v}_{\tau,t}|^2 \, d\mu_{\tau,t}\, dt = \int_0^T \int_{\mathscr{V}} |\overline{\boldsymbol{V}}_{\tau,t}|^2 \, d\overline{M}_{\tau,t}\, dt \qquad \forall T > 0. \tag{4.64}$$

Proof. We check first that $\boldsymbol{T}_{\tau,t}$ is a flow relative to $\boldsymbol{v}_{\tau,t}$. Clearly $t \mapsto \boldsymbol{T}_{\tau,t}(x)$ is continuous and piecewise linear; in any interval $\big((k-1)\tau, k\tau\big)$, with k integer, we have that its derivative is given by

$$\frac{d}{dt}\boldsymbol{T}_{\tau,t}(x) = \frac{1}{\tau}\big(\boldsymbol{T}^k(x) - \boldsymbol{T}^{k-1}(x)\big), \tag{4.65}$$

so that (4.61) and (4.58) yield

$$\boldsymbol{v}_{\tau,t}(\boldsymbol{T}_{\tau,t}(x)) = \frac{1}{\tau}(\boldsymbol{i} - \boldsymbol{s}^k) \circ \boldsymbol{t}^{\sigma,k}(\boldsymbol{T}_{\tau,t}(x)) = \frac{1}{\tau}(\boldsymbol{i} - \boldsymbol{s}^k) \circ \boldsymbol{T}^k(x) = \frac{1}{\tau}(\boldsymbol{T}^k(x) - \boldsymbol{T}^{k-1}(x)),$$

and therefore the ODE $\frac{d}{dt}\boldsymbol{T}_{\tau,t}(x) = \boldsymbol{v}_{\tau,t}(\boldsymbol{T}_{\tau,t}(x))$ is satisfied.

Finally, by the definition of $\boldsymbol{v}_{\tau,t}$ and taking into account (4.59) we get

$$\int_{\mathscr{V}} |\boldsymbol{v}_{\tau,t}|^2 \, d\mu_{\tau,t} = \int_{\mathscr{V}} |\overline{\boldsymbol{V}}_{\tau,t}(\boldsymbol{t}^{\sigma,k}(x))|^2 \, d\mu_{\tau,t}(x) = \int_{\mathscr{V}} |\overline{\boldsymbol{V}}_{\tau,t}|^2 \, d\mu^k \qquad t \in \big[(k-1)\tau, k\tau\big),$$

and this immediately gives (4.64) after an integration in time. □

Theorem 4.12 (Convergence of forward iterated transport maps) *Let $\mu_{\tau,t}$, $\boldsymbol{v}_{\tau,t}$, $\boldsymbol{T}_{\tau,t}$ be defined as in* (4.59), (4.61) *and* (4.58) *respectively starting from $\bar\mu = \bar\beta\mathscr{L}^d \in \mathrm{Dom}\,(\phi)$. Assume that ψ and $\bar\beta$ satisfy all the assumptions of Theorem 4.7, including the upper gradient inequality* (4.26)*, and that the tangent velocity field $\boldsymbol{v}_t$ relative to the gradient flow μ_t of ϕ with initial condition $\bar\mu$ satisfies* (3.11) *(or at least one of the conditions illustrated in Remarks* 3.6, 3.7, and 3.8*).*

Then

$$\lim_{\tau\downarrow 0} \int_{\mathscr{V}} \max_{t\in[0,T]} |\boldsymbol{T}_{\tau,t}(x) - \boldsymbol{X}(t,x)|^2 \, d\bar\mu(x) = 0 \qquad \forall T > 0, \tag{4.66}$$

where $\boldsymbol{X}$ is the flow associated to $\boldsymbol{v}_t$.

Proof. We apply Theorem 3.5. Notice that the convergence of $\mu_{\tau,t}$ to μ_t as $\tau \downarrow 0$ is ensured by Theorem 4.7 (that gives the convergence of $\overline{M}_{\tau,t}$) and the $C^{0,1/2}$ estimate (4.62). Moreover, (4.64) and (4.29) give

$$\limsup_{\tau\downarrow 0} \int_0^T \int_{\mathscr{V}} |\boldsymbol{v}_{\tau,t}|^2 \, d\mu_{\tau,t} \, dt = \limsup_{\tau\downarrow 0} \int_0^T \int_{\mathscr{V}} |\overline{\boldsymbol{V}}_{\tau,t}|^2 \, d\overline{M}_{\tau,t} \, dt = \int_0^T |\boldsymbol{v}_t|^2 \, d\mu_t \, dt$$

for any $T > 0$. Therefore, taking also into account that $\boldsymbol{T}_{\tau,t}(x)$ are flows relative to $\boldsymbol{v}_{t,\tau}$, all the assumptions of Theorem 3.5 are fulfilled and (4.66) is the conclusion of that theorem. □

The assumption (3.11) (and its variants considered in Remark 3.6, Remark 3.7 and Remark 3.8) may not be satisfied if the initial datum $\bar\mu$ is not sufficiently smooth. For this reason it is also interesting to consider the behaviour of the inverses of $\boldsymbol{T}_{\tau,t}$, reversing also the time variable. In the following theorem we focus on the case of a Sobolev regularity of $\boldsymbol{v}$ with respect to the space variable, leaving all other variants (for instance the BV ones) to the interested reader.

Recalling (3.8), given $T > 0$ we define the *backward* flow $\tilde{\boldsymbol{X}}(t,x)$ associated to $\boldsymbol{v}_t$ as $\tilde{\boldsymbol{X}}(t,x) := \boldsymbol{X}(t,T,x)$. Under the assumption

$$\boldsymbol{v} \in L^1_{\rm loc}\left((0,T];W^{1,1}_{\rm loc}(\mathscr{V})\right), \quad [\nabla_x \cdot \boldsymbol{v}]^+ \in L^1_{\rm loc}\left((0,T];L^\infty(\mathscr{V})\right), \tag{4.67}$$

considered up to a time reversal in Remark 3.7, the backward flow is well defined up to $t=0$ and produces for $t>0$ densities $\mu_t = \tilde{\boldsymbol{X}}(t,\cdot)_\#\bar\mu$ in $L^\infty(\mathscr{V})$ for any $\bar\mu \in L^\infty(\mathscr{V})$.

Analogously, we define

$$\tilde{\boldsymbol{T}}_{\tau,t} := \boldsymbol{T}_{\tau,t} \circ \boldsymbol{T}^{-1}_{\tau,T}, \tag{4.68}$$

mapping $\mu_{\tau,T}$ to $\mu_{\tau,t}$. Finally, as in Theorem 3.5, we have to take into account a correction term due to the optimal map $\boldsymbol{d}^\tau$ between μ_T and $\mu_{\tau,T}$.

Theorem 4.13 (Convergence of backward iterated transport maps) *Let $\tilde{\boldsymbol{T}}_{\tau,t}$ be defined as in (4.68) starting from $\bar\mu \in \mathrm{Dom}\,(\phi)$. Assume that ϕ satisfies all the assumptions of Theorem 4.7 and that the tangent velocity field $\boldsymbol{v}_t$ relative to the gradient flow μ_t of ϕ with initial condition $\bar\mu$ satisfies (4.67). Then*

$$\lim_{\tau\downarrow 0} \int_{\mathscr{V}} \max_{t\in[0,T]} \left|\tilde{\boldsymbol{T}}_{\tau,t}(\boldsymbol{d}^\tau) - \tilde{\boldsymbol{X}}(t,x)\right|^2 d\mu_T = 0.$$

Proof. We have just to notice that $\tilde{\boldsymbol{T}}_{\tau,t}$ is the backward flow associated to the velocity field $\boldsymbol{v}_{\tau,t}$ defined in (4.61) and then, using the same estimates used in the proof of Theorem 4.12, apply Theorem 3.5 and Remark 3.7 in a time reversed situation. □

Now we conclude the discussion relative to the basic example 4.4 we are interested to, namely the *internal energy* functional. See for instance [33], [41], [6] for more general examples.

The case of the internal energy functional. Let $\psi : [0,+\infty) \to (-\infty,+\infty]$ and $\bar\mu = \bar\beta\mathscr{L}^d$ be satisfying assumptions (ψ 1), (ψ 2), (ψ 3), and (ψ 4). Theorem 4.7 applies, yielding a unique gradient flow $\mu_t = \beta_t\mathscr{L}^d$ relative to ϕ which satisfies the nonlinear PDE

$$\frac{\partial}{\partial t}\beta_t = D_x \cdot (\nabla L_\psi(\beta_t)) \qquad \text{in } \mathscr{D}'\left((0,T)\times\mathscr{V}\right) \tag{4.69}$$

with homogeneous Neumann boundary conditions. Since $L'_\psi(s) = s\psi''(s)$, under suitable regularity assumptions on β_t the chain rule gives the alternative formulations

$$\frac{d}{dt}\beta_t = D_x \cdot \left(\beta_t\psi''(\beta_t)\nabla\beta_t\right) = D_x \cdot \left(\beta_t\nabla\psi'(\beta_t)\right). \tag{4.70}$$

If we want to apply to this example the results of Theorem 4.12, we should check if the tangent velocity field $\boldsymbol{v}$ given by

$$\boldsymbol{v}_t = -\frac{\nabla L_\psi(\beta_t)}{\beta_t} = -\nabla\psi'(\beta_t) \qquad \text{for } t>0, \tag{4.71}$$

satisfies (3.11) or one of the conditions discussed in Remarks 3.6, 3.7, 3.8, which are strictly related to the regularity of the solution β_t.

Besides (3.11), in the following discussion we focus our attention on Remark 3.8 (and its variants 3.10, 3.11). We distinguish some cases:

[(a) $\mathscr{V}$ is bounded and of class $C^{2,\alpha}$ and $\bar\beta \in C^\alpha$ is bounded away from 0.] When

$$\begin{cases} \mathscr{V} \text{ is bounded and of class } C^{2,\alpha} \text{ for some } \alpha > 0, \text{ and} \\ \bar\beta \in C^\alpha(\mathscr{V}), \quad 0 < \beta_{\min} \le \bar\beta(x) \le \beta_{\max} \quad \text{for } \mathscr{L}^d\text{-a.e. } x \in \mathscr{V}, \end{cases} \tag{4.72}$$

the maximum principle shows that the solution β_t satisfies the same bound

$$0 < \beta_{\min} \le \beta_t(x) \le \beta_{\max} \quad \forall\, x \in \mathscr{V},\, t \in [0,T]; \tag{4.73}$$

by [29, Thm. 10.1, Chap. III;Thm. 7.1, Chap. V] the variational solution of (4.69) is Hölder continuous in $[0,T] \times \overline{\mathscr{V}}$.

The smooth transformation $\rho := L_\psi(\beta)$ shows that ρ is a solution of the linear parabolic equation

$$\frac{\partial}{\partial t}\rho - a(t,x)\Delta\rho = 0 \quad \text{in } (0,T) \times \mathscr{V}$$

with homogeneous Neumann boundary conditions, where $a(t,x) = L'_\psi(\beta(t,x))$ is Hölder continuous and satisfies $0 < a_{\min} \le a \le a_{\max} < +\infty$ in $(0,T] \times \overline{\mathscr{V}}$.

Standard parabolic regularity theory [29, Chap. IV] yields $D^2_x\rho \in C^{\bar\alpha}((0,T) \times \overline{\mathscr{V}})$ for some $\bar\alpha > 0$; moreover, since $\rho_0 \in C^\alpha(\mathscr{V})$, too, then the intermediate Schauder estimates of [32, Thm 6.1] yield for every $x_0 \in \mathscr{V}$ the existence of $\varepsilon, \delta > 0$ such that

$$\sup_{(t,x)\in(0,T)\times B_\varepsilon(x_0)} t^{1-\delta}\|D^2_x\rho_t(x)\| < +\infty, \tag{4.74}$$

so that $\boldsymbol{v}$ satisfies (3.11).

[(b) The Heat/Porous medium equation in $\mathscr{V} = \mathbb{R}^d$.]

Let us consider the Heat/Porous medium equation in $\mathbb{R}^d$, corresponding to the choice

$$L_\psi(\beta) = \beta^m, \qquad \psi(s) := \begin{cases} \frac{1}{m-1}s^m & \text{if } m > 1, \\ s\log s & \text{if } m = 1. \end{cases} \tag{4.75}$$

Following the approach of Remarks 3.8 and 3.11, we assume here that an open set $P_0 \subset \mathbb{R}^d$ exists such that

$$\bar\beta \in C^\alpha(P_0), \quad \bar\beta > 0 \quad \text{in } P_0, \quad \bar\beta \equiv 0 \quad \text{in } \mathbb{R}^d \setminus P_0. \tag{4.76}$$

We know (see [16, 18, 42] and also [23, Chap. 5, Thm. 3.1, 3.3]) that the solution β is Hölder continuous in $\big((0,T]\times\mathbb{R}^d\big)\cup\big(\{0\}\times P_0\big)$; still applying the local regularity theory we mentioned in point **(a)**, we obtain that $D^2_x\beta$, $D_x\psi'(\beta) = D_x\boldsymbol{v}$ are Hölder continuous in P, too, thus showing that (3.35) is satisfied.

The Hölder assumption on $\bar\beta$ shows that β is locally Hölder continuous in P up to the initial time $t = 0$, too. Arguing as before, we obtain (4.74) for every $x_0 \in P_0$, which entails (3.36).

In order to check (3.44) we will invoke the ARONSON-BENILAN estimate [7], [23, Chap. 5, Lemma 2.1]

$$\Delta\psi'(\beta_t) \geq -\frac{k}{t}, \quad \text{with} \quad k = \frac{1}{m-1+(2/d)}, \quad \text{i.e.} \quad D_x \cdot \boldsymbol{v}_t(x) \leq \frac{k}{t}, \quad \forall\,(t,x) \in P. \tag{4.77}$$

[**(c)** $\mathscr{V}$ **is** $C^{2,\alpha}$**, possibly unbounded, and** $\Delta L_\psi(\bar\beta)$ **is a finite measure.**] The local regularity results we used in the previous points **(a), (b)** depend, in fact, only on the local behavior of L'_ψ around 0; let us thus assume that, in a suitable neighborhood $(0,\varepsilon_0)$, the function L'_ψ has a "power like" behaviour

$$c_0\beta^{\kappa_0} \leq L'_\psi(\beta) \leq c_1\beta^{\kappa_1} \quad \forall\,\beta \in (0,\varepsilon_0), \tag{4.78}$$

for given positive constants $0 < c_0 \leq c_1$, and $0 < \kappa_1 \leq \kappa_0$. Under this assumption and (4.76), the regularity results of [37] (see also [19, Page 76]) yield the Hölder continuity of β in $\Big((0,T) \times \overline{\mathscr{V}}\Big) \cup \Big(\{0\} \times P_0\Big)$. Arguing as in the previous points we still get $D_x^2\beta \in C^0(P)$, $D_x\psi'(\beta) = D_x\boldsymbol{v} \in C^0(P)$ and (4.74).

The only difference here is that we cannot invoke the regularizing effect (4.77), which seems to depend on the particular form (4.75) of L_ψ.

In this case, we should impose extra regularity properties on $\bar\beta$ which guarantee (3.42) (and therefore (3.37)): one possibility is to assume, in addition to (4.76), that

$$\Delta L_\psi(\bar\beta) \text{ is a finite measure.} \tag{4.79}$$

(4.79) and standard results for contraction semigroups in L^1 yield

$$\sup_{t>0} \int_{\{x:\beta_t(x)>0\}} |\partial_t\beta(t,x)|\,dx < +\infty; \tag{4.80}$$

for, the BREZIS-STRAUSS resolvent estimates [11] and CRANDALL-LIGGETT [17] generation Theorem show that the nonlinear operator $u \mapsto -\Delta L_\psi(u)$ with domain $D := \big\{u \in L^1(\mathscr{V}) : -\Delta L_\psi(u) \in L^1(\mathscr{V})\big\}$ generates a contraction semigroup in L^1, whose trajectories satisfy (4.80) if (4.79) holds.

Since L_ψ is an increasing function, we have

$$\boldsymbol{v}_t \cdot \nabla\beta_t = -\frac{\nabla L_\psi(\beta_t)}{\beta_t} \cdot \nabla\beta_t \leq 0 \quad \text{in } P;$$

taking into account (4.80), in order to prove (3.42) we should show that

$$\iint_P \frac{\nabla L_\psi(\beta_t)}{\beta_t} \cdot \nabla\beta_t\,dx\,dt < +\infty. \tag{4.81}$$

We argue by approximation as in the proof of Proposition 4.8 and we consider the same kind of regularized functions ℓ_ε obtained from the entropy $\ell(\beta) := \beta\log\beta$ according to (4.42), which satisfy the convexity condition (4.54) and

$$\ell_\varepsilon(\beta) - \varepsilon \leq \beta\log\beta \leq \ell_\varepsilon(\beta) \quad \forall\,\beta > 0. \tag{4.82}$$

(4.13) yields $L_\varepsilon(\beta) := L_\varepsilon := \varepsilon\ell_\varepsilon'(\beta) - \ell_\varepsilon(\beta) \quad \in L^1(0,T;W^{1,1}(\mathscr{V}))$ with

$$\int_0^T \int_{\mathscr{V}} \frac{|\nabla L_\varepsilon(\beta)|^2}{\beta^2}\beta\,dx\,dt < +\infty \quad \forall\,\varepsilon > 0.$$

It follows from the geodesic convexity of the functional $H_\varepsilon(\beta) := \int_{\mathscr{V}} \ell_\varepsilon(\beta)\,dx$ and the "Wasserstein chain rule" (see §10.1.2 in [6]) that the map $t \mapsto H_\varepsilon(\beta_t)$ is absolutely continuous and its time derivative is

$$\frac{d}{dt}H_\varepsilon(\beta_t) = \int_{\mathscr{V}} \boldsymbol{v}_t \cdot \nabla L_\varepsilon(\beta)\,dx = -\int_{\mathscr{V}} \frac{\nabla L_\psi(\beta_t)}{\beta_t}\frac{DL_\varepsilon(\beta_t)}{\beta_t}\beta_t\,dx \quad \text{for } \mathscr{L}^1\text{-a.e. } t > 0.$$

Upon an integration in time, (4.82) and $\varepsilon < e^{-1}$ yield

$$\int_{\mathscr{V}} \beta_T \log\beta_T\,dx + \iint_{P\cap\{\beta>\varepsilon\}} \frac{\nabla L_\psi(\beta_t)}{\beta_t}\nabla\beta_t\,dx\,dt \le \int_{\mathscr{V}\cap\{\bar\beta>1-e\varepsilon\}} \bar\beta\big([\log\bar\beta]^+ + \frac{\varepsilon}{1-e\varepsilon}\big)\,dx. \tag{4.83}$$

Since $\bar\beta$ is bounded, the right hand side of (4.83) is uniformly bounded as $\varepsilon \downarrow 0$; moreover, the finiteness of the second moment $\int_{\mathscr{V}} |x|^2\beta_T(x)\,dx < +\infty$ and Hölder inequality yield that the entropy $\int_{\mathscr{V}} \beta_T \log\beta_T$ cannot take the value $-\infty$ (see for instance Remark 9.3.7 in [6]); hence passing to the limit as $\varepsilon \downarrow 0$ we obtain (4.81).

We collect the above discussion and Theorem 4.12 in the following result

Corollary 4.14 (Convergence of iterated transport maps for diffusion equations) *Let $\mathscr{V} \subset \mathbb{R}^d$ be an open set of class $C^{2,\alpha}$, let $\bar\mu = \bar\beta\mathscr{L}^d \in \mathscr{P}_2(\mathscr{V})$ and $\psi : [0,+\infty) \to (-\infty,+\infty]$ be satisfying $(\psi\,1)$, $(\psi\,2)$, $(\psi\,3)$, $(\psi\,4)$.*
We assume that there exists an open set $P_0 \subset \mathscr{V}$ such that (4.76) *holds and that* at least one *of the following conditions is satisfied:*

$$P_0 = \mathscr{V} \text{ bounded of class } C^{2,\alpha} \quad \text{and} \quad 0 < \beta_{\min} \le \bar\beta(y) \quad \forall\,y \in \mathscr{V}, \tag{4.84a}$$

$$\mathscr{V} = \mathbb{R}^d \text{ and } L_\psi(\beta) = \beta^m \text{ for some } m \ge 1, \tag{4.84b}$$

$$\partial\mathscr{V} \in C^{2,\alpha}, \text{ (4.78) holds, and } \Delta L_\psi(\bar\beta) \text{ is a finite measure in } \mathscr{V}. \tag{4.84c}$$

If $\boldsymbol{v}_t$ is the velocity vector field associated through (4.71) *to the solution β_t of the nonlinear diffusion PDE* (4.69) *with initial condition $\bar\beta$, then for every final time $T > 0$ there exist an open set $\mathscr{V}_0$ with $\bar\beta\mathscr{L}^d(\mathscr{V} \setminus \mathscr{V}_0) = 0$ and a unique $\overline{\mathscr{V}}$-valued forward flow $\boldsymbol{X}$ which solves* (3.8) *for every $x \in \mathscr{V}_0$.*
$\boldsymbol{X} \in C^1([0,T] \times \mathscr{V}_0; \overline{\mathscr{V}})$, $\boldsymbol{X}(t,\cdot)\bar\beta\mathscr{L}^d = \beta_t\mathscr{L}^d$ and the iterated transport maps $\boldsymbol{T}_\tau$ constructed as in (4.57), (4.58) *from the solution of the variational algorithm* (4.21) *converge to $\boldsymbol{X}$:*

$$\lim_{\tau\downarrow 0} \int_{\mathscr{V}} \max_{[0,T]} |\boldsymbol{T}_{\tau,\cdot}(x) - \boldsymbol{X}(\cdot,x)|^2\bar\beta(x)\,dx = 0 \qquad \forall T > 0. \tag{4.85}$$

In the cases (4.84a) *and* (4.84b) *we can always choose $\mathscr{V}_0 \equiv P_0$.*

Proof. We can apply Theorem 4.12 in a time interval $(0, T')$ with $T' > T$: conditions (ψ 1), (ψ 2), (ψ 3), (ψ 4) together with Proposition 4.8 ensure that we are in the "Wasserstein gradient flow" setting; each of the assumptions (4.84a,b,c) provides enough regularity on the limit vector field

$$\boldsymbol{v}_t = -\frac{\nabla L_\psi(\beta_t)}{\beta_t}$$

in order to check (3.11) (in the case (4.84a)), or Remark 3.11 (in the case (4.84b)), or Remark 3.10 (in the case (4.84c))

Concerning the regularity of $\boldsymbol{X}$, it follows by classical results on differential equations in the first case and by (3.44) and Remark 3.11 in the second one (together with the identification $P_0 = \mathscr{V}_0$). In the third case, we can still apply (3.43) of Remark 3.11, by choosing $\varepsilon < T' - T$. □

5 An application to gradient flows of polyconvex functionals

Basic notation for vector calculus and first variations

In this section we will deal with vector valued maps $\boldsymbol{u} = (\boldsymbol{u}^i)_{i=1}^d : \mathscr{U} \to \mathscr{V} \subset \mathbb{R}^d$ defined in an open subset $\mathscr{U}$ of $\mathbb{R}^d$ with values in an open set $\mathscr{V}$. In order to distinguish between the systems of coordinates in $\mathscr{U}$ and in $\mathscr{V}$, we will use the greek letters $\alpha, \beta, \ldots$ for variables in $\mathscr{U}$ and latin letters $i, j, \ldots$ for components in $\mathscr{V}$.

We denote by A^i_α the elements of the matrix representing a linear map A in $\mathbb{L}(\mathbb{R}^d; \mathbb{R}^d)$; $(\mathsf{A}^T)^\alpha_i = \mathsf{A}^i_\alpha$ is the usual transposed matrix, $\mathsf{A} \cdot \mathsf{B} = \mathrm{tr}(\mathsf{A}^T\mathsf{B}) = \sum_{\alpha,i} \mathsf{A}^i_\alpha \mathsf{B}^i_\alpha$ denotes the scalar product. The *cofactor* matrix of A is denoted by $\mathrm{cof}\,\mathsf{A} = (\mathrm{cof}\,\mathsf{A})^i_\alpha$ and it satisfies the Laplace identities

$$(\mathrm{cof}\,\mathsf{A})^T \mathsf{A} = (\det \mathsf{A})\mathsf{Id}, \quad \text{i.e.} \quad \sum_i (\mathrm{cof}\,\mathsf{A})^i_\alpha \mathsf{A}^i_\beta = \delta_{\alpha,\beta} \det \mathsf{A}. \tag{5.1}$$

When A is invertible, $(\mathrm{cof}\,\mathsf{A})^T = (\det \mathsf{A})\mathsf{A}^{-1}$.

For a given sufficiently regular vector map $\boldsymbol{u}$, we denote by $D\boldsymbol{u}$ the matrix

$$D\boldsymbol{u} = (D\boldsymbol{u})^i_\alpha, \quad \text{with} \quad (D\boldsymbol{u})^i_\alpha := \frac{\partial \boldsymbol{u}^i}{\partial x_\alpha}; \tag{5.2}$$

the *divergence* of a matrix valued map $\mathsf{A} : \mathscr{U} \to \mathbb{L}(\mathbb{R}^d; \mathbb{R}^d)$ is defined as

$$(\mathrm{div}\,\mathsf{A})^i := \sum_\alpha \frac{\partial}{\partial x_\alpha} A^i_\alpha. \tag{5.3}$$

We recall that any map $\boldsymbol{u} \in W^{1,d-1}_{\mathrm{loc}}$ satisfies

$$\mathrm{div}\,\big(\mathrm{cof}\,D\boldsymbol{u}\big) = 0 \quad \text{in } \mathscr{U}. \tag{5.4}$$

Let F be a real map defined on (an open subset of) $\mathbb{L}(\mathbb{R}^d;\mathbb{R}^d)$; the differential of F and its action on a matrix B can be represented as

$$DF(\mathsf{A})_i^\alpha = \frac{\partial F}{\partial A_\alpha^i}(\mathsf{A}), \quad \text{so that} \quad DF(\mathsf{A})\mathsf{B} = DF(\mathsf{A})^T \cdot \mathsf{B} = \sum_{\alpha,i} \frac{\partial F}{\partial A_\alpha^i}(\mathsf{A})\mathsf{B}_\alpha^i. \tag{5.5}$$

In particular,

$$\text{if} \quad F(\mathsf{A}) := \Phi(\det \mathsf{A}), \quad \text{then} \quad DF(\mathsf{A}) = \Phi'(\det \mathsf{A})(\operatorname{cof} \mathsf{A})^T. \tag{5.6}$$

Let us now consider the functional

$$I(\boldsymbol{u}) := \int_{\mathscr{U}} F(D\boldsymbol{u})\,dx = \int_{\mathscr{U}} \Phi(\det D\boldsymbol{u})\,dx. \tag{5.7}$$

If $\boldsymbol{u}$ is sufficiently regular, the first variation $\delta I(\boldsymbol{u};\boldsymbol{\xi})$ along a smooth vector field $\boldsymbol{\xi} : \mathscr{U} \to \mathbb{R}^d$ with $\operatorname{cof} D\boldsymbol{u}\boldsymbol{n}_{\mathscr{U}}$ orthogonal to $\boldsymbol{\xi}$ on $\partial\mathscr{U}$ (here $\boldsymbol{n}_{\mathscr{U}}$ denotes the exterior unit normal to $\partial\mathscr{U}$) yields

$$\begin{aligned} \delta I(\boldsymbol{u};\boldsymbol{\xi}) &= \frac{d}{ds} I(\boldsymbol{u}+s\boldsymbol{\xi})\Big|_{s=0} = \int_{\mathscr{U}} \Phi'(\det D\boldsymbol{u})(\operatorname{cof} D\boldsymbol{u}) \cdot D\boldsymbol{\xi}\,dx \\ &= -\int_{\mathscr{U}} \operatorname{div}\Big(\Phi'(\det D\boldsymbol{u})(\operatorname{cof} D\boldsymbol{u})\Big) \cdot \boldsymbol{\xi}\,dx. \end{aligned} \tag{5.8}$$

Other kind of variations will play a crucial role in the following: here we are considering the *variation of the deformed state* [25, Chap. 2, 1.5] induced by a given vector field $\boldsymbol{\eta} \in C^1(\overline{\mathscr{V}};\overline{\mathscr{V}})$ with $\boldsymbol{\eta}\cdot\boldsymbol{n}_{\mathscr{V}} = 0$ and $\boldsymbol{u}(\mathscr{U}) \subset \mathscr{V}$, and the induced flow in $\mathscr{V}$

$$\begin{cases} \dfrac{d}{ds}\boldsymbol{Y}(s,y) = \boldsymbol{\eta}\left(\boldsymbol{Y}(s,y)\right) \\ \boldsymbol{Y}(0,y) = y \end{cases} \qquad \text{for } y \in \mathscr{V}.$$

Thus we can evaluate

$$\bar{\delta} I(\boldsymbol{u};\boldsymbol{\eta}) := \frac{d}{ds} I(\boldsymbol{Y}(s,\boldsymbol{u}))\Big|_{s=0} = \int_{\mathscr{U}} \Phi'(\det D\boldsymbol{u}) \det D\boldsymbol{u}\ \operatorname{tr}(D_y\boldsymbol{\eta}(\boldsymbol{u}))\,dx \tag{5.9}$$

which, in the case $\boldsymbol{u}$ is a diffeomorphism between $\overline{\mathscr{U}}$ and $\overline{\mathscr{V}}$, corresponds to the usual Euler-Lagrange first variation (5.8) with $\boldsymbol{\xi} := \boldsymbol{\eta}\circ\boldsymbol{u}$; for, the fact that $\boldsymbol{u}$ is a C^1 diffeomorphism between $\mathscr{U}$ and $\mathscr{V}$ gives that the normal to $\mathscr{V}$ at $\boldsymbol{u}(x)$, with $x \in \partial\mathscr{U}$, is parallel to $\operatorname{cof} D\boldsymbol{u}(x)\boldsymbol{n}_{\mathscr{U}}(x)$, so that $\boldsymbol{\xi} = \boldsymbol{\eta}(\boldsymbol{u})$ is orthogonal to $\operatorname{cof} D\boldsymbol{u}(x)\boldsymbol{n}_{\mathscr{U}}(x)$ on $\partial\mathscr{U}$. Then, (5.1) yields

$$\begin{aligned} \delta I(\boldsymbol{u};\boldsymbol{\eta}\circ\boldsymbol{u}) &= \int_{\mathscr{U}} \Phi'(\det D\boldsymbol{u})(\operatorname{cof} D\boldsymbol{u}) \cdot (D_y\boldsymbol{\eta}(\boldsymbol{u})D\boldsymbol{u})\,dx \\ &= \int_{\mathscr{U}} \Phi'(\det D\boldsymbol{u}) \det D\boldsymbol{u}((D\boldsymbol{u})^{-1})^T \cdot (D_y\boldsymbol{\eta}(\boldsymbol{u})Du)\,dx \\ &= \int_{\mathscr{U}} \Phi'(\det D\boldsymbol{u}) \det D\boldsymbol{u} \operatorname{tr}(D_y^T\boldsymbol{\eta}(\boldsymbol{u}))\,dx = \bar{\delta} I(\boldsymbol{u};\boldsymbol{\eta}), \end{aligned} \tag{5.10}$$

because $((D\boldsymbol{u})^{-1})^T \cdot (D_y\eta(\boldsymbol{u})D\boldsymbol{u}) = ((D\boldsymbol{u})^T)^{-1} \cdot (D_y\eta(\boldsymbol{u})D\boldsymbol{u}) = \operatorname{trace}\big(((D\boldsymbol{u})^T)^{-1}(D\boldsymbol{u})^T D_y^T\eta(\boldsymbol{u})\big) = \operatorname{trace}(D_y^T\eta(\boldsymbol{u})$.

Gradient flow in L^2

In [22], the authors build a smooth solution $\boldsymbol{u}$ of the nonlinear parabolic PDE (here we adopt the same notation of that paper)

$$\begin{cases} \dfrac{\partial}{\partial t}\boldsymbol{u} = \operatorname{div}\,\big(DF(D\boldsymbol{u})^T\big) = \operatorname{div}\Big(\Phi'(\det D\boldsymbol{u})(\operatorname{cof} D\boldsymbol{u})\Big), \\ \boldsymbol{u}(0,\cdot) = \bar{\boldsymbol{u}}, \end{cases} \tag{5.11}$$

in $(0,T)\times\mathscr{U}$, corresponding to the gradient flow with respect to the L^2 metric of the functional I, since (5.11) has, at least formally, a natural variational formulation as

$$\frac{d}{dt}\int_{\mathscr{U}} \boldsymbol{u}\cdot\boldsymbol{\xi}\,dx = -\delta I(\boldsymbol{u};\boldsymbol{\xi}) \quad \text{for every } \boldsymbol{\xi}\in C^1(\overline{\mathscr{U}};\mathbb{R}^d) \text{ with } (\operatorname{cof} D\boldsymbol{u})\boldsymbol{n}_{\mathscr{U}} \perp \boldsymbol{\xi} \text{ on } \partial\mathscr{U}. \tag{5.12}$$

Here $\mathscr{U},\mathscr{V}\subset\mathbb{R}^d$ are bounded open sets with a smooth boundary, $\bar{\boldsymbol{u}}:\overline{\mathscr{U}}\to\overline{\mathscr{V}}$ belongs to $\mathrm{Diff}\,(\overline{\mathscr{U}};\overline{\mathscr{V}})$, the class of C^1 diffeomorphisms mapping $\partial\mathscr{U}$ onto $\partial\mathscr{V}$ with strictly positive determinant, $\Phi:(0,+\infty)\to\mathbb{R}$ is a smooth convex function with $\Phi''>0$, and the solution $\boldsymbol{u}$ of (5.11) is built in such a way that the same properties are satisfied by $\boldsymbol{u}(t,\cdot)$ for any $t\ge 0$. Precisely, they show that the scalar quantity

$$\beta(t,y) := \det D\boldsymbol{w}(t,y) = \frac{1}{\det D\boldsymbol{u}(t,\boldsymbol{w}(t,y))}, \quad y\in\mathscr{V}, \quad \text{with} \quad \boldsymbol{w}(t,\cdot) := [\boldsymbol{u}(t,\cdot)]^{-1}, \tag{5.13}$$

which is the Lebesgue density of the measure $\mu_t := (\boldsymbol{u}_t)_\#\mathscr{L}^d\llcorner\mathscr{U}$, can be built solving the nonlinear boundary value problem of diffusion type

$$\begin{cases} \dfrac{\partial}{\partial t}\beta = \operatorname{div}\Big(\Phi''\Big(\dfrac{1}{\beta}\Big)\dfrac{D\beta}{\beta^2}\Big) = \Delta\Big(-\Phi'\Big(\dfrac{1}{\beta}\Big)\Big) & \text{in } (0,+\infty)\times\mathscr{V}, \\ D\beta(t,\cdot)\boldsymbol{n}_{\mathscr{V}}(\cdot) = 0 & \text{on } (0,T)\times\partial\mathscr{V}, \\ \beta(0,\cdot) = \bar\beta := \dfrac{1}{\det D\bar{\boldsymbol{u}}}\circ\bar{\boldsymbol{u}}^{-1} & \text{in } \mathscr{V}, \end{cases} \tag{5.14}$$

(notice that the map $s\mapsto -\Phi'(1/s)$ is monotone increasing, thus the problem is parabolic). This can be explained as follows: setting

$$\psi(s) := s\Phi\Big(\frac{1}{s}\Big), \quad s>0, \quad \psi(0) = \lim_{s\downarrow 0} s\Phi\Big(\frac{1}{s}\Big) = \lim_{r\uparrow+\infty}\frac{\Phi(r)}{r}, \tag{5.15}$$

a change of variables gives $I(\boldsymbol{u}) = \int_{\mathscr{V}}\psi(\beta[\boldsymbol{u}])\,dy$, and therefore we expect the gradient flow of I to be related to the gradient flow of

$$\phi(\beta) := \int_{\mathscr{V}}\psi(\beta)\,dy. \tag{5.16}$$

Using the identities

$$L_\psi(s) = s\psi'(s) - \psi(s) = -\Phi'\Big(\frac{1}{s}\Big), \quad \psi''(s) = s^{-3}\Phi''(\frac{1}{s}), \tag{5.17}$$

the PDE (5.14) can be written as

$$\frac{\partial}{\partial t}\beta = \operatorname{div}\big(\beta D\psi'(\beta)\big) = \Delta(L_\psi(\beta)) \qquad \text{in } (0,+\infty)\times\mathscr{V}. \tag{5.18}$$

As for the metric, again a change of variables (see [22] and also Lemma 5.5 below) shows that it becomes W_2 at the level of β, and Otto's calculus (see [35], [36] and Example 4.4) shows that (5.18) is indeed the gradient flow of ϕ with respect to W_2.

A direct derivation is also possible, starting from the variational formulation of (5.11): first of all we observe that changing variables inside the integral (5.9) yields

$$\bar\delta I(\boldsymbol{u};\boldsymbol{\eta}) = \int_{\mathscr{V}} \Phi'(\frac{1}{\beta})\operatorname{tr}(D\boldsymbol{\eta})\,dy = -\int_{\mathscr{V}} L_\psi(\beta)\operatorname{tr}(D\boldsymbol{\eta})\,dy. \tag{5.19}$$

Now we choose a test vector field $\boldsymbol{\eta}$ of the form

$$\boldsymbol{\eta} = (D\zeta)^T, \quad \text{with} \quad \zeta \in C^2_n(\overline{\mathscr{V}}) := \Big\{\zeta \in C^2(\overline{\mathscr{V}}) : D\zeta(y)\boldsymbol{n}_{\mathscr{V}}(y) = 0 \quad \text{on } \partial\mathscr{V}\Big\}, \tag{5.20}$$

and we observe that in this case $\boldsymbol{\xi} = (D\zeta)^T \circ \boldsymbol{u}$ and

$$\int_{\mathscr{U}} \partial_t\boldsymbol{u}\cdot\boldsymbol{\xi}\,dx = \int_{\mathscr{U}} \partial_t\boldsymbol{u}\cdot(D\zeta)^T\circ\boldsymbol{u}\,dx = \frac{d}{dt}\int_{\mathscr{U}}\zeta(\boldsymbol{u})\,dx = \frac{d}{dt}\int_{\mathscr{V}}\zeta\beta\,dy; \tag{5.21}$$

on the other hand (5.12) and (5.10) yield $\int_{\mathscr{U}} \partial_t\boldsymbol{u}\cdot\boldsymbol{\xi}\,dx = \bar\delta I(\boldsymbol{u};\boldsymbol{\eta})$, so that (5.19) provides

$$\frac{d}{dt}\int_{\mathscr{V}}\zeta\beta\,dy = \int_{\mathscr{V}} L_\psi(\beta)\Delta\zeta\,dy \qquad \forall\,\zeta\in C^2_n(\overline{\mathscr{V}}), \tag{5.22}$$

which is just the weak formulation of (5.14).

Having built β, the remarkable fact is that one is able to build $\boldsymbol{u}$ solving a *first order* ODE associated to the vectorfield

$$\boldsymbol{V}^i(t,y) := -\frac{\partial}{\partial y_i}\psi'(\beta(t,y)), \quad \text{so that} \quad \boldsymbol{V} := -\psi''(\beta)\big(D_y\beta\big)^T = -\frac{DL_\psi(\beta)^T}{\beta}, \tag{5.23}$$

and precisely solving

$$\begin{cases} \boldsymbol{Y}'(t,\bar y) = \boldsymbol{V}(t,\boldsymbol{Y}(t,\bar y)), \\ \boldsymbol{Y}(0,\bar y) = \bar y \end{cases} \tag{5.24}$$

and setting $\boldsymbol{u}_t(x) = \boldsymbol{Y}(t,\bar{\boldsymbol{u}}(x))$. This is still a consequence of (5.12) by choosing an arbitrary vector field of the form $\boldsymbol{\xi} := \boldsymbol{\eta}\circ\boldsymbol{u}$ and applying (5.10) and (5.19): we obtain

$$\begin{aligned}\int_{\mathscr{U}} \partial_t\boldsymbol{u}(x)\cdot\boldsymbol{\eta}(\boldsymbol{u}(x))\,dx &= -\int_{\mathscr{U}} \frac{(DL_\psi(\beta))^T}{\beta}\circ\boldsymbol{u}\cdot\eta\circ\boldsymbol{u}\,dx = -\int_{\mathscr{V}} DL_\psi(\beta)\cdot\boldsymbol{\eta}\,dy \\ &= \int_{\mathscr{V}} L_\psi(\beta)\operatorname{tr}(D\boldsymbol{\eta})\,dy = -\bar\delta I(\boldsymbol{u};\boldsymbol{\eta}) = -\delta I(\boldsymbol{u};\boldsymbol{\xi}).\end{aligned} \tag{5.25}$$

When $\mathscr{V}$ is bounded and of class $C^{2,\alpha}$ the main result of [22], obtained along the lines of the previous discussion (see also Example 4.4(a)), can be stated as follows.

Theorem 5.1 (Evans-Gangbo-Savin) *Let us assume that $\mathscr{V}$ is a bounded open set of class $C^{2,\alpha}$; if $\boldsymbol{u}_0 \in C^{1,\alpha}(\overline{\mathscr{U}};\overline{\mathscr{V}}) \cap \mathrm{Diff}\,(\overline{\mathscr{U}};\overline{\mathscr{V}})$ then there exists a unique solution $\boldsymbol{u}_t \in \mathrm{Diff}\,(\overline{\mathscr{U}};\overline{\mathscr{V}})$ with $\partial_t \boldsymbol{u} \in L^2((0,T)\times\mathscr{U};\mathbb{R}^d)$ of the (distributional formulation of) (5.11): $\boldsymbol{u}_t$ admits the representation $\boldsymbol{u}_t(x) := \boldsymbol{Y}(t,\bar{\boldsymbol{u}}(x))$, where $\boldsymbol{Y}$ is defined (5.24), (5.23) and (5.14).*

When $\mathscr{V}$ is unbounded or $\bar{\boldsymbol{u}}$ is not surjective or $\det D\bar{\boldsymbol{u}}$ is not bounded away from 0 and $+\infty$, we should consider a wider class of maps and a weaker notion of solution, which are closely related to the notion of weak diffeomorphisms of [25, Chap. 2]. In this case, the first variations of the functional (5.7) introduced in (5.8) and (5.9) are not equivalent and the right notion of solution should take into account both the approaches, as it should be clear by the previous calculations.

Definition 5.2 (Weak diffeomorphism) *We say that a Borel map $\boldsymbol{u} : \mathscr{U} \to \mathscr{V}$ is a* weak diffeomorphism *and we write $\widetilde{\mathrm{diff}}\,(\mathscr{U};\mathscr{V})$ if there exists a Borel set $\mathscr{U}_0 \subset \mathscr{U}$ such that*

$$\begin{gathered}\mathscr{L}^d(\mathscr{U}\setminus\mathscr{U}_0) = 0, \quad \boldsymbol{u} \textit{ is differentiable at all points of } \mathscr{U}_0,\\ \boldsymbol{u} : \mathscr{U}_0 \to \mathscr{V} \textit{ is injective and } \det D\boldsymbol{u}(x) > 0 \textit{ for all } x \in \mathscr{U}_0.\end{gathered} \tag{5.26}$$

We say that a weak diffeomorphism $\boldsymbol{u}$ belongs to $\widetilde{\mathrm{Diff}}\,(\mathscr{U};\mathscr{V})$ if there exists an open *set $\mathscr{U}_0 \subset \mathscr{U}$ satisfying (5.26) with $\boldsymbol{u}|_{\mathscr{U}_0} \in C^1(\mathscr{U}_0;\mathscr{V})$.*

Definition 5.3 (Weak solutions of the gradient flow) *We say that a family $\boldsymbol{u}_t \in \widetilde{\mathrm{diff}}\,(\mathscr{U};\mathscr{V})$ with $\partial_t\boldsymbol{u}_t \in L^2((0,T)\times\mathscr{U};\mathbb{R}^d)$ is a solution of (5.11) if $\Phi'(\det D\boldsymbol{u}_t)\det D\boldsymbol{u}_t \in L^1(\mathscr{U})$ and*

$$\int_{\mathscr{U}} \partial_t\boldsymbol{u}_t\cdot\boldsymbol{\eta}(\boldsymbol{u}_t)\,dx = -\int_{\mathscr{U}} \Phi'(\det D\boldsymbol{u}_t)\,\mathrm{tr}(D_y\boldsymbol{\eta}(\boldsymbol{u}_t))\det D\boldsymbol{u}_t\,dx \quad \mathscr{L}^1\textit{-a.e. in } (0,T), \tag{5.27}$$

for every vector field $\boldsymbol{\eta} \in C^1(\overline{\mathscr{V}};\overline{\mathscr{V}})$ with $\boldsymbol{\eta}\cdot\boldsymbol{n}_{\mathscr{V}} = 0$ on $\partial\mathscr{V}$.

Notice that the area formula [6, Lemma 5.5.3] gives that if $\boldsymbol{u} \in \widetilde{\mathrm{diff}}\,(\mathscr{U};\mathscr{V})$ then

$$\boldsymbol{u}_\#\big(\mathscr{L}^d \llcorner \mathscr{U}\big) = \beta[\boldsymbol{u}]\mathscr{L}^d\llcorner\mathscr{V} \quad \text{with} \quad \beta[\boldsymbol{u}](y) := \begin{cases} \dfrac{1}{\det D\boldsymbol{u}(x)} & \text{if } y = \boldsymbol{u}(x) \text{ for } x \in \mathscr{U}_0,\\ 0 & \text{if } y \in \mathscr{V}\setminus\boldsymbol{u}(\mathscr{U}_0).\end{cases} \tag{5.28}$$

Moreover, it is easy to check that

$$I(\boldsymbol{u}) = \phi(\beta[\boldsymbol{u}]) \quad \text{if either } \psi(0) = 0 \text{ or } \mathscr{L}^d\left(\{y\in\mathscr{V} : \beta[\boldsymbol{u}](y) = 0\}\right) = 0. \tag{5.29}$$

The next result is the natural generalization of Theorem 5.1 when $\mathscr{V}$ is unbounded or the initial datum $\bar{\boldsymbol{u}}$ is only in $\widetilde{\mathrm{Diff}}\,(\mathscr{U};\mathscr{V})$ (thus allowing for the degeneracy of $\beta[\boldsymbol{u}]$).

Theorem 5.4 (Existence and uniqueness of weak solutions) *Let us assume that*

$$\bar{\boldsymbol{u}} \in \widetilde{\mathrm{Diff}}\,(\mathscr{U};\mathscr{V}) \cap L^2(\mathscr{U};\mathbb{R}^d), \qquad \int_{\mathscr{U}_0} \Phi(\det D\bar{\boldsymbol{u}}(x))\,dx < +\infty, \quad \bar{\boldsymbol{u}}|_{\mathscr{U}_0} \in C^{1,\alpha}, \tag{5.30}$$

where $\mathscr{U}_0$ is an open set with $\mathscr{L}^d(\mathscr{U} \setminus \mathscr{U}_0) = 0$ and that at least one *of the following conditions is satisfied:*

$$\mathscr{V} \text{ is bounded of class } C^{2,\alpha}, \quad 0 < \beta_{\min} \le \bar{\beta} = \beta[\bar{\boldsymbol{u}}] \le \beta_{\max} < +\infty \quad \forall\, y \in \mathscr{V}, \tag{5.31a}$$

$$\mathscr{V} = \mathbb{R}^d, \quad \bar{\beta} = \beta[\bar{\boldsymbol{u}}] \in L^\infty(\mathscr{V}), \quad L_\psi(\beta) = \beta^m \text{ for some } m \ge 1, \tag{5.31b}$$

$$\partial\mathscr{V} \in C^{2,\alpha}, \quad (4.78) \text{ holds}, \quad \bar{\beta} = \beta[\bar{\boldsymbol{u}}] \in L^\infty(\mathscr{V}), \quad \Delta\bar{\beta} \text{ is a finite measure in } \mathscr{V}. \tag{5.31c}$$

Then there exists a unique solution $\boldsymbol{u}_t \in \widetilde{\mathrm{Diff}}\,(\mathscr{U};\mathscr{V})$ with $\partial_t \boldsymbol{u} \in L^2((0,T)\times\mathscr{U};\mathbb{R}^d)$ of (5.11) *according to Definition 5.3; it admits the representation $\boldsymbol{u}_t(x) := \boldsymbol{Y}(t,\bar{\boldsymbol{u}}(x))$, where $\boldsymbol{Y}$ is defined by* (5.24), (5.23) *and* (5.14).

Proof. We follow the same construction we already described in the "regular" case. Applying Corollary 4.14 we find an open set $\mathscr{V}_0 \subset \boldsymbol{u}(\mathscr{U}_0)$ with $\mathscr{L}^d(\mathscr{U}_0 \setminus \boldsymbol{u}^{-1}(\mathscr{V}_0)) = 0$ and a flow $\boldsymbol{Y} : [0,T]\times\mathscr{V}_0 \to \overline{\mathscr{V}}$ solving (5.24) for the velocity $\boldsymbol{V}$ given by (5.23).

It is not restrictive to assume $\mathscr{V}_0 = \boldsymbol{u}(\mathscr{U}_0)$; setting $\boldsymbol{u}(t,x) := \boldsymbol{Y}(t,\bar{\boldsymbol{u}}(x))$, $x \in \mathscr{U}_0$, we find $\beta(t,\cdot) = \beta[\boldsymbol{u}(t,\cdot)]$ and therefore

$$\begin{aligned}\int_0^T\int_{\mathscr{U}_0} |\partial_t\boldsymbol{u}|^2\,dx\,dt &= \int_0^T\int_{\mathscr{U}_0} |\boldsymbol{V}(t,\boldsymbol{u}(t,x))|^2\,dx\,dt\\ &= \int_0^T\int_{\mathscr{V}_0} |\boldsymbol{V}(t,y)|^2\,\beta(t,y)\,dy\,dt < +\infty,\end{aligned}$$

by the energy estimate (4.13) in the Wasserstein space of equation (5.14).

We then have for $\rho \in C^1_0(0,T)$ and $\boldsymbol{\eta} \in C^1(\overline{\mathscr{V}};\overline{\mathscr{V}})$ with $\boldsymbol{\eta}\cdot\boldsymbol{n}_{\mathscr{V}} = 0$,

$$\begin{aligned}\int_0^T\int_{\mathscr{U}_0} \rho(t)\partial_t\boldsymbol{u}\cdot\boldsymbol{\eta}(\boldsymbol{u})\,dx\,dt &= \int_0^T\int_{\mathscr{U}_0} \rho(t)\boldsymbol{V}(t,\boldsymbol{u}(t,x))\cdot\boldsymbol{\eta}(\boldsymbol{u}(t,x))\,dx\,dt\\ &= \int_0^T \rho(t)\int_{\mathscr{V}_0} \frac{-(DL_\psi(\beta))^T}{\beta}\cdot\boldsymbol{\eta}\beta\,dy\,dt\\ &= \int_0^T \rho(t)\int_{\mathscr{V}} L_\psi(\beta)\,\mathrm{tr}(D\boldsymbol{\eta})\,dy\,dt\\ &= \int_0^T \rho(t)\int_{\mathscr{U}} -\Phi'(\det D\boldsymbol{u})\,\mathrm{tr}(D\boldsymbol{\eta}(\boldsymbol{u}))\det D\boldsymbol{u}\,dx\,dt,\end{aligned}$$

which yields (5.27), being ρ arbitrary. The uniqueness follows easily: for any other solution $\boldsymbol{v}$, choosing $\boldsymbol{\eta} = (D\zeta)^T$ as in (5.20) we find that $\beta[\boldsymbol{v}]$ is the unique weak solution $\beta = \beta[\boldsymbol{u}]$ of (5.14). Arguing as above we obtain

$$\partial_t\boldsymbol{v} = \boldsymbol{V}(t,\boldsymbol{v}) \quad \mathscr{L}^d\text{-a.e. in } \mathscr{U}$$

for $\mathscr{L}^1$-a.e. $t \in (0,T)$. Possibly redefining $\boldsymbol{v}$ in a space-time negligible set we can assume that $\boldsymbol{v}(\cdot,x)$ is absolutely continuous, with derivative $\mathscr{L}^1$-a.e. equal in $(0,T)$ to $\partial_t\boldsymbol{v}(t,x)$, for $\mathscr{L}^d$-a.e. $x \in \mathscr{U}$. On the other hand, Fubini's theorem gives that

$$\partial_t\boldsymbol{v} = \boldsymbol{V}(t,\boldsymbol{v}) \quad \mathscr{L}^1\text{-a.e. in } (0,T)$$

for $\mathscr{L}^d$-a.e. $x \in \mathscr{U}$. As a consequence, $\boldsymbol{v}(\cdot, x)$ solves the ODE for $\mathscr{L}^d$-a.e. $x \in \mathscr{U}$. Invoking the uniqueness of the flow (5.24) for $\bar{y} \in \mathscr{V}_0$ of Corollary 4.14 we conclude that $\boldsymbol{v}(t,x) = \boldsymbol{u}(t,x)$.

□

Convergence of the variational approximation scheme

In [22] the authors raise the problem of the convergence of the variational formulation of the Euler implicit scheme to their solution. Let $\tau > 0$ be a time step and let $\boldsymbol{u}^k$ be recursively defined by minimizing in $\widetilde{\mathrm{diff}}\,(\mathscr{U};\mathscr{V})$ the functional

$$\boldsymbol{u} \mapsto \mathscr{F}_\tau(\boldsymbol{u};\boldsymbol{u}^{k-1}) := \frac{1}{2\tau}\int_{\mathscr{U}} |\boldsymbol{u} - \boldsymbol{u}^{k-1}|^2\,dx + I(\boldsymbol{u}) \tag{5.32}$$

with the initial condition $\boldsymbol{u}^0 = \bar{\boldsymbol{u}}$.

We denote by β^k the sequence of measure densities arising from the analogous recursive minimization of the functional

$$\beta \mapsto \mathscr{G}_\tau(\beta;\beta^{k-1}) = \frac{1}{2\tau}W_2^2(\beta,\beta^{k-1}) + \phi(\beta), \tag{5.33}$$

with ϕ as in (5.16) and the initial condition $\beta^0 = \bar{\beta} = \beta[\bar{\boldsymbol{u}}]$ (see the previous section and, in particular, Example 4.4).

Since the functional I (resp. ϕ) differs only by a constant if we perturb Φ (resp. ψ) by adding a linear term λs (resp. a constant term λ), we can always assume that

$$\psi(0) = \lim_{s\downarrow 0}\psi(s) = \lim_{r\to+\infty}\frac{\Phi(r)}{r} \quad \text{is either } 0 \text{ or } +\infty, \tag{5.34}$$

and we have

$$I(\boldsymbol{u}) = \phi(\beta[\boldsymbol{u}]) \quad \text{if either } \psi(0) = 0 \text{ or } \phi(\beta) < +\infty. \tag{5.35}$$

In this correspondence, the (usual) convexity of ϕ is equivalent to the *polyconvexity* of I, whereas the *geodesic convexity* (4.52), (4.54) of ϕ corresponds to the condition that the map $s \mapsto \Phi(s^d)$ is convex and nonincreasing in $(0,+\infty)$, and it is equivalent to the convexity of I along a special class of perturbations (closely related to the variation $\bar{\delta}$ of (5.9), see also (5.20), (5.21)), namely

$$t \mapsto I(\boldsymbol{\eta}_t \circ \boldsymbol{u}) \quad \text{is convex whenever } \boldsymbol{\eta}_t := (1-t)\boldsymbol{i} + t(D\zeta)^T, \text{ with } \zeta : \mathbb{R}^d \to \mathbb{R} \text{ convex.} \tag{5.36}$$

The following existence result for $\boldsymbol{u}^k$, together with some uniform estimates, has been proved in [22] in the "regular" case, when $\mathscr{V}$ is bounded and $0 < \beta_{\min} \le \bar{\beta} \le \beta_{\max} < +\infty$.

Lemma 5.5 (Discrete estimates) *Assume that*

(i) $\mathscr{U}$ is an open set in $\mathbb{R}^d$ whose finite Lebesgue measure is normalized to 1, $\mathscr{V}$ is an open (possibly unbounded) set of $\mathbb{R}^d$ with $C^{2,\alpha}$ boundary;

(ii) $s \mapsto s\Phi(1/s)$ *is smooth and convex in* $(0, +\infty)$*, and*

$$\lim_{s\downarrow 0} \Phi(s) = +\infty, \qquad \lim_{s\to+\infty} \frac{\Phi(s)}{s} \in \{0, +\infty\};$$

(iii) $\bar{\boldsymbol{u}} \in \widetilde{\mathrm{diff}}\,(\mathscr{U}; \mathscr{V}) \cap L^2(\mathscr{U}; \mathbb{R}^d)$*,* $\bar\beta = \beta[\bar{\boldsymbol{u}}]$*,* $\Phi(\bar{\boldsymbol{u}}) = \phi(\bar\beta) < +\infty$*.*

Then the variational problems (5.32) *and* (5.33) *can be iteratively solved in* $\widetilde{\mathrm{diff}}\,(\mathscr{U}; \mathscr{V})$ *and* $\mathscr{P}_2(\mathscr{V})$ *respectively, they admit a unique solution with*

$$\beta^k = \beta[\boldsymbol{u}^k], \quad \boldsymbol{u}^k = \boldsymbol{t}^{k-1} \circ \boldsymbol{u}^{k-1}, \quad \mathscr{F}_\tau(\boldsymbol{u}^k; \boldsymbol{u}^{k-1}) = \mathscr{G}_\tau(\beta^k; \beta^{k-1}) \qquad \forall k \in \mathbb{N}, \tag{5.37}$$

where $\boldsymbol{t}^{k-1}$ *is the optimal transport map between* β^{k-1} *and* β^k*.*
Finally, if $0 < \beta_{\min} \le \bar\beta(x) \le \beta_{\max} < +\infty$*, then all* β^k *are Lipschitz continuous in* $\overline{\mathscr{V}}$*,*

$$0 < \beta_{\min} \le \beta^k \le \beta_{\max} \qquad \forall k \in \mathbb{N} \tag{5.38}$$

and $\boldsymbol{u}^k \in \mathrm{Diff}\,(\overline{\mathscr{U}}; \overline{\mathscr{V}})$ *for all* $k \ge 1$*.*

Proof. First of all, we observe that the map $\boldsymbol{u} \mapsto \beta[\boldsymbol{u}]\mathscr{L}^d$ between $\widetilde{\mathrm{diff}}\,(\mathscr{U}; \mathscr{V})$ (with the L^2-norm) and $\mathscr{P}_2(\mathscr{V})$ is non expansive, i.e.

$$W_2(\beta[\boldsymbol{u}_1]\mathscr{L}^d, \beta[\boldsymbol{u}_2]\mathscr{L}^d) \le \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{L^2(\mathscr{U};\mathbb{R}^d)}; \tag{5.39}$$

for, we simply apply the very definition of Kantorovich-Wasserstein distance (2.5) with the plan $\gamma := (\boldsymbol{u}_1, \boldsymbol{u}_2)_\# \mathscr{L}^d \llcorner \mathscr{U}$. In particular, (5.39) and (5.35) show that

$$\mathscr{F}_\tau(\boldsymbol{u}, \boldsymbol{u}^{k-1}) \ge \mathscr{G}_\tau(\beta[\boldsymbol{u}]; \beta^{k-1}); \qquad \inf_{\boldsymbol{u} \in \widetilde{\mathrm{diff}}\,(\mathscr{U};\mathscr{V})} \mathscr{F}_\tau(\boldsymbol{u}, \boldsymbol{u}^{k-1}) \ge \inf_{\beta \in \mathscr{P}_2(\mathscr{V})} \mathscr{G}_\tau(\beta[\boldsymbol{u}]; \beta^{k-1}). \tag{5.40}$$

We also observe that for every $\boldsymbol{u}_1 \in \widetilde{\mathrm{diff}}\,(\mathscr{U}; \mathscr{V})$ with $\beta_1\mathscr{L}^d := \beta[\boldsymbol{u}_1]\mathscr{L}^d \in \mathscr{P}_2(\mathscr{V})$ and $\beta_2\mathscr{L}^d\mathscr{P}_2(\mathscr{V})$ there exists a unique weak diffeomorphism $\boldsymbol{u}_2$ such that

$$\beta_2 = \beta[\boldsymbol{u}_2] \quad \text{and} \quad \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{L^2(\mathscr{U};\mathbb{R}^d)} = W_2(\beta[\boldsymbol{u}_1]\mathscr{L}^d, \beta[\boldsymbol{u}_2]\mathscr{L}^d). \tag{5.41}$$

In fact, Brenier Theorem yields the existence of a *unique* optimal transport $\boldsymbol{t} \in L^2(\beta_1; \mathbb{R}^d)$ such that $\boldsymbol{t}_\#\beta_1\mathscr{L}^d = \beta_2\mathscr{L}^d$: thus, if $\boldsymbol{u}_2$ satisfies (5.41), then the plan $\gamma := (\boldsymbol{u}_1, \boldsymbol{u}_2)_\#\mathscr{L}^d \llcorner \mathscr{U}$ is optimal and therefore $\boldsymbol{u}_2 = \boldsymbol{t} \circ \boldsymbol{u}_1$. Moreover, $\boldsymbol{t}$ is $\mathscr{L}^d$-essentially injective and differentiable at $\beta_1\mathscr{L}^d$-a.e. point of $\mathscr{V}$ [24], so that $\boldsymbol{u}_2$ still belongs to $\widetilde{\mathrm{diff}}\,(\mathscr{U}; \mathscr{V})$. It follows that

$$\mathscr{F}_\tau(\boldsymbol{u}, \boldsymbol{u}^{k-1}) = \mathscr{G}_\tau(\beta[\boldsymbol{u}]; \beta^{k-1}) \quad \Leftrightarrow \quad \boldsymbol{u} = \boldsymbol{t} \circ \boldsymbol{u}^{k-1} \quad \boldsymbol{t}_\#\beta^{k-1}\mathscr{L}^d = \beta[\boldsymbol{u}]\mathscr{L}^d, \quad \boldsymbol{t} \text{ is optimal.} \tag{5.42}$$

Therefore, the minimization problem (5.32) associated to $\mathscr{F}_\tau$ is completely reduced to the analogous one associated to $\mathscr{G}_\tau$ (5.33). By (iii) we know that $\bar\beta\mathscr{L}^d \in \mathscr{P}_2(\mathscr{V})$; since the function ψ defined by (5.15) is convex and lower semicontinuous,

$$\lim_{s\to\infty} \psi(s) = \lim_{s\to\infty} s\Phi(1/s) = \lim_{r\to 0^+} \frac{\Phi(r)}{r} = +\infty,$$

and $\mathscr{G}_\tau(\beta^{k-1};\beta^{k-1}) = \phi(\beta^{k-1}) < +\infty$, it follows that the minimum problem for $\mathscr{G}_\tau(\cdot;\beta^{k-1})$ always admits a unique solution β^k with $\beta^k\mathscr{L}^d \in \mathscr{P}_2(\mathscr{V})$ (notice that, in the case $\psi(0) > 0$, the assumption $\phi(\bar\beta) < +\infty$ forces $\mathscr{L}^d(\mathscr{V}) < +\infty$).

The estimate (5.38) (here $\mathscr{V}$ should be bounded) can be proved arguing as in [36], [1], while the Hölder continuity of β^k follows by elliptic regularity theory and the Euler-Lagrange equation, which reads (see for instance [6, Lemma 10.1.2, Thm. 10.4.6])

$$DL_\psi(\beta^k) = \frac{\boldsymbol{s}^{k-1} - \boldsymbol{i}}{h}\beta^k \in L^\infty(\mathscr{V}), \quad \boldsymbol{s}^{k-1} := (\boldsymbol{t}^{k-1})^{-1}.$$

Since $L'_\psi(s) > 0$ in $[\beta_{\min}, \beta_{\max}]$ we conclude that β^k are Lipschitz. The map $\boldsymbol{u}^k$ defined in (5.37) belongs to Diff $(\overline{\mathscr{U}}, \overline{\mathscr{V}})$ by the Caffarelli-Urbas regularity theory (see [12, 14, 13, 15], [39, 40]). $\square$

Before stating our final result, concerning the convergence of the discrete scheme, let us collect some remarks which easily follows by the above proof and which we briefly discussed at the end of the Introduction.

Remark 5.6 (Slope comparison) Recalling that [6, Lemma 3.1.5]

$$\frac{1}{2}|\partial\phi|^2(\beta) = \limsup_{\tau\downarrow 0} \tau^{-1}\Big(\phi(\beta) - \inf_\rho \mathscr{G}_\tau(\rho;\beta)\Big), \tag{5.43}$$

and, analogously,

$$\frac{1}{2}|\partial I|^2(\boldsymbol{u}) = \limsup_{\|\boldsymbol{v}-\boldsymbol{u}\|_2\to 0} \left(\frac{\big(I(\boldsymbol{u}) - I(\boldsymbol{v})\big)^+}{\|\boldsymbol{u}-\boldsymbol{v}\|_2}\right)^2 = \limsup_{\tau\downarrow 0} \tau^{-1}\Big(I(\boldsymbol{u}) - \inf_{\boldsymbol{v}} \mathscr{F}_\tau(\boldsymbol{v};\boldsymbol{u})\Big), \tag{5.44}$$

(5.42) shows that

$$|\partial I|(\boldsymbol{u}) = |\partial\phi|(\beta) \quad \text{if } \boldsymbol{u} \in \widetilde{\mathrm{diff}}\,(\mathscr{U};\mathscr{V}), \quad \beta = \beta[\boldsymbol{u}], \quad I(\boldsymbol{u}) = \phi(\beta) < +\infty. \tag{5.45}$$

On the other hand, when ψ satisfies (4.54), $\mathscr{V}$ is *convex*, and therefore the functional ϕ is *displacement convex* according to (4.52), then [6, Thm. 10.4.9]

$$|\partial\phi|(\beta) < +\infty \quad \Longleftrightarrow \quad L_\psi(\beta) \in W^{1,1}_{\mathrm{loc}}(\mathscr{V}), \quad \nabla L_\psi(\beta) = \beta\boldsymbol{w} \quad \text{for } \boldsymbol{w} \in L^2(\beta\mathscr{L}^d;\mathbb{R}^d) \tag{5.46}$$

and

$$|\partial\phi|^2(\beta) = \int_{\mathscr{V}} |\boldsymbol{\xi}|^2\beta\,dy = \int_{\mathscr{V}} \left|\frac{\nabla L_\psi(\beta)}{\beta}\right|^2 \beta\,dy. \tag{5.47}$$

If $\beta \in L^\infty(\mathscr{V})$ we can approximate $\boldsymbol{w}$ in $L^2(\beta\mathscr{L}^d;\mathbb{R}^d)$ by smooth vector field with compact support: a further integration by parts shows that

$$|\partial\phi|(\beta) = \sup\Big\{ -\int_{\mathscr{V}} L_\psi(\beta)\cdot\mathrm{tr}(D\boldsymbol{\eta})\,dy : \boldsymbol{\eta} \in C^1_c(\mathscr{V};\mathbb{R}^d),\ \int_{\mathscr{V}} |\boldsymbol{\eta}|^2\,\beta\,dy \le 1\Big\}. \tag{5.48}$$

Recalling (5.19) and (5.10), if $\boldsymbol{u} \in \mathrm{Diff}\,(\overline{\mathscr{U}};\overline{\mathscr{V}})$ we get

$$|\partial\phi|(\beta) = \sup\Big\{\bar{\delta}I(\boldsymbol{u};\boldsymbol{\eta}) = \delta I(\boldsymbol{u};\boldsymbol{\eta}\circ\boldsymbol{u}) : \int_{\mathscr{U}} |\boldsymbol{\eta}(\boldsymbol{u}(x))|^2\,dx \le 1\Big\}, \tag{5.49}$$

and therefore

$$|\partial\phi|(\beta) = |\partial I|(\boldsymbol{u}) \le \|\,\mathrm{div}\,DF(D\boldsymbol{u})\|_{L^2(\mathscr{U})}. \tag{5.50}$$

The converse inequality follows easily by taking variations of I along smooth vector field $\boldsymbol{\xi} \in C^\infty_c(\mathscr{U};\mathbb{R}^d)$ and yields the identity

$$|\partial I|(\boldsymbol{u}) = \|\,\mathrm{div}\,DF(D\boldsymbol{u})\|_{L^2(\mathscr{U})} \quad \forall\,\boldsymbol{u} \in \mathrm{Diff}\,(\overline{\mathscr{U}};\overline{\mathscr{V}}). \tag{5.51}$$

Theorem 5.7 (Convergence of $\boldsymbol{u}^{[t/\tau]}$) *Let $\mathscr{U}$, $\mathscr{V}$ be open sets in $\mathbb{R}^d$ with $\mathscr{L}^d(\mathscr{U}) < +\infty$. Assume that*

(a) $\bar{\boldsymbol{u}}$ fulfils (5.30) *for some open set $\mathscr{U}_0 \subset \mathscr{U}$ with full measure in $\mathscr{U}$;*

(b) at least one of the conditions (5.31a,b,c) *holds;*

(c) Φ satisfies the assumption (ii) in Lemma 5.5.

Then

$$\lim_{\tau\downarrow 0} \boldsymbol{u}^{[t/\tau]} = \boldsymbol{u}_t \qquad \text{in } L^2(\mathscr{U}),\ \text{locally uniformly in } [0,+\infty)$$

and $\boldsymbol{u}_t$ is the unique weak solution of (5.11)*, according to Definition 5.3 and Theorem 5.4.*

Proof. The discussion of Example 4.4 shows that the initial datum $\bar{\beta}$ and the form of the equation satisfy the conditions of Theorem 4.12.

Recall that $\boldsymbol{u} = \boldsymbol{Y}(t,\bar{\boldsymbol{u}})$, with $\boldsymbol{Y}$ flow of the vectorfield $\boldsymbol{v}_t$. By applying (5.37) repeatedly we obtain $\boldsymbol{u}^k = \boldsymbol{T}^k \circ \bar{\boldsymbol{u}}$, where $\boldsymbol{T}^k$ is the iterated transport map defined in (4.58).

Then, Theorem 4.12 ensures (recall that $\boldsymbol{t}_{\tau,t} = \boldsymbol{T}^{t/\tau}$ when t/τ is an integer)

$$\lim_{\tau\downarrow 0} \int_{\mathscr{V}} \max_{t\in[0,T]} |\boldsymbol{T}^{[t/\tau]}(\cdot) - \boldsymbol{Y}(t,\cdot)|^p \bar{\beta}(y)\,dy = 0 \qquad \forall T > 0. \tag{5.52}$$

By taking a right composition with $\bar{\boldsymbol{u}}$ in (5.52) the proof is achieved. □

References

[1] M. Agueh, *Existence of solutions to degenerate parabolic equations via the monge-kantorovich theory*, Advances in Differential Equations, to appear, (2002).

[2] L. Ambrosio, *Lecture notes on transport equation and cauchy problem for bv vector fields and applications*, To appear in the proceedings of the School on Geometric Measure Theory, Luminy, October 2003, available at http://cvgmt.sns.it/people/ambrosio/, (2004).

[3] ——, *Transport equation and cauchy problem for bv vector fields*, Inventiones Mathematicae, 158 (2004), pp. 227–260.

[4] L. Ambrosio, E. Crippa, and S. Maniglia, *Traces and fine properties of a bd class of vector fields and applications*, (2004).

[5] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, Clarendon Press, Oxford, 2000.

[6] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows – in Metric Spaces and in the Space of Probability Measures*, Birkhäuser, Basel, 2005.

[7] D. G. Aronson and P. Bénilan, *Régularité des solutions de l'équation des milieux poreux dans $\mathbf{R}^N$*, C. R. Acad. Sci. Paris Sér. A-B, 288 (1979), pp. A103–A105.

[8] F. Bouchut and F. James, *One-dimensional transport equations with discontinuous coefficients*, Nonlinear Anal., 32 (1998), pp. 891–933.

[9] F. Bouchut, F. James, and S. Mancini, *Uniqueness and weak stability for multi-dimensional transport equations with one-sided lipschitz coefficient*, Annali SNS, (5) 4 (2005), pp. 1–25.

[10] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math., 44 (1991), pp. 375–417.

[11] H. Brézis and W. A. Strauss, *Semi-linear second-order elliptic equations in L^1*, J. Math. Soc. Japan, 25 (1973), pp. 565–590.

[12] L. A. Caffarelli, *Some regularity properties of solutions of Monge Ampère equation*, Comm. Pure Appl. Math., 44 (1991), pp. 965–969.

[13] ——, *Boundary regularity of maps with convex potentials*, Comm. Pure Appl. Math., 45 (1992), pp. 1141–1151.

[14] ——, *The regularity of mappings with a convex potential*, J. Amer. Math. Soc., 5 (1992), pp. 99–104.

[15] ——, *Boundary regularity of maps with convex potentials. II*, Ann. of Math. (2), 144 (1996), pp. 453–496.

[16] L. A. Caffarelli and A. Friedman, *Continuity of the density of a gas flow in a porous medium*, Trans. Amer. Math. Soc., 252 (1979), pp. 99–113.

[17] M. G. Crandall and T. M. Liggett, *Generation of semi-groups of nonlinear transformations on general Banach spaces*, Amer. J. Math., 93 (1971), pp. 265–298.

[18] E. DiBenedetto, *Continuity of weak solutions to a general porous medium equation*, Indiana Univ. Math. J., 32 (1983), pp. 83–118.

[19] ——, *Degenerate parabolic equations*, Universitext, Springer-Verlag, New York, 1993.

[20] R. J. DiPerna and P.-L. Lions, *Ordinary differential equations, transport theory and Sobolev spaces*, Invent. Math., 98 (1989), pp. 511–547.

[21] 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.

[22] L. C. Evans, W. Gangbo, and O. Savin, *Diffeomorphisms and nonlinear heat flows*, (2004).

[23] A. Friedman, *Variational principles and free-boundary problems*, Pure and Applied Mathematics, John Wiley & Sons Inc., New York, 1982. A Wiley-Interscience Publication.

[24] W. Gangbo and R. J. McCann, *The geometry of optimal transportation*, Acta Math., 177 (1996), pp. 113–161.

[25] M. Giaquinta, G. Modica, and J. Souček, *Cartesian currents in the calculus of variations. II*, vol. 38 of Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], Springer-Verlag, Berlin, 1998. Variational integrals.

[26] M. Hauray, *On two-dimensional Hamiltonian transport equations with L^p_{loc} coefficients*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 20 (2003), pp. 625–644.

[27] ——, *On Liouville transport equation with force field in BV_{loc}*, Comm. Partial Differential Equations, 29 (2004), pp. 207–217.

[28] 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).

[29] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural′ceva, *Linear and quasilinear equations of parabolic type*, Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23, American Mathematical Society, Providence, R.I., 1967.

[30] C. Le Bris and P.-L. Lions, *Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications*, Ann. Mat. Pura Appl. (4), 183 (2004), pp. 97–130.

[31] N. Lerner, *Transport equation with partially bv velocities*, (2004).

[32] G. M. Lieberman, *Intermediate Schauder theory for second order parabolic equations. II. Existence, uniqueness, and regularity*, J. Differential Equations, 63 (1986), pp. 32–57.

[33] R. J. McCann, *A convexity principle for interacting gases*, Adv. Math., 128 (1997), pp. 153–179.

[34] F. Otto, *Doubly degenerate diffusion equations as steepest descent*, Manuscript, (1996).

[35] ——, *Evolution of microstructure in unstable porous media flow: a relaxational approach*, Comm. Pure Appl. Math., 52 (1999), pp. 873–915.

[36] ——, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations, 26 (2001), pp. 101–174.

[37] M. M. Porzio and V. Vespri, *Hölder estimates for local solutions of some doubly nonlinear degenerate parabolic equations*, J. Differential Equations, 103 (1993), pp. 146–178.

[38] F. Poupaud and M. Rascle, *Measure solutions to the linear multi-dimensional transport equation with non-smooth coefficients*, Comm. Partial Differential Equations, 22 (1997), pp. 337–358.

[39] J. I. E. Urbas, *Global Hölder estimates for equations of Monge-Ampère type*, Invent. Math., 91 (1988), pp. 1–29.

[40] ——, *Regularity of generalized solutions of Monge-Ampère equations*, Math. Z., 197 (1988), pp. 365–393.

[41] C. Villani, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.

[42] W. P. Ziemer, *Interior and boundary continuity of weak solutions of degenerate parabolic equations*, Trans. Amer. Math. Soc., 271 (1982), pp. 733–748.