METHODS AND APPLICATIONS OF ANALYSIS.
Vol. 15, No. 1, pp. 001–030, March 2008

CONSTRUCTION OF THE PARALLEL TRANSPORT IN THE WASSERSTEIN SPACE*

LUIGI AMBROSIO† AND NICOLA GIGLI‡

Abstract. In this paper we study the problem of parallel transport in the Wasserstein spaces $\mathscr{P}_2(\mathbb{R}^d)$. We show that the parallel transport exists along a class of curves whose velocity field is sufficiently smooth, and that we call regular. Furthermore, we show that the class of regular curves is dense in the class of absolutely continuous curves and discuss the problem of parallel transport along geodesics. Most results are extracted from the PhD thesis [8].

Key words. Optimal transport, Wasserstein space, curvature tensor.

AMS subject classifications. 28A33, 35K55, 47J35

1. Introduction.

In the last few years, starting from the seminal papers [14, 4, 12, 9], the geometric and differential properties of the space $\mathscr{P}_2(\mathbb{R}^d)$ of probability measures in $\mathbb{R}^d$ with finite quadratic moments, endowed with the quadratic optimal transportation distance, have been deeply investigated. Motivations for this analysis come from PDE's, Functional Inequalities, Riemannian Geometry. We refer to [16] for a comprehensive presentation of this wide and continuously expanding research field.

A complete theory of the first-order differential properties of $\mathscr{P}_2(\mathbb{R}^d)$ has been estabilished in [1] (starting from the heuristics developed in [14]), without any extra regularity assumption, either on the measures involved, or on the velocity fields. These results lead to a complete theory of gradient flows in $\mathscr{P}_2(\mathbb{R}^d)$ which extends, as a matter of fact, also to the case when $\mathbb{R}^d$ is replaced by more general spaces (see for instance [3, 13, 15]), for instance an infinite-dimensional Hilbert space. We recall the basic facts of the first-order theory in Section 2.

On the other hand, much less is known on the second-order properties of $\mathscr{P}_2(\mathbb{R}^d)$: the only paper we are aware of is [10], where the parallel transport equation and the curvature tensor of $\mathscr{P}(M)$ are computed, mostly at a formal level, when M is a compact Riemannian manifold; in Section 7 we borrow some computations of the sectional curvature of $\mathscr{P}_2(\mathbb{R}^d)$ from [10].

In this paper, whose content is essentially extracted from Chapter 6 of [8], we focus on some analytic aspects: we introduce a class of curves μ_t in $\mathscr{P}_2(\mathbb{R}^d)$ along which the parallel transport of tangent vectors can be defined. In the case when $\mu_t = \rho_t \mathscr{L}^d$ ($\mathscr{L}^d$ being the Lebesgue measure), the PDE corresponding to the parallel transport of a gradient vector field $\nabla\varphi_t$ is, in accordance with [10],

$$\nabla \cdot \Big((\partial_t \nabla \varphi_t + \nabla^2 \varphi_t \cdot v_t)\rho_t \Big) = 0. \tag{1.1}$$

Existence and uniqueness for this evolution problem can presumably studied by direct PDE methods, although difficulties obviously are due to the degeneracy of ρ_t, which results in a lack of uniform ellipticity. Moreover, additional difficulties appear if one is willing to consider unbounded densities ρ_t, and even (in the same spirit of the theory in [1]) measures μ_t that have a singular part with respect to $\mathscr{L}^d$. For these

*Received April 15, 2008; accepted for publication April 28, 2008.
†Scuola Normale Superiore di Pisa, Piazza dei Cavalieri 7, 56126 Pisa, Italy (l.ambrosio@sns.it).
‡McKinsey & Co. (nicolagigli@googlemail.com). This author is a former PhD student at SNS.

reasons, using a suitable Riemannian analogy described in Section 3, we provide a geometric construction of solutions to (1.1). The advantage of this construction is that it provides easily the properties that parallel transport should have. Nevertheless, our construction still requires some regularity condition on the tangent velocity field to μ_t. However, we prove in Section 6 that our class of "regular" curves is dense in the class of all absolutely continuous curves. We discuss also in detail the problem of parallel transport along geodesics, see Proposition 5.19 for a positive result in the case of "forward" transport, and Example 5.20 for a counterexample in the case of "backward" transport.

Finally, in Section 7 we introduce the covariant derivative starting from the parallel transport (in contrast with the usual procedure on manifolds) and explain why this covariant derivative should be qualified as the Levi-Civita derivative on $\mathscr{P}_2(\mathbb{R}^d)$. Finally, we discuss the possibility of defining a distance in the tangent bundle of $\mathscr{P}_2(\mathbb{R}^d)$.

2. First order differentiable calculus in Wasserstein spaces. In this section we recall the main features of the first order differentiable calculus in Wasserstein spaces. We assume that the reader is already familiar with the basic facts regarding optimal transportation and Wasserstein distance, and we shall denote by

$$\Gamma(\mu,\nu) := \big\{\gamma \in \mathscr{P}(\mathbb{R}^d \times \mathbb{R}^d) \,:\, \gamma(A\times\mathbb{R}^d) = \mu(A),\ \gamma(\mathbb{R}^d\times B) = \nu(B)\big\}$$

the set of admissible plans between μ, $\nu \in \mathscr{P}_2(\mathbb{R}^d)$, and by $\Gamma_0(\mu,\nu)$ the set of *optimal* plans, i.e.

$$\gamma \in \Gamma_0(\mu,\nu) \qquad \Longleftrightarrow \qquad \gamma \in \Gamma(\mu,\nu) \text{ and } \int |x-y|^2 d\gamma(x,y) = W_2^2(\mu,\nu).$$

We shall denote by Id the identity map, and use the notation $T_\#$ for the push-forward operator from $\mathscr{P}(X)$ into $\mathscr{P}(Y)$ induced by a Borel map $T : X \to Y$. We also use extensively the short notation L^2_μ and $\|u\|_\mu$ for $L^2(\mu;\mathbb{R}^d)$ and $\|u\|_{L^2(\mu;\mathbb{R}^d)}$ respectively.

Let (E,d) be a metric space. Recall that a curve $x_t : [0,T] \to E$ is said to be *absolutely continuous* if there exists $g \in L^1(0,T)$ satisfying

$$d(x(s),x(t)) \le \int_s^t g(r)dr \qquad \forall s,\, t \in [0,T],\ s \le t.$$

It turns out that for absolutely continuous curves there exists a minimal function g (of course up to Lebesgue negligible sets) with this property, the so-called *metric derivative*, given for a.e. t by (see for instance [1, 1.1.2])

$$|x'|(t) := \lim_{h\to 0} \frac{d(x(t+h),x(t))}{|h|}.$$

In order to describe the differentiable structure of the Wasserstein space we start with purely heuristic considerations, as in [14]: the continuity equation

$$\frac{d}{dt}\mu_t + \nabla\cdot(v_t\mu_t) = 0 \tag{2.1}$$

describes the evolution of a time-dependent mass distribution μ_t under the action of a velocity field v_t. In this perspective Otto suggested to consider the tangent space at μ as $-\nabla\cdot(v\mu)$, where v runs in L^2_μ; furthermore, since optimal transport maps

are gradients, when looking for "minimal" velocity fields it is natural to restrict the admissible velocities to be gradients only. Otto suggested to endow the tangent bundle with the metric inherited from L^2_μ:

$$\langle -\nabla\cdot(v\mu), -\nabla\cdot(w\mu)\rangle_\mu := \int \langle v, w\rangle \, d\mu.$$

We shall consider the tangent space at μ directly as a subset of L^2_μ, retaining the link with the continuity equation. The following result, proved in [1, 8.3.1], provides a complete differential characterization of the class of absolutely continuous curves in the Wasserstein space and makes rigorous this picture.

THEOREM 2.1. *Let $\mu_t : [0,T] \to \mathscr{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve. Then there exists a velocity field $v_t \in L^2_{\mu_t}$ with $\|v_t\|_{\mu_t} \in L^1(0,T)$ such that the continuity equation* (2.1) *holds and*

$$\|v_t\|_{\mu_t} \le |\mu_t'| \qquad \text{for a.e. } t \in (0,T). \tag{2.2}$$

Conversely, if (μ_t, v_t) satisfy (2.1) *and $\|v_t\|_{\mu_t} \in L^1(0,T)$, then μ_t is absolutely continuous and*

$$\|v_t\|_{\mu_t} \ge |\mu_t'| \qquad \text{for a.e. } t \in (0,T). \tag{2.3}$$

The previous result shows that, among all velocity fields v_t compatible with μ_t (in the sense that the continuity equation holds) there exists a distinguished one, of minimal $L^2_{\mu_t}$ norm. This vector field is clearly unique (up to a negligible set of times), thanks to the linearity with respect to v_t of the continuity equation and to the strict convexity of the $L^2_{\mu_t}$ norms.

It turns out that the "optimal" vector field constructed in the proof of the first statement of Theorem 2.1 satisfies, besides (2.2), also

$$v_t \in \overline{\{\nabla\varphi \,:\, \varphi \in C^\infty_c(\mathbb{R}^d)\}}^{L^2_{\mu_t}} \qquad \text{for a.e. } t \in (0,T). \tag{2.4}$$

This, and the previous heuristic remarks, motivate the following definition.

DEFINITION 2.2 (Tangent bundle of $\mathscr{P}_2(\mathbb{R}^d)$). *Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. We define*

$$\mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d)) := \overline{\{\nabla\varphi \,:\, \varphi \in C^\infty_c(\mathbb{R}^d)\}}^{L^2(\mu)}.$$

We shall call *tangent velocity* field the vector field v_t provided by Theorem 2.1 and we shall denote by $P_\mu : L^2_\mu \to \mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d))$ the orthogonal projection.

It turns out that v_t, besides the metric characterization based on (2.2), has also a differential characterization based on (2.4).

PROPOSITION 2.3. *Let (μ_t, v_t) be such that* (2.1) *holds and $\|v_t\|_{\mu_t} \in L^1(0,T)$. Then v_t is tangent if and only if $v_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ for a.e. $t \in (0,T)$.*

Proof. We already said that the tangent vector field satisfies $v_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ for a.e. $t \in (0,T)$. Conversely, if this property holds and w_t is the tangent velocity field, then $\nabla\cdot((v_t - w_t)\mu_t) = 0$ as a space-time distribution. This easily implies that

$$\nabla\cdot((v_t - w_t)\mu_t) = 0 \text{ in } \mathbb{R}^d, \qquad \text{for a.e. } t \in (0,T),$$

so that $v_t - w_t$ is orthogonal in $L^2(\mu_t)$ to all functions $\nabla\varphi$, $\varphi \in C_c^\infty(\mathbb{R}^d)$. But since $v_t - w_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$, this proves that $v_t = w_t$. □

Having defined a tangent velocity field, a satisfactory theory of evolution problems in $\mathscr{P}_2(\mathbb{R}^d)$ based on these concepts can be built on these grounds. We refer to Chapters 10 and 11 of [1] (see also [5, 16]) and we just mention in particular the characterization of gradient flows for convex functionals $F : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{+\infty\}$, based on the evolution variational inequalities

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t,\sigma) + F(\mu_t) \le F(\sigma) \qquad \text{in } (0,T), \text{ for all } \sigma \in \mathscr{P}_2(\mathbb{R}^d).$$

The link between this formulation and the most classical ones is provided by the following purely geometric results (see [1, 8.4.6] and [1, 8.4.7]). The first result relates the tangent field to the infinitesimal behaviour of optimal transport maps (or plans) along the curve; the second result, which is actually a consequence of the first one, provides an explicit formula for the derivative of the Wasserstein distance.

THEOREM 2.4. *Let μ_t be an absolutely continuous curve and let v_t be its tangent velocity field. Then:*

(i) for a.e. $t \in (0,T)$, for any choice of plans $\gamma_h \in \Gamma_0(\mu_t, \mu_{t+h})$, the rescaled transport plans

$$\tilde\gamma_h := (x, \frac{y-x}{h})_\#\gamma_h$$

converge in $\mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ to $(Id \times v_t)_\#\mu_t$.

(ii) for all $\sigma \in \mathscr{P}_2(\mathbb{R}^d)$ and a.e. $t \in (0,T)$ we have

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t,\sigma) = \int \langle v_t(x), x-y\rangle d\gamma(x,y) \qquad \forall \gamma \in \Gamma_0(\mu_t,\sigma).$$

In the particular case when the transport plans γ_h are induced by transport maps T_h (i.e. $(Id \times T_h)_\#\mu_t = \gamma_h$), statement (i) is equivalent to

$$\lim_{h\to 0} \frac{T_h - Id}{h} = v_t \qquad \text{in } L^2_{\mu_t}. \tag{2.5}$$

3. The case of a manifold embedded in $\mathbb{R}^d$. Throughout this section, M will be a C^∞ manifold embedded in $\mathbb{R}^d$ with the induced Riemannian structure. We describe a possible construction of the parallel transport in M, in order to exemplify the construction that will be performed in the Wasserstein space.

Let $\boldsymbol\gamma(t) : [0,1] \to M$ be a fixed C^∞ curve and let $v(t) = \dot{\boldsymbol\gamma}(t) \in T_{\boldsymbol\gamma(t)}M$, $t \in [0,1]$, be the velocity vector of $\boldsymbol\gamma(t)$. We will think to the tangent space $V_t := T_{\boldsymbol\gamma(t)}M$ at the point $\boldsymbol\gamma(t)$ as a linear subspace of $\mathbb{R}^d$ (i.e. we *translate* it to let the origin be included) and we denote by $P_t : \mathbb{R}^d \to V_t$ the orthogonal projection of $\mathbb{R}^d$ onto V_t.

Let $u(t) : [0,1] \to V_t$ be a regular vector field along the curve. In this setting the Levi-Civita derivative of $u(t)$ along $v(t)$ is given by:

$$\nabla_{v(t)}u(t) := P_t\left(\frac{du}{dt}(t)\right). \tag{3.1}$$

More generally, if u, v are vector fields in M and P_x denotes the orthogonal projection on T_xM, $\nabla_v u(x)$ can be defined as $P_x(\tilde u'(0))$, where $\tilde u(t) = u(\boldsymbol\gamma(t))$ and $\boldsymbol\gamma(t)$ is uniquely determined by the conditions $\boldsymbol\gamma(0) = x$ and $\dot{\boldsymbol\gamma}(t) = v(\boldsymbol\gamma(t))$.

Recall that this covariant derivative is uniquely identified, among the other connections, by the following two properties, called *compatibility with the metric* and *torsion free identity*:

$$\frac{d}{dt}\langle u^1(\boldsymbol{\gamma}(t)), u^2(\boldsymbol{\gamma}(t))\rangle_{\boldsymbol{\gamma}(t)} = \langle \nabla_{v(t)} u^1(\boldsymbol{\gamma}(t)), u^2(\boldsymbol{\gamma}(t))\rangle_{\boldsymbol{\gamma}(t)} + \langle u^1(\boldsymbol{\gamma}(t)), \nabla_{v(t)} u^2(\boldsymbol{\gamma}(t))\rangle_{\boldsymbol{\gamma}(t)}, \tag{3.2a}$$

$$\nabla_{u^1} u^2 - \nabla_{u^2} u^1 = [u^1, u^2], \tag{3.2b}$$

where u^1, u^2 are tangent vector fields. The fact that there is at most one connection $\nabla_v u$ for which the previous equations are satisfied is a consequence of the Koszul formula:

$$2\langle \nabla_v u, w\rangle = v(\langle u, w\rangle) + u(\langle v, w\rangle) - w(\langle u, v\rangle) + \langle [u, v], w\rangle - \langle [u, w], v\rangle - \langle [v, w], u\rangle,$$

valid for any vector fields u, v, w defined on the whole of M, and any connection $\nabla_v u$ satisfying equations (3.2). Given that the formula expresses the covariant derivative in terms of the Riemannian metric and the Lie bracket only, the uniqueness follows.

The vector field $u(t)$ is said to be the parallel transport of the vector $u(0)$ along $\boldsymbol{\gamma}(t)$ if

$$P_t\left(\frac{du}{dt}(t)\right) = 0. \tag{3.3}$$

Observe that it is easy to prove the uniqueness of the solution of this equation: indeed by linearity it is sufficient to show that the norm is preserved in time, and this follows by:

$$\frac{d}{dt}|u(t)|^2 = 2\langle \frac{d}{dt}u(t), u(t)\rangle = 2\langle P_t\Big(\frac{d}{dt}u(t)\Big), u(t)\rangle = 0.$$

Therefore the problem is to show the existence of a solution of (3.3) for a given initial datum $u(0)$. This is usually done by using coordinates and solving an appropriate system of differential equations. However, this technique cannot be applied to the space $\mathscr{P}_2(\mathbb{R}^d)$ (we have neither Christoffel symbols, nor coordinates). Here we are going to show how the parallel transport can be constructed using tools which have a Wasserstein analogous.

Let us start with a useful concept.

Definition 3.1 (Angle between subspaces). *Let V_0, $V_1 \subset \mathbb{R}^d$ be two given subspaces, and let P_i, $i = 0, 1$, be the orthogonal projections of $\mathbb{R}^n$ onto V_i. Then the angle $\theta(V_0, V_1) \in [0, \pi/2]$ is defined by:*

$$\cos\theta(V_0, V_1) = \inf_{\substack{v_0 \in V_0 \\ |v_0|=1}} |P_1(v_0)|.$$

It is not difficult to see that, letting $V_i^\perp$, $i = 0, 1$, be the orthogonal complement of V_i, it holds

$$\begin{aligned}\sin\theta(V_0, V_1) &= \sup_{\substack{v_0 \in V_0 \\ |v_0|=1}} |v_0 - P_1(v_0)| = \|P_1^\perp|_{V_0}\| \\ &= \sup_{\substack{v_0 \in V_0,\ |v_0|=1 \\ v_1^\perp \in V_1^\perp,\ |v_1^\perp|=1}} \langle v_0, v_1^\perp\rangle = \sin\theta(V_1^\perp, V_0^\perp),\end{aligned}$$

where $P_i^\perp$, $i = 0, 1$, is the projection onto $V_i^\perp$.

In general $\theta(V_0, V_1) = \theta(V_1, V_0)$ does not hold: for instance, if $V_0 \subset V_1$ we have $\theta(V_0, V_1) = 0$, while $\theta(V_1, V_0) = \pi/2$ if the inclusion is strict. By applying this concept to a smooth curve on M, we clearly have that both functions $(t, s) \mapsto \theta(V_t, V_s)$, $(t, s) \mapsto \theta(V_s, V_t)$ are Lipschitz. Therefore, for some constant C depending on γ, we have:

$$|u - P_s(u)| \le C|u||s - t|, \qquad \forall t,\ s \in [0, 1] \text{ and } u \in V_t, \tag{3.4a}$$

$$|P_s(u^\perp)| \le C|u^\perp||s - t|, \qquad \forall t,\ s \in [0, 1] \text{ and } u^\perp \in V_t^\perp. \tag{3.4b}$$

The idea of the construction is based on the identity:

$$\nabla_{v(0)} P_t(u) = 0, \quad \forall u \in V_0. \tag{3.5}$$

That is: the vectors $P_t(u)$ are a first order approximation at $t = 0$ of the parallel transport. Taking (3.1) into account, (3.5) is equivalent to

$$|P_0(u - P_t(u))| = o(t), \qquad u \in V_0. \tag{3.6}$$

Equation (3.6) follows by applying inequalities (3.4) (note that $u - P_t(u) \in V_t^\perp$):

$$|P_0(u - P_t(u))| \le Ct|u - P_t(u)| \le C^2 t^2 |u|.$$

Now, let $\mathfrak{P}$ be the direct set of all the partitions of $[0, 1]$, where, for $\mathcal{P}$, $\mathcal{Q} \in \mathfrak{P}$, $\mathcal{P} \ge \mathcal{Q}$ if $\mathcal{P}$ is a refinement of $\mathcal{Q}$. For $\mathcal{P} = \{0 = t_0 < t_1 < \cdots < t_N = 1\} \in \mathfrak{P}$ and $u \in V_0$ define $\mathcal{P}(u) \in V_1$ as:

$$\mathcal{P}(u) := P_{t_N}(P_{t_{N-1}}(\cdots(P_{t_0}(u))).$$

Our first goal is to prove that the limit $\mathcal{P}(u)$ for $\mathcal{P} \in \mathfrak{P}$ exists. This will naturally define a curve $t \to u_t \in V_t$ by taking partitions of $[0, t]$ instead of $[0, 1]$: the final goal is to show that this curve is actually the parallel transport of u along the curve γ.

The proof is based on the following lemma.

LEMMA 3.2. *Let $0 \le s_1 \le s_2 \le s_3 \le 1$ be given numbers. Then it holds:*

$$\left|P_{s_3}(u) - P_{s_3}(P_{s_2}(u))\right| \le C^2|u||s_1 - s_2||s_2 - s_3|, \quad \forall u \in V_{s_1}.$$

Proof. Since $P_{s_3}(u) - P_{s_3}(P_{s_2}(u)) = (P_{s_3}(Id - P_{s_2}))(u)$, the proof is a straightforward application of inequalities (3.4). □

From this lemma, an easy induction shows that for any $0 \le s_1 < \cdots < s_N \le 1$ and $u \in V_{s_1}$ we have

$$\begin{aligned}
&\left|P_{s_N}(u) - P_{s_N}(P_{s_{N-1}}(\cdots(P_{s_2}(u)))\right| \\
&\le \left|P_{s_N}(u) - P_{s_N}(P_{s_{N-1}}(u))\right| + \left|P_{s_{N-1}}(u) - (P_{s_{N-1}}(\cdots(P_{s_2}(u)))\right| \\
&\le \cdots \\
&\le C^2|u| \sum_{i=2}^{N-1} |s_1 - s_i||s_i - s_{i+1}| \le C^2|u||s_1 - s_N|^2.
\end{aligned} \tag{3.7}$$

With this result, we can prove existence of the limit of $P(u)$ as P varies in $\mathfrak{P}$.

THEOREM 3.3. *For any $u \in V_0$ there exists the limit of $\mathcal{P}(u)$ as $\mathcal{P}$ varies in $\mathfrak{P}$.*

Proof. We have to prove that, given $\varepsilon > 0$, there exists a partition $\mathcal{P}$ such that

$$|\mathcal{P}(u) - \mathcal{Q}(u)| \leq |u|\varepsilon, \quad \forall \mathcal{Q} \geq \mathcal{P}. \tag{3.8}$$

In order to do so, it is sufficient to find $0 = t_0 < t_1 < \cdots < t_N = 1$ such that $\sum_i |t_{i+1} - t_i|^2 \leq \varepsilon / C^2$, and repeatedly apply equation (3.7) to all partitions induced by $\mathcal{Q}$ in the intervals (t_i, t_{i+1}) (see Section 5 for a more detailed proof in the Wasserstein setting). ☐

Now, for $s \leq t$ we can introduce the maps $T_s^t : V_s \to V_t$ which associate to the vector $u \in V_s$ the limit of the process just described (taking into account partitions of $[s, t]$).

THEOREM 3.4. *For any $t_1 \leq t_2 \leq t_3 \in [0, 1]$ it holds*

$$T_{t_2}^{t_3} \circ T_{t_1}^{t_2} = T_{t_1}^{t_3}. \tag{3.9}$$

Moreover, for any $u \in V_0$ the curve $t \to u_t := T_0^t(u) \in V_t$ is the parallel transport of u along γ.

Proof. We consider those partitions of $[t_1, t_3]$ which contain t_2 and pass to the limit first on $[t_1, t_2]$ and then on $[t_2, t_3]$. To prove the second part of the statement, observe that due to (3.9) it is sufficient to check that the covariant derivative vanishes at 0. Note that from (3.7) it follows that $|P_t(u) - u_t| \leq C^2 t^2$, therefore the thesis follows from (3.5). ☐

4. Angle between tangent spaces in $\mathscr{P}_2(\mathbb{R}^d)$. The construction we did on regular manifolds embedded in $\mathbb{R}^d$ shows that the key step which allows to prove the existence of the parallel transport is the Lipschitz property of the angle between tangent spaces. In this section we introduce the analogous notion of angle for the space $\mathscr{P}_2(\mathbb{R}^d)$ and analyze its properties.

An important difference with the case of a manifold embedded in $\mathbb{R}^d$ is that the two spaces $\mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d))$ and $\mathrm{Tan}_\nu(\mathscr{P}_2(\mathbb{R}^d))$ are not (affine) subspaces of a larger Hilbert space, therefore we cannot directly imitate the definition of angle given in the previous section. However, a natural way to embed L^2_ν into L^2_μ is given by the composition with a map T pushing μ into ν. Thus, we give the following general definition.

DEFINITION 4.1 (Translation of vectors through a map). *For any pair of measures $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$ and any transport map T between μ and ν we define the* translation τ_T *of a vector $f \in L^2_\nu$ into the vector $\tau_T(f) \in L^2_\mu$ as the map $\tau_T(f) := f \circ T$.*

Clearly the translation through a map T is an isometry from L^2_μ to L^2_ν. The definition of angle comes out naturally.

DEFINITION 4.2 (Angle between tangent spaces through a map). *Let $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$ and let T be a transport map from μ to ν. Then the angle $\theta_T(\mu, \nu) \in [0, \pi/2]$ between the tangent spaces at μ and ν through the map T is given by*

$$\cos \theta_T(\mu, \nu) := \inf \|P_\mu(v \circ T)\|_\mu,$$

where the infimum is taken among all $v \in \mathrm{Tan}_\nu(\mathscr{P}_2(\mathbb{R}^d))$ such that $\|v\|_\nu = 1$.

It is important to note that the angle between the tangent spaces at two measures, strongly depends on the transport maps used. Observe also that, even if we assume

that the transport map T is invertible, the angle $\theta_T(\mu,\nu)$ is in general *n*ot equal to the angle $\theta_{T^{-1}}(\nu,\mu)$: this corresponds to the fact that there exist two angles between subspaces V_1 and V_2 of $\mathbb{R}^d$, depending on whether we are considering projections from V_1 onto V_2 or from V_2 to V_1.

The fundamental bound on the angle we are going to use in the sequel is given by the following proposition: the key requirement is the Lipschitz property of the transport map, while there is *n*o regularity assumption on the measures involved.

PROPOSITION 4.3. *Let $\mu,\nu\in\mathscr{P}_2(\mathbb{R}^d)$ and let $T\in L^2_\mu$ be a transport map from μ to ν. Suppose that T is Lipschitz. Then it holds*

$$\sin\theta_T(\mu,\nu)\le \mathrm{Lip}(T-Id). \tag{4.1}$$

Proof. The statement is equivalent to

$$\|\nabla\varphi\circ T-P_\mu(\nabla\varphi\circ T)\|_\mu\le\|\nabla\varphi\|_\nu\,\mathrm{Lip}(T-Id),\qquad\forall\varphi\in C^\infty_c(\mathbb{R}^d). \tag{4.2}$$

Let us suppose first that $T-Id\in C^\infty_c(\mathbb{R}^d)$. In this case the map $\varphi\circ T$ is in $C^\infty_c(\mathbb{R}^d)$, too, and therefore $\nabla(\varphi\circ T)=\nabla T\cdot(\nabla\varphi)\circ T$ belongs to $\mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d))$. From the minimality properties of the projection we get:

$$\begin{aligned}\|\nabla\varphi\circ T-P_\mu(\nabla\varphi\circ T)\|_\mu&\le\|\nabla\varphi\circ T-\nabla T\cdot(\nabla\varphi)\circ T\|_\mu\\&=\left(\int|(I-\nabla T(x))\cdot\nabla\varphi(T(x))|^2d\mu(x)\right)^{1/2}\\&\le\left(\int|\nabla\varphi(T(x))|^2\|\nabla(Id-T)(x)\|^2_{op}d\mu(x)\right)^{1/2}\\&\le\|\nabla\varphi\|_\nu\,\mathrm{Lip}(T-Id),\end{aligned}$$

where I is the identity matrix and $\|\nabla(Id-T)(x)\|_{op}$ is the operator norm of the linear functional from $\mathbb{R}^d$ to $\mathbb{R}^d$ given by $v\mapsto\nabla(Id-T)(x)\cdot v$.

Now turn to the general case. Find a sequence $(T_n-Id)\subset C^\infty_c(\mathbb{R}^d)$ such that $T_n\to T$ uniformly on compact sets and $\overline{\lim}_n\mathrm{Lip}(T_n-Id)\le\mathrm{Lip}(T-Id)$. It is clear that for such a sequence it holds $\|T-T_n\|_\mu\to0$, and we have

$$\begin{aligned}\|\nabla\varphi\circ T-P_\mu(\nabla\varphi\circ T)\|_\mu&\le\|\nabla\varphi\circ T-\nabla(\varphi\circ T_n)\|_\mu\\&\le\|\nabla\varphi\circ T-\nabla\varphi\circ T_n\|_\mu+\|\nabla\varphi\circ T_n-\nabla(\varphi\circ T_n)\|_\mu\\&\le\mathrm{Lip}(\nabla\varphi)\|T-T_n\|_\mu+\|\nabla\varphi\circ T_n\|_\mu\,\mathrm{Lip}(T_n-Id).\end{aligned}$$

Letting $n\to+\infty$ we get the thesis. □

5. Regular curves and parallel transport along them. In this section we introduce a class of sufficiently regular curves in the Wasserstein space along which a parallel transport can be defined.

Having the Riemannian analogy in mind (see in particular (3.6)), we would like to say that $u_t\in\mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ is a parallel transport if $\|P^h_t(u(t+h))-u(t)\|_{\mu_t}=o(h)$, where P^h_t are suitable projections from $\mathrm{Tan}_{\mu_{t+h}}(\mathscr{P}_2(\mathbb{R}^d))$ to $\mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ induced by maps pushing μ_t to μ_{t+h}, as in Definition 4.1. It is natural to relate these maps to the tangent vector of the curve, see also Remark 5.15.

We know from the classical Cauchy-Lipschitz theory that, if the tangent vector v_t of μ_t satisfies $\int_0^1\mathrm{Lip}(v_t)dt<+\infty$, then the flow maps exist and are Lipschitz

functions of the space variable. More precisely, there exists a unique family of maps $\mathbf{T}(s,t,x) : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, which we call the *flow of the curve* μ_t, absolutely continuous with respect to t and Lipschitz with respect to x, satisfying

$$\begin{cases} \mathbf{T}(s,s,x) = x, \\ \dfrac{d}{d\tau}\mathbf{T}(s,\tau,x)_{|\tau=t} = v_t(\mathbf{T}(s,t,x)), \\ \mathbf{T}(t,r,\mathbf{T}(s,t,x)) = \mathbf{T}(s,r,x), \\ \mathbf{T}(s,t,\cdot)_{\#}\mu_s = \mu_t. \end{cases} \tag{5.1}$$

Here all the equations except the second one hold for all $x \in \mathbb{R}^d$ and $s,\, t \in [0,1]$. The second one holds, given $x \in \mathbb{R}^d$ and $s \in [0,1]$, for a.e. t; it can be written in a pointwise way, including also the first one, as $\mathbf{T}(s,t,x) - x = \int_s^t v_\tau(\mathbf{T}(s,\tau,x))d\tau$.

Definition 5.1 (Regular curves). *Let $\mu_t : [0,1] \to \mathscr{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve and let $v_t \in L^2_{\mu_t}$ be its tangent velocity field. We say that μ_t is regular if*

$$\int_0^T \mathrm{Lip}(v_t)dt < +\infty.$$

Observe that we are making *no* regularity assumption on the measures μ_t. Strictly speaking, in the definition of regularity we mean that v_t has, for almost every t, a Lipschitz continuous version, and that the (smallest) Lipschitz constant of this version is integrable in time (recall that v_t are uniquely determined only up to μ_t-negligible sets).

In the following we will always assume (this is not really restrictive, up to a reparameterization) that the regular curve is parameterized in $[0,1]$.

The key property of regular curves needed to prove the existence of the parallel transport is the following bound on the Lipschitz constant of $\mathbf{T}(s,t,\cdot) - Id$:

$$\mathrm{Lip}(\mathbf{T}(s,t,\cdot) - Id) \le \exp\left(\left|\int_t^s \mathrm{Lip}(v_r)dr\right|\right) - 1, \qquad t,\, s \in [0,1]. \tag{5.2}$$

This inequality is a simple consequence of equations (5.1), and we recall its proof for the sake of completeness.

Proposition 5.2. *Let $\mathbf{T}(s,t,\cdot)$ be the flow maps of a regular curve μ_t. Then:*

$$\mathrm{Lip}(\mathbf{T}(s,t,\cdot)) \le \exp\left(\left|\int_t^s \mathrm{Lip}(v_r)dr\right|\right), \qquad t,\, s \in [0,1],$$
$$\mathrm{Lip}(\mathbf{T}(s,t,\cdot) - Id) \le \exp\left(\left|\int_t^s \mathrm{Lip}(v_r)dr\right|\right) - 1, \qquad t,\, s \in [0,1].$$

Proof. The first equation follows by a direct application of Gronwall lemma to the differential inequality

$$\begin{aligned} \frac{d}{dt}|\mathbf{T}(s,t,x) - \mathbf{T}(s,t,y)|^2 =& 2\langle \mathbf{T}(s,t,x) - \mathbf{T}(s,t,y), v_t(\mathbf{T}(s,t,x)) - v_t(\mathbf{T}(s,t,y))\rangle \\ \le& 2|\mathbf{T}(s,t,x) - \mathbf{T}(s,t,y)|^2\,\mathrm{Lip}(v_t). \end{aligned}$$

For the second one, observe that

$$\begin{aligned}&\frac{d}{dt}|\mathbf{T}(s,t,x)-x-\mathbf{T}(s,t,y)+y|^2\\ &\le 2\langle \mathbf{T}(s,t,x)-x-\mathbf{T}(s,t,y)+y, v_t(\mathbf{T}(s,t,x))-v_t(\mathbf{T}(s,t,y))\rangle\\ &\le 2|\mathbf{T}(s,t,x)-x-\mathbf{T}(s,t,y)+y||x-y|\,\mathrm{Lip}(v_t)\,\mathrm{Lip}(\mathbf{T}(s,t,\cdot)),\end{aligned}$$

therefore the conclusion follows by integrating from s to t the inequality

$$\frac{d}{dt}|\mathbf{T}(s,t,x)-x-\mathbf{T}(s,t,y)+y|\le |x-y|\,\mathrm{Lip}(v_t)\exp\left(\left|\int_t^s \mathrm{Lip}(v_r)dr\right|\right).$$

□

Definition 5.3 (Absolutely continuous vector fields). *Let μ_t be a regular curve and let $u_t\in L^2_{\mu_t}$ be a vector field along it. We say that u_t is absolutely continuous if the maps $u_t\circ\mathbf{T}(s,t,\cdot)\in L^2_{\mu_s}$ are absolutely continuous for any $s\in[0,1]$, where $\mathbf{T}(s,t,\cdot)$ are the flow maps of μ_t.*

For an absolutely continuous vector field u_t, we will write $\frac{d}{dt}u_t\in L^2_{\mu_t}$ for its derivative, defined by:

$$\frac{d}{dt}u_t:=\lim_{h\to0}\frac{u_{t+h}\circ\mathbf{T}(t,t+h,\cdot)-u_t}{h}=\frac{d}{dr}(u_r\circ\mathbf{T}(s,r,\cdot))_{|r=t}\circ\mathbf{T}(t,s,\cdot),\quad\forall s\in[0,1]. \tag{5.3}$$

Given that the right composition with $\mathbf{T}(s,t,\cdot)$ is an isometry from $L^2_{\mu_t}$ to $L^2_{\mu_s}$, it is clear that a vector field u_t is absolutely continuous if and only if for *some* $s\in[0,1]$ the curve $t\mapsto u_t\circ\mathbf{T}(s,t,\cdot)\in L^2_{\mu_s}$ is absolutely continuous. Using the second identity in (5.3) one can easily prove the chain rules

$$\frac{d}{dt}\langle u_t^1,u_t^2\rangle_{\mu_t}=\langle\frac{d}{dt}u_t^1,u_t^2\rangle_{\mu_t}+\langle u_t^1,\frac{d}{dt}u_t^2\rangle_{\mu_t}, \tag{5.4}$$

$$\frac{d}{dt}\langle\nabla\eta,u_t\rangle_{\mu_t}=\langle\nabla^2\eta\cdot v_t,u_t\rangle_{\mu_t}+\langle\nabla\eta,\frac{d}{dt}u_t\rangle_{\mu_t}\qquad\forall\eta\in C^\infty_c(\mathbb{R}^d). \tag{5.5}$$

for a.e. $t\in(0,1)$, whenever u^1, u^2, u are absolutely continuous. Notice also that $t\mapsto\|u_t\|_{\mu_t}$ is absolutely continuous whenever u_t is absolutely continuous.

It is important to underline that the definition of derivative of an absolutely continuous vector field allows us to take derivative of a function u_t whose range belongs to different L^2 spaces as t varies: actually these spaces can be quite different from each other, if the support of μ_t does depend on time.

In the following we will use, without explicit mention, the fact that for any $\varphi\in C^\infty_c(\mathbb{R}^d)$ the vector field $t\mapsto\nabla\varphi\in L^2_{\mu_t}$ is tangent and absolutely continuous. Its derivative in the sense of the above definition is easily seen to be equal to $\nabla^2\varphi\cdot v_t$.

We now give a precise definition of parallel transport along a regular curve.

Definition 5.4 (Parallel transport along regular curves). *Let μ_t be a regular curve and let $u_t\in\mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ a vector field defined along the curve. We say that u_t is a parallel transport if it is absolutely continuous and*

$$P_{\mu_t}\left(\frac{d}{dt}u_t\right)=0\qquad\text{for a.e. } t\in(0,1). \tag{5.6}$$

Equation (5.6) may be equivalently written as:

$$\lim_{h\to 0} P_{\mu_t}\left(\frac{u_{t+h}\circ \mathbf{T}(t,t+h,\cdot)-u_t}{h}\right)=0 \qquad \text{in } L^2_{\mu_t}, \text{ for a.e. } t\in(0,1). \tag{5.7}$$

Another equivalent characterization, thanks to (5.5), is:

$$\frac{d}{dt}\langle \nabla\eta, u_t\rangle_{\mu_t} = \langle \nabla^2\eta\cdot v_t, u_t\rangle_{\mu_t} \qquad \text{for a.e. } t\in(0,1), \text{ for all } \eta\in C^\infty_c(\mathbb{R}^d), \tag{5.8}$$

or, in integral form

$$\langle \nabla\eta, u_s\rangle_{\mu_s} - \langle \nabla\eta, u_t\rangle_{\mu_t} = \int_t^s \langle \nabla^2\eta\cdot v_r, u_r\rangle_{\mu_r} dr \qquad \text{for all } t<s\in(0,1),\ \eta\in C^\infty_c(\mathbb{R}^d). \tag{5.9}$$

Observe that this equations makes sense even if the underlying curve μ_t is not regular, but only absolutely continuous. We will come back to this point at the end of section 7.

It is also easy to check that also the concept of parallel transport is invariant under reparameterization: if μ_t is a regular curve, u_t is a parallel transport along it and $r:[0,R]\to[0,1]$ is a Lipschitz reparameterization of $[0,1]$, then $\tilde\mu_s := \mu_{r(s)}$ is regular in $[0,R]$ and $\tilde u_s := u_{r(s)}$ is a parallel transport along it.

Proposition 5.5 (Linearity and conservation of norm). *Let u_t, u^1_t, u^2_t be parallel transports along a regular curve μ_t and let λ^1, $\lambda^2\in\mathbb{R}$. Then $\|u_t\|_{\mu_t}$ is constant and $\lambda^1 u^1_t+\lambda^2 u^2_t$ is a parallel transport.*

Proof. The claim on $\lambda^1u^1_t+\lambda^2u^2_t$ follows directly by the linearity of equation (5.6). To prove that the norm is constant, just recall that $t\mapsto \|u_t\|^2_{\mu_t}$ is absolutely continuous and that, thanks to (5.4), its derivative is given by

$$\frac{d}{dt}\|u_t\|^2_{\mu_t} = 2\langle u_t, \frac{d}{dt}u_t\rangle_{\mu_t} = 2\langle u_t, P_{\mu_t}\left(\frac{d}{dt}u_t\right)\rangle_{\mu_t} = 0.$$

□

As a direct consequence we get the uniqueness of the parallel transport and the conservation of the scalar product.

Corollary 5.6 (Uniqueness of parallel transport). *Let μ_t be a regular curve and let $u_0\in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^d))$. Then there exists at most one parallel transport u_t along μ_t such that $u_0=u$.*

Corollary 5.7 (Conservation of scalar product). *Let u^1_t, u^2_t be parallel transports along the regular curve μ_t. Then $t\mapsto\langle u^1_t,u^2_t\rangle_{\mu_t}$ is constant.*

Observe that for parallel transports we have an explicit bound on the norm of $\frac{d}{dt}u_t$ which depends only on the Lipschitz constant of the vectors v_t.

Proposition 5.8. *Let μ_t be a regular curve and let u_t be a parallel transport along it. Then*

$$\left\|\frac{d}{dt}u_t\right\|_{\mu_t} \le \|u_0\|_{\mu_0}\,\mathrm{Lip}(v_t) \qquad \text{for a.e. } t\in(0,1). \tag{5.10}$$

Proof. We will prove that equation (5.10) is fulfilled at any Lebesgue point t of the function $s \to \mathrm{Lip}(v_s)$. Fix such t and observe that

$$\|u_s \circ \mathbf{T}(t,s,\cdot) - u_t\|_{\mu_t} \le \|P_{\mu_t}(u_s \circ \mathbf{T}(t,s,\cdot)) - u_t\|_{\mu_t} + \|P^\perp_{\mu_t}(u_s \circ \mathbf{T}(t,s,\cdot))\|_{\mu_t}.$$

Dividing by $|s-t|$ and letting $s \to t$ we have that the first term goes to 0 by definition of parallel transport, while for the second one we have the following estimate, based on Proposition 4.3 and Proposition 5.2:

$$\begin{aligned}\limsup_{s\to t^+} \frac{\|P^\perp_{\mu_t}(u_s \circ \mathbf{T}(t,s,\cdot))\|_{\mu_t}}{s-t} &\le \limsup_{s\to t^+} \frac{1}{s-t}\|u_t\|_{\mu_t}\,\mathrm{Lip}(\mathbf{T}(t,s,\cdot) - Id)\\ &\le \limsup_{s\to t^+} \|u_0\|_{\mu_0}\frac{e^{\int_t^s \mathrm{Lip}(v_r)dr} - 1}{s-t} = \|u_0\|_{\mu_0}\,\mathrm{Lip}(v_t).\end{aligned}$$

The case $s \to t^-$ is analogous. □

Now we turn to the proof of the existence of the parallel transport: μ_t will be a fixed regular curve, v_t its tangent vector field and $\mathbf{T}(t,s,x)$ its flow. In order to enlighten the notation we define

$$\begin{aligned}D(t,s) &:= e^{\int_t^s \mathrm{Lip}(v_r)dr} - 1, \quad 0 \le t \le s \le 1,\\ D(t,s) &:= D(s,t) \qquad 0 \le s \le t \le 1.\end{aligned}$$

Then we denote by τ_t^s be the linear isometry from $L^2_{\mu_t}$ to $L^2_{\mu_s}$ given by the right composition with $\mathbf{T}(s,t,\cdot)$. Note that from the group property of $\mathbf{T}(s,t,\cdot)$ it follows

$$\tau_t^r = \tau_s^r \circ \tau_t^s, \quad \forall t,\, s,\, r \in [0,1]. \tag{5.11}$$

Moreover we define

$$\mathscr{P}_t^s(u) := P_{\mu_s}\Big(\tau_t^s(u)\Big).$$

Observe that the maps $\mathscr{P}_t^s$ are non-expansive and that, by inequality (5.2) and Proposition 4.3 we get:

$$\|\mathscr{P}_s^t(w)\|_{\mu_t} \le \|w\|_{\mu_s} D(t,s), \qquad t,\, s \in [0,1],\ w \in \mathrm{Tan}^\perp_{\mu_s}(\mathscr{P}_2(\mathbb{R}^d)), \tag{5.12a}$$

$$\|\tau_t^s(u) - \mathscr{P}_t^s(u)\|_{\mu_s} \le \|u\|_{\mu_t} D(t,s), \qquad t,\, s \in [0,1],\ u \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d)). \tag{5.12b}$$

To prove the existence of the transport we proceed as in the first section: let $\mathfrak{P}$ be the direct set of all partitions of $[0,1]$, where, for $\mathcal{P}$, $\mathcal{Q} \in \mathfrak{P}$, $\mathcal{Q} \ge \mathcal{P}$ if $\mathcal{Q}$ is a refinement of $\mathcal{P}$. For $\mathcal{P} = \{0 = t_0 < t_1 < \cdots < t_N = 1\} \in \mathfrak{P}$ and $u \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^d))$ define $\mathcal{P}(u) \in \mathrm{Tan}_{\mu_1}(\mathscr{P}_2(\mathbb{R}^d))$ as:

$$\mathcal{P}(u) := \mathscr{P}^1_{t_{N-1}}(\mathscr{P}^{t_{N-1}}_{t_{N-2}}(\cdots(\mathscr{P}^{t_1}_0(u)))).$$

Finally, we set $D^2(\mathcal{P}) := \sum_i D^2(t_i, t_{i+1})$.

We will prove first that there exists a unique limit $\mathcal{T}_0^1(u) \in \mathrm{Tan}_{\mu_1}(\mathscr{P}_2(\mathbb{R}^d))$ of $\mathcal{P}(u)$ as $\mathcal{P}$ varies in $\mathfrak{P}$; then we will define a curve u_t with $u_t = \mathcal{T}_0^t(u) \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ by considering partitions of $[0,t]$, and finally prove that this curve is the parallel transport of u along the curve μ_t.

LEMMA 5.9. *It holds*

$$D(t_1,s_1) \le D(t_2,s_2), \qquad \forall [t_1,s_1] \subset [t_2,s_2] \subset [0,1], \tag{5.13a}$$

$$\sum_{i=1}^{n-1} D(t_i,t_{t+1}) \le D(t,s), \quad t=t_1<\ldots<t_n=s, \tag{5.13b}$$

$$\lim_{s\to t} \frac{D^2(t,s)}{|s-t|} = 0 \qquad \text{for a.e. } t\in[0,1], \tag{5.13c}$$

$$\lim_{\mathcal{P}\in\mathfrak{P}} D^2(\mathcal{P}) = 0. \tag{5.13d}$$

Proof. Equation (5.13a) is clear. For (5.13b) we need to prove that $e^a-1+e^b-1 \le e^{a+b}-1$ for positive a, b, which is obvious.

The convexity of $c \to e^c-1$ in $[0, \int_0^1 \mathrm{Lip}(v_r)dr]$ gives

$$D(t,s) \le \left(\frac{e^{\int_0^1 \mathrm{Lip}(v_r)dr}-1}{\int_0^1 \mathrm{Lip}(v_r)dr} \right) \int_t^s \mathrm{Lip}(v_r)dr, \tag{5.14}$$

from which, taking the integrability of $\mathrm{Lip}(v_t)$ into account, (5.13c) follows at every Lebesgue point of $t \mapsto \mathrm{Lip}(v_t)$. Finally, from (5.14) we get

$$\begin{aligned}\sum_{i=0}^{N-1} D^2(t_{i+1},t_i) &\le C \sum_{i=0}^{N-1} \left(\int_{t_i}^{t_{i+1}} \mathrm{Lip}(v_r)dr \right)^2 \\ &\le C \max_i \left\{ \int_{t_i}^{t_{i+1}} \mathrm{Lip}(v_r)\,dr \right\} \int_0^1 \mathrm{Lip}(v_r)\,dr,\end{aligned}$$

from which (5.13d) follows, taking the absolute continuity property of the integral into account. □

The following lemma corresponds to Lemma 3.2:

LEMMA 5.10. *Let $0 \le s_1 \le s_2 \le s_3 \le 1$ and let $u \in \mathrm{Tan}_{\mu_{s_1}}(\mathscr{P}_2(\mathbb{R}^d))$. Then:*

$$\left\| \mathscr{P}_{s_1}^{s_3}(u) - \mathscr{P}_{s_2}^{s_3}(\mathscr{P}_{s_1}^{s_2}(u)) \right\|_{\mu_{s_3}} \le \|u\|_{\mu_{s_1}} D(s_1,s_2) D(s_2,s_3). \tag{5.15}$$

Proof. Observe that, thanks to the semigroup property (5.11), we have

$$\mathscr{P}_{s_1}^{s_3}(u) - \mathscr{P}_{s_2}^{s_3}(\mathscr{P}_{s_1}^{s_2}(u)) = \mathscr{P}_{s_2}^{s_3}(\tau_{s_1}^{s_2}(u) - \mathscr{P}_{s_1}^{s_2}(u)),$$

and that $\tau_{s_1}^{s_2}(u) - \mathscr{P}_{s_1}^{s_2}(u) \in \mathrm{Tan}^{\perp}_{\mu_{s_2}}(\mathscr{P}_2(\mathbb{R}^d))$. Therefore the thesis follows by a direct application of inequalities (5.12). □

COROLLARY 5.11. *Let $\mathcal{P} = \{t = t_0 < t_1 < \cdots < t_n = s\}$ be a partition of $[t,s] \subset [0,1]$ and let $\mathcal{Q}$ be a refinement of $\mathcal{P}$. Then:*

$$\|\mathcal{P}(u) - \mathcal{Q}(u)\|_{\mu_s} \le \|u\|_{\mu_t} D^2(\mathcal{P}) \qquad \text{for every } u \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d)). \tag{5.16}$$

Proof. Without loss of generality we may assume $[t,s]=[0,1]$. Fix $i<n$ such that (t_i,t_{i+1}) contains some element of $\mathcal{Q}$ and write $\mathcal{Q}\cap[t_i,t_{i+1}] = \{t_i = s_{i,0} < s_{i,1} < \cdots < s_{i,k(i)} = t_{i+1}\}$ for some $k(i) \ge 1$. Now, we claim that

$$\|\mathscr{P}_{s_{i,0}}^{s_{i,k(i)}}(u_{t_i}) - \mathscr{P}_{s_{i,k_i-1}}^{s_{i,k_i}}(\mathscr{P}_{s_{i,k_i-2}}^{s_{i,k_i-1}}(\cdots(\mathscr{P}_{s_{i,0}}^{s_{i,1}}(u_{t_i}))))\|_{\mu_{t_{i+1}}} \le \|u_{t_i}\|_{\mu_{t_i}} D^2(t_i,t_{i+1}) \tag{5.17}$$

for all $u_{t_i} \in \mathrm{Tan}_{\mu_{t_i}}(\mathscr{P}_2(\mathbb{R}^d))$. Indeed, the right hand side of (5.17) can be estimated by

$$\begin{aligned}
&\|\mathscr{P}^{s_{i,k(i)}}_{s_{i,0}}(u_{t_i}) - \mathscr{P}^{s_{i_{k(i)}}}_{s_{i,k(i)-1}}(\mathscr{P}^{s_{i,k(i)-1}}_{s_{i,0}}(u_{t_i}))\|_{\mu_{t_{i+1}}}\\
&+ \|\mathscr{P}^{s_{i,k(i)}}_{s_{i,k(i)-1}}(\mathscr{P}^{s_{i,k(i)-1}}_{s_{i,0}}(u_{t_i})) - \mathscr{P}^{s_{i,k(i)}}_{s_{i,k(i)-1}}(\mathscr{P}^{s_{i,k(i)-1}}_{s_{i,k(i)-2}}(\cdots(\mathscr{P}^{s_{i,1}}_{s_{i,0}}(u_{t_i}))))\|_{\mu_{t_{i+1}}}\\
&\le \|u_{t_i}\|_{\mu_{t_i}} D(s_{i,0}, s_{i,k(i)-1}) D(s_{i,k(i)-1}, s_{i,k(i)})\\
&+ \|\mathscr{P}^{s_{i,k(i)-1}}_{s_{i,0}}(u_{t_i}) - \mathscr{P}^{s_{i,k(i)-1}}_{s_{i,k(i)-2}}(\mathscr{P}^{s_{i,k(i)-2}}_{s_{i,k(i)-3}}(\cdots(\mathscr{P}^{s_{i,0}}_{t_i}(u_{t_i}))))\|_{\mu_{t_{i+1}}}\\
&\le \cdots\\
&\le \|u_{t_i}\|_{\mu_{t_i}} \sum_{j=0}^{k(i)-1} D(s_{i,0}, s_{i,j}) D(s_{i,j}, s_{i,j+1}) \le \|u_{t_i}\|_{\mu_{t_i}} D(t_i, t_{i+1}) \sum_{j=0}^{k(i)-1} D(s_{i,j}, s_{i,j+1})\\
&\le \|u_{t_i}\|_{\mu_{t_i}} D^2(t_i, t_{i+1}).
\end{aligned}$$

Now, let us assume that (t_0, t_1) contains some element of $\mathcal{Q}$ and let $\mathcal{P}' = [t_1, 1] \cap \mathcal{P}$, $\mathcal{Q}' = [t_1, 1] \cap \mathcal{Q}$, $u \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^d))$ and $v, w \in \mathrm{Tan}_{\mu_{t_1}}(\mathscr{P}_2(\mathbb{R}^d))$ be such that $\mathcal{P}(u) = \mathcal{P}'(v)$ and $\mathcal{Q}(u) = \mathcal{Q}'(w)$. Then, the inequality (5.17) with $i = 0$ reads

$$\|v - w\|_{\mu_{t_1}} \le \|u\|_{t_0} D^2(t_0, t_1),$$

(the estimate is trivial if $\mathcal{Q}' \cap (t_0, t_1) = \emptyset$, because $v = w$) so that

$$\begin{aligned}
\|\mathcal{P}(u) - \mathcal{Q}(u)|_{\mu_{t_n}} &\le \|\mathcal{P}'(v) - \mathcal{Q}'(v)\|_{\mu_{t_n}} + \|\mathcal{Q}'(v) - \mathcal{Q}'(w)\|_{\mu_{t_n}}\\
&\le \|\mathcal{P}'(v) - \mathcal{Q}'(v)\|_{\mu_{t_n}} + \|u\|_{t_0} D^2(t_0, t_1).
\end{aligned}$$

Since $\|v\|_{t_1} \le \|u\|_{t_0}$ we can apply repeatedly (5.17) in the intervals (t_i, t_{i+1}) to obtain $\|\mathcal{P}(u) - \mathcal{Q}(u)\|_{\mu_1} \le \|u\|_{\mu_0} D^2(\mathcal{P})$. □

The following result follows directly from the previous corollary and from (5.13d).

THEOREM 5.12 (Existence of the limit of $\mathcal{P}(u_0)$). *Let μ_t be a regular curve and let $u_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^d))$. Then $\lim_{P \in \mathfrak{P}} \mathcal{P}(u_0)$ exists.*

Define $\mathcal{T}^1_0(u_0)$ as the vector obtained by the limit process described above and observe that, by repeating the arguments to the restriction of μ_t to the interval $[t, s]$, we can define a map $\mathcal{T}^s_t : \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d)) \to \mathrm{Tan}_{\mu_s}(\mathscr{P}_2(\mathbb{R}^d))$ whenever $t \le s$. Furthermore, by considering the curve $t \to \mu_{1-t}$, we can define the maps $\mathcal{T}^s_t$ even for $t > s$.

PROPOSITION 5.13 (Group property). *Let μ_t be a regular curve and let $\mathcal{T}^s_t : \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d)) \to \mathrm{Tan}_{\mu_s}(\mathscr{P}_2(\mathbb{R}^d))$ be defined as above. Then*

$$\mathcal{T}^s_t \circ \mathcal{T}^t_r = \mathcal{T}^s_r, \qquad \forall r,\, s,\, t \in [0, 1]. \tag{5.18}$$

Proof. Let us first assume $r \le t \le s$. In this case it is sufficient to observe that, by definition of limit over a direct set, the limit over all partitions coincides with the limit over all partitions which contain the point t. The thesis then follows easily. For the general case it is sufficient to prove that $\mathcal{T}^s_t = (\mathcal{T}^t_s)^{-1}$, or, without loss of generality, that $\mathcal{T}^1_0 = (\mathcal{T}^0_1)^{-1}$. The latter equation will follow if we show that

$$\lim_{\mathcal{P} \in \mathfrak{P}} \|u - \mathcal{P}^{-1}(\mathcal{P}(u))\|_{\mu_0} = 0 \qquad \forall u \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^d)), \tag{5.19}$$

where $\mathcal{P}^{-1} : \mathrm{Tan}_{\mu_1}(\mathscr{P}_2(\mathbb{R}^d)) \to \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^d))$ is defined by

$$\mathcal{P}^{-1}(u) := \mathscr{P}_{t_1}^0(\mathscr{P}_{t_2}^{t_1}(\cdots \mathscr{P}_1^{t_n-1}(u)))$$

for the partition $\mathcal{P} = \{0 = t_0 < t_1 < \cdots < t_n = 1\}$ (and, in particular, it is *not* the functional inverse of $u \to \mathcal{P}(u)$). Observe that for any $u \in \mathrm{Tan}_{\mu_{t_i}}(\mathscr{P}_2(\mathbb{R}^d))$ the identities $u = \mathscr{P}_{t_{i+1}}^{t_i}(\tau_{t_i}^{t_{i+1}}(u))$ and $\mathscr{P}_{t_i}^{t_{i+1}}(u) - \tau_{t_i}^{t_{i+1}}(u) \in \mathrm{Tan}^{\perp}_{\mu_{t_{i+1}}}(\mathscr{P}_2(\mathbb{R}^d))$, in conjunction with inequalities (5.12), yield

$$\begin{aligned}\|\mathscr{P}_{t_{i+1}}^{t_i}(\mathscr{P}_{t_i}^{t_{i+1}}(u)) - u\|_{\mu_{t_i}} =& \|\mathscr{P}_{t_{i+1}}^{t_i}(\mathscr{P}_{t_i}^{t_{i+1}}(u) - \tau_{t_i}^{t_{i+1}}(u))\|_{\mu_{t_i}} \\ \leq& \|\mathscr{P}_{t_i}^{t_{i+1}}(u) - \tau_{t_i}^{t_{i+1}}(u)\|_{\mu_{t_i}} D(t_i, t_{i+1}) \\ \leq& \|u\|_{\mu_{t_i}} D^2(t_i, t_{i+1}).\end{aligned}$$

For any $u \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^d))$ we obtain

$$\begin{aligned}&\|u - \mathscr{P}_{t_1}^0(\cdots(\mathscr{P}_1^{t_{n-1}}(\mathcal{P}(u)))\|_{\mu_0} \\ &\qquad \leq \|u - \mathscr{P}_{t_1}^0(\mathscr{P}_0^{t_1}(u))\|_{\mu_0} + \|\mathscr{P}_{t_1}^0(\mathscr{P}_0^{t_1}(u)) - \mathscr{P}_{t_1}^0(\cdots(\mathscr{P}_1^{t_{n-1}}(\mathcal{P}(u))))\|_{\mu_0} \\ &\qquad \leq \|u\|_{\mu_0} D^2(0, t_1) + \|v - \mathscr{P}_{t_2}^{t_1}(\cdots(\mathscr{P}_{t_{n-1}}^1(\mathcal{P}'(v))))\|_{\mu_{t_1}},\end{aligned}$$

where $v = \mathscr{P}_0^{t_1}(u)$ and $\mathcal{P}' = \{t_1 < \cdots < t_n\}$ (so that $\mathcal{P}'(v) = \mathcal{P}(u)$). Since $\|v\|_{\mu_{t_1}} \leq \|u\|_{\mu_0}$ we can continue in this way, to arrive at

$$\|u - \mathscr{P}_{t_1}^0(\cdots(\mathscr{P}_1^{t_{n-1}}(\mathcal{P}(u))))\|_{\mu_0} \leq \|u\|_{\mu_0} D^2(\mathcal{P})$$

and this, taking (5.13d) into account, leads to (5.19). ∎

THEOREM 5.14 (Existence of the parallel transport). *Let μ_t be a regular curve, let $u_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^d))$ and let $\mathcal{T}_t^s$ be the maps defined as above. Then the vector field $u_t := \mathcal{T}_0^t(u_0)$ is the parallel transport of u_0 along the curve.*

Proof. Consider any interval $[t, s] \subset [0, 1]$, its trivial partition $\mathcal{P} = \{t, s\}$ and any (finer) partition Q. Applying inequality (5.16) and passing to the limit on $\mathcal{Q}$ we get

$$\|\mathscr{P}_t^s(u) - \mathcal{T}_t^s(u)\|_{\mu_s} \leq \|u\|_{\mu_t} D^2(t, s) \qquad \forall u \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d)). \tag{5.20}$$

Coupling this equation with inequality (5.12b) we get

$$\begin{aligned}\|\tau_t^s(u) - \mathcal{T}_t^s(u)\|_{\mu_s} \leq& \|\tau_t^s(u) - \mathscr{P}_t^s(u)\|_{\mu_s} + \|\mathscr{P}_t^s(u) - \mathcal{T}_t^s(u)\|_{\mu_s} \\ \leq& \|u\|_{\mu_t} D(t, s)\,(1 + D(0, 1)) \qquad\qquad \forall u \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d)),\end{aligned}$$

which gives the absolute continuity of $t \mapsto \mathcal{T}_0^t(u_0)$.

Now, pick a Lebesgue point t of the function $t \mapsto \mathrm{Lip}(v_t)$ and observe that inequality (5.20) gives

$$\lim_{s \to t} \frac{\|\mathscr{P}_t^s(u) - \mathcal{T}_t^s(u)\|_{\mu_s}}{|s - t|} = 0 \qquad \text{for all } u \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d)).$$

In particular, choosing $u = \mathcal{T}_0^t(u_0)$, we obtain $\|\mathscr{P}_t^s(u) - u_s\|_{\mu_s} = o(s - t)$. Therefore, to conclude it is sufficient to prove that

$$\lim_{s \to t} P_{\mu_t}\left(\frac{\tau_s^t(\mathscr{P}_t^s(u)) - u}{s - t}\right) = 0 \qquad \forall u \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d)).$$

Observe that $\mathscr{P}_t^s(u) - \tau_t^s(u) \in \mathrm{Tan}^{\perp}_{\mu_s}(\mathscr{P}_2(\mathbb{R}^d))$, therefore from inequalities (5.12) we get

$$\begin{aligned}\|P_{\mu_t}(\tau_s^t(\mathscr{P}_t^s(u)) - u)\|_{\mu_t} &= \|\mathscr{P}_s^t(\mathscr{P}_t^s(u) - \tau_t^s(u))\|_{\mu_t} \\ &\le \|\mathscr{P}_t^s(u) - \tau_t^s(u)\|_{\mu_t} D(t,s) \le \|u\|_{\mu_t} D^2(t,s).\end{aligned}$$

□

REMARK 5.15 (Parallel transport along a flow). *Observe that the approximation argument presented in Section 5 to build the parallel transport, works as well if, instead of assuming that the curve μ_t is regular, we assume the existence of a family of maps $X(t,s,x)$ having the group property $X(t,s,X(r,t,x)) = X(r,s,x)$, satisfying $X(t,s,\cdot)_\#\mu_t = \mu_s$ and such that the Lipschitz constant of $X(t,s,\cdot) - Id$ is bounded by a function $D(t,s)$ having the properties* (5.13)*. In other words, we drop the requirement that X is the flow of the* tangent *vector field, but just choose a vector field $\tilde v_t$ with $\int_0^1 \mathrm{Lip}(\tilde v_t)\,dt < \infty$ for which the continuity equation holds. Recall also that*

$$-\nabla\cdot(v_t\mu_t) = \frac{d}{dt}\mu_t = -\nabla\cdot(\tilde v_t\mu_t) \qquad \text{in the sense of distributions in } (0,1)\times\mathbb{R}^d$$

implies $\nabla\cdot((\tilde v_t - v_t)\mu_t) = 0$ for a.e. t, i.e. $P_{\mu_t}(\tilde v_t) = v_t$ for a.e. $t \in (0,1)$.

Using X we would obtain tangent fields $\tilde u_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$, which we call parallel transport along the flow X*, such that $t \mapsto u_t \circ X(s,t,\cdot) \in L^2_{\mu_s}$ is absolutely continuous for every $s \in [0,1]$ and satisfying*

$$P_{\mu_t}\left(\lim_{h\to 0}\frac{u_{t+h}\circ X(t,t+h,\cdot) - u_t}{h}\right) = 0 \qquad \text{for a.e. } t \in (0,1)$$

and

$$\frac{d}{dt}\langle\nabla\eta, u_t\rangle_{\mu_t} = \langle\nabla^2\eta\cdot\tilde v_t, u_t\rangle_{\mu_t} + \langle\nabla\eta, \frac{d}{dt}u_t\rangle_{\mu_t} \qquad \forall\eta\in C_c^\infty(\mathbb{R}^d) \tag{5.21}$$

for a.e. $t \in (0,1)$. However, we shall prove in Section 7 that the choice of the tangent vector field is more natural.

In the rest of the section we analyze some simple examples of parallel transport.

EXAMPLE 5.16 (Equation in the smooth case). Assume that $u_t(x) = \nabla\varphi_t(x)$ for some smooth functions φ_t smoothly varying in time. Then it is easy to see that equation (5.6) becomes:

$$\nabla\cdot\Big((\partial_t\nabla\varphi_t + \nabla^2\varphi_t\cdot v_t)\mu_t\Big) = 0. \tag{5.22}$$

Equivalently

$$\int_0^1\int\langle\partial_t\nabla\varphi_t + \nabla^2\varphi_t\cdot v_t, \nabla\eta\rangle d\mu_t dt = 0 \qquad \forall\eta\in C_c^\infty((0,1)\times\mathbb{R}^d). \tag{5.23}$$

EXAMPLE 5.17 (Constant vector fields). Let μ_t be a regular curve, let $v \in \mathbb{R}^d$ and let C_v be the function on $\mathbb{R}^d$ constantly equal to v. Define $u_0 := C_v \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^d))$. The parallel transport u_t of u_0 along μ_t is given by $u_t = C_v$, for all $t \in [0,1]$. The proof is immediate: it is sufficient to observe that $u_t(x) = \nabla\varphi(t,x)$, where $\varphi(t,x) = \langle x, v\rangle$ and to verify that φ satisfies equation (5.22).

EXAMPLE 5.18 (Geodesics). Consider a geodesic μ_t defined on the interval $[0,1]$: we want to prove that in any interval of the form $[\varepsilon, 1-\varepsilon]$ with $\varepsilon > 0$ the geodesic is regular. Fix $t \in (0,1)$; it is well-known (see [1, 7.2.1] for instance) that there exists only one optimal plan between μ_t and μ_1 and that this plan is induced by a Lipschitz map T_t with Lipschitz constant bounded by t^{-1}. We know also [1, 7.2.2] that for $s \in [t,1]$ it holds $\mu_s = (Id + \frac{s-t}{1-t}(T_t - Id))_\# \mu_t$, the transport map T_t being optimal. Computing the velocity vector v_t as limit of the optimal transport maps as in (2.5), we get $v_t = (1-t)^{-1}(T_t - Id)$, therefore its Lipschitz constant is bounded by $(1+t)(t(1-t))^{-1}$. Our claim on the regularity of μ_t in all intervals $[\varepsilon, 1-\varepsilon]$ follows.

Now assume that the geodesic $[0,1] \ni t \mapsto \mu_t$ is regular. Such a μ_t must be induced by a Lipschitz optimal map T. In this case its flow is given by

$$\mathbf{T}(t,s,\cdot) = (Id + s(T(\cdot) - Id)) \circ (Id + t(T(\cdot) - Id))^{-1},$$

and the velocity vectors satisfy

$$\begin{aligned} v_0 &= T - Id, \\ v_s &= v_t \circ T(t,s,\cdot), \end{aligned}$$

therefore a direct calculation shows that v_t is a parallel transport.

Let us consider now a *locally regular* curve in $(0,T]$, i.e. a curve μ_t such that the function $\mathrm{Lip}(v_t)$ belongs to $L^1_{\mathrm{loc}}((0,T])$: for instance, this is the case of constant speed geodesics in $[0,1]$, that are locally regular in $(0,T]$ for all $T \in (0,1)$. In the following proposition we show how existence of the "forward" parallel transport can still be achieved along locally regular curves.

PROPOSITION 5.19 (Forward parallel transport along locally regular curves). *Let μ_t be an absolutely continuous curve in $[0,T]$, locally regular in $(0,T]$, and let $u_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^d))$. Then the parallel transport of u_0 along μ_t exists, i.e. a locally absolutely continuous vector field $u_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ in $(0,T]$ which is a parallel transport in $(0,T]$ and satisfies $u_t \mu_t \rightharpoonup u_0 \mu_0$ as $t \downarrow 0$ and $\|u_t\|_{\mu_t} = \|u_0\|_{\mu_0}$.*

Proof. We will use the inequality

$$\|\mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi\|_{\mu_s} \le \mathrm{Lip}(\nabla\varphi)\mathcal{L}_t^s(\mu_r), \tag{5.24}$$

where $\mathcal{L}_t^s(\mu_r)$ is the length of μ_r restricted to the interval $[t,s]$. This inequality is a consequence of the fact that $s \mapsto \mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi \in \mathrm{Tan}_{\mu_s}(\mathscr{P}_2(\mathbb{R}^d))$ is an absolutely continuous vector field in $[t,T]$ and of the differential inequality

$$\begin{aligned} \frac{d}{ds}\|\mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi\|_{\mu_s}^2 &= 2\langle \mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi, \frac{d}{ds}\big(\mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi\big)\rangle_{\mu_s} \\ &= 2\langle \mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi, \frac{d}{ds}\big(\mathcal{T}_t^s(\nabla\varphi)\big) - \nabla^2\varphi \cdot v_s\rangle_{\mu_s} \\ &= 2\langle \mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi, P_{\mu_s}\big(\frac{d}{ds}(\mathcal{T}_t^s(\nabla\varphi)) - \nabla^2\varphi \cdot v_s\big)\rangle_{\mu_s} \\ &= -2\langle \mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi, P_{\mu_s}(\nabla^2\varphi \cdot v_s)\rangle_{\mu_t} \\ &= -2\langle \mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi, \nabla^2\varphi \cdot v_s\rangle_{\mu_t} \\ &\le 2\|\mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi\|_{\mu_s}\,\mathrm{Lip}(\nabla\varphi)\|v_s\|_{\mu_s}. \end{aligned}$$

Start assuming that u_0 is the gradient of $\varphi \in C_c^\infty(\mathbb{R}^d)$. Fix $\varepsilon > 0$, think $\nabla\varphi$ as a vector in $\mathrm{Tan}_{\mu_\varepsilon}(\mathscr{P}_2(\mathbb{R}^d))$ and define the vectors $u_t^\varepsilon := \mathcal{T}_\varepsilon^t(\nabla\varphi)$ for any $t \in [\varepsilon, T]$, so that we have $u_\varepsilon^\varepsilon = \nabla\varphi$. From

$$\|u_t^{\varepsilon'} - u_t^\varepsilon\|_{\mu_t} = \|u_\varepsilon^{\varepsilon'} - u_\varepsilon^\varepsilon\|_{\mu_\varepsilon} = \|\mathcal{T}_{\varepsilon'}^\varepsilon(\nabla\varphi) - \nabla\varphi\|_{\mu_\varepsilon} \le \mathrm{Lip}(\nabla\varphi)\omega(\varepsilon) \quad 0 < \varepsilon' \le \varepsilon \le t \le T,$$

with $\omega(\varepsilon) := \int_0^\varepsilon \|v_t\|_{\mu_t} dt$, we get that for any t, the family $\{u_t^\varepsilon\}$ converges in $\mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$, as $\varepsilon \to 0$, to a vector u_t satisfying $\|u_t^\varepsilon - u_t\|_{\mu_t} \le \mathrm{Lip}(\nabla\varphi)\omega(\varepsilon)$. The limit vector field u_t is easily seen to be a parallel transport in the interval $(0, T]$: indeed from the uniform bound (5.8) we get its local absolute continuity, and we conclude by the stability of the solutions of (5.9).

From

$$\|u_t\|_{\mu_t} = \lim_\varepsilon \|u_t^\varepsilon\|_{\mu_t} = \lim_\varepsilon \|u_\varepsilon^\varepsilon\|_{\mu_\varepsilon} = \lim_\varepsilon \|\nabla\varphi\|_{\mu_\varepsilon}$$

we get that the norm of u_t is constant, and equal to $\|\nabla\varphi\|_{\mu_0}$. Finally it holds

$$\langle u_\varepsilon, \eta\rangle_{\mu_\varepsilon} = \langle u_\varepsilon - u_\varepsilon^\varepsilon, \eta\rangle_{\mu_\varepsilon} + \langle u_\varepsilon^\varepsilon, \eta\rangle_{\mu_\varepsilon} = R_\varepsilon + \langle \nabla\varphi, \eta\rangle_{\mu_\varepsilon} \qquad \forall \eta \in C_c^\infty(\mathbb{R}^d, \mathbb{R}^d),$$

where the term R_ε is bounded by $\|u_t - u_t^\varepsilon\|_{\mu_t} \sup|\eta| \le \omega(\varepsilon)\,\mathrm{Lip}(\nabla\varphi)\sup|\eta|$.

For the general case, just approximate u_0 with smooth gradients u_0^n, apply the construction above to obtain the existence of forward parallel transports u_t^n of u_0^n and use the fact that (clearly) $\|u_t^n - u_t^m\|_{\mu_t} = \|u_0^n - u_0^m\|_{\mu_0}$ to get that for any t the sequence (u_t^n) strongly converges to some u_t satisfying $\|u_t\|_{\mu_t} = \|u_0\|_{\mu_0}$. By the stability argument used above we get that u_t is a parallel transport on $(0, T]$, so we need just to prove that u_t weakly converge to u_0 as $t \to 0$. To prove this, observe that since $[0, T] \ni t \mapsto u_t^n$ is a forward parallel transport, passing to the limit as $t \to 0$ in (5.9) we get

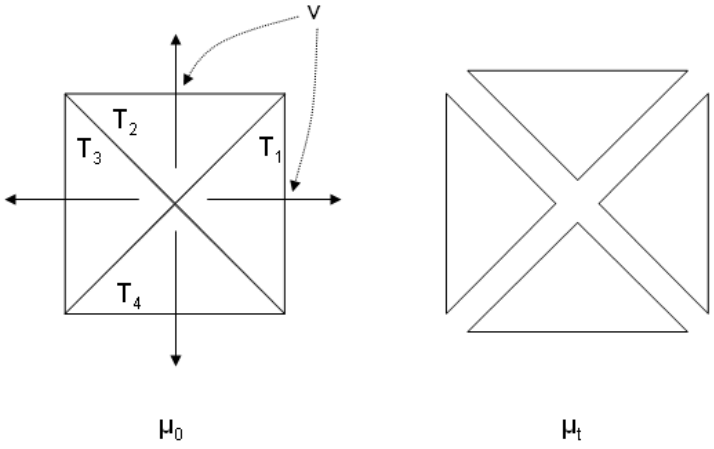
$$\langle u_t^n, \nabla\eta\rangle_{\mu_t} - \langle u_0^n, \nabla\eta\rangle_{\mu_0} = \int_0^t \langle u_r^n, \nabla^2\eta\cdot v_r\rangle_{\mu_r} dr \le \|u_t^n\|_{\mu_t}\,\mathrm{Lip}(\nabla\eta)\omega(t) \quad \forall \eta \in C_c^\infty(\mathbb{R}^d).$$

Letting $n \to \infty$ in the above inequality the weak convergence follows. □

Now we give an example of a geodesic along which a parallel transport does not exist globally: we proved in Proposition 5.19 that the forward parallel transport exists in $[0, 1)$ for all constant speed geodesics $\mu_t : [0, 1] \to \mathscr{P}_2(\mathbb{R}^d)$. We will see, on the other hand, that for $T \in (0, 1)$ the backward problem of a transporting $u_T \in \mathrm{Tan}_{\mu_T}(\mathscr{P}_2(\mathbb{R}^d))$ to some $u_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^d))$ does not have solution in general. The obstruction to the construction we made of the parallel transport is the fact that $\mathrm{supp}(\mu_0)$ is not necessarily homeomorphic to $\mathrm{supp}(\mu_t)$ for $t > 0$. This change of topology cannot happen along a regular curve: indeed, the flow maps $\mathbf{T}(t, s, \cdot)$ are (actually, can be extended to) bi-Lipschitz homeomorphisms of $\mathbb{R}^d$ into itself. Therefore, since $\mathrm{supp}(\mu_s) = \mathbf{T}(t, s, \mathrm{supp}(\mu_t))$, the supports of μ_t, as t varies, are all homeomorphic. We will see that, in this situation, not only the parallel transport can't be built with flow maps, but also that it may happen that the parallel transport does not exist.

Example 5.20. *Let $Q = [0, 1] \times [0, 1]$ be the unit square in $\mathbb{R}^2$ and let T_i, $i = 1, 2, 3, 4$, be the four open triangles in which Q is divided by its diagonals. Let $\mu_0 := \chi_Q \mathscr{L}^2$ and define the function $v : Q \to \mathbb{R}^2$ as the gradient of the convex map $\max\{|x|, |y|\}$, as in the figure. Set also $w = v^\perp$, the rotation by $\pi/2$ of v, in Q and $w = 0$ out of Q. Notice that w is a divergence-free vector field in the whole of $\mathbb{R}^2$.*

Set $\mu_t := (Id + tv)_{\#}\mu_0$ and observe that, for positive t, the support Q_t of μ_t is made of 4 connected components, each one the translation of one of the sets T_i, and that $\mu_t = \chi_{Q_t}\mathscr{L}^2$.



It is immediate to check that μ_t is a geodesic in $[0,\infty)$, so that from Example 5.18 we know that the restriction of μ_t to any interval $[\varepsilon,1]$ with $\varepsilon > 0$ is regular. Fix $\varepsilon > 0$ and note that, by construction, the flow maps of μ_t in $[\varepsilon,1]$ are given by

$$\mathbf{T}(t,s,\cdot) = (Id + sv) \circ (Id + tv)^{-1}, \quad \forall t,s \in [\varepsilon,1].$$

Now, set $w_t := w \circ \mathbf{T}(t,0,\cdot)$ and notice that w_t is tangent at μ_t, because w_t is constant in the connected components of the support of μ_t. Since $w_{t+h} \circ \mathbf{T}(t,t+h,\cdot) = w_t$, from (5.7) we obtain that w_t is a parallel transport in $[\varepsilon,1]$. Furthermore, since $\nabla \cdot (w\mu_0) = 0$, we have $w_0 = w \notin \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(\mathbb{R}^2))$. Therefore there is no way to extend w_t to a continuous tangent *vector field on the whole $[0,1]$.*

6. Density of regular curves. Aim of this section is the proof of a density result for regular curves. It is well-known that the set

$$\mathscr{P}_2^a(\mathbb{R}^d) := \left\{\mu \in \mathscr{P}_2(\mathbb{R}^d) \,:\, \mu \ll \mathscr{L}^d\right\}$$

is a *geodesic* subspace of $\mathscr{P}_2(\mathbb{R}^d)$ (i.e. any geodesic between two points in $\mathscr{P}_2^a(\mathbb{R}^d)$ is entirely contained in $\mathscr{P}_2^a(\mathbb{R}^d)$) and the same is true for the subsets $\{\mu = \rho\mathscr{L}^d \,:\, \|\rho\|_\infty \le C\}$. Our approximation will be obtained with measures in this class, and preserves these upper bounds on the densities, if any.

The delicate point in our approximation result is due to the fact that regularity imposes a Lipschitz condition on the *tangent* velocity field. The typical approximation schemes for solutions to the continuity equation, on the other hand, produce a regularized vector field that is compatible with the regularized density, but it is not tangent in general. Therefore a further projection of the regularized velocity on the tangent space is needed.

The following lemma will be used in the reduction to compactly supported measures.

LEMMA 6.1 (Monotone approximation). *Let $\mu_t : [0,1] \to \mathscr{P}_2(\mathbb{R}^d)$ be absolutely continuous and let v_t be its tangent velocity field. Then there exist absolutely continuous curves $\mu_t^n : [0,1] \to \mathscr{P}_2(\mathbb{R}^d)$ and $z_n \uparrow 1$ satisfying:*

(i) $z_n\mu_t^n \uparrow \mu_t$ *for all* $t \in [0,1]$ *and* $\sup_t W_2(\mu_t^n, \mu_t) \to 0$ *as* $n \to \infty$*;*
(ii) the tangent velocity field of μ_t^n *is* v_t*, and there exists a closed ball* B_n *such that* $\operatorname{supp}\mu_t^n \subset B_n$ *for all* $t \in [0,1]$*.*

Proof. Let Ω be the Banach space of continuous maps from $[0,1]$ to $\mathbb{R}^d$ and let $e_t : \Omega \to \mathbb{R}^d$ be the evaluation maps at time t, i.e. $e_t(\omega) = \omega(t)$. According to [1, 8.2.1], we can represent μ_t as the law under e_t of a suitable probability measure η in Ω, concentrated in the set of absolutely continuous solutions of the equation $\dot\omega = v_t(\omega)$.

Let

$$\Omega_n := \left\{\omega \in \Omega : \omega \text{ is absolutely continuous, } |\omega(0)| \le n,\ \int_0^1 |\dot\omega| dt \le n\right\}$$

and set $\eta_n = \chi_{\Omega_n}\eta$, $z_n = \eta(\Omega_n) = \eta_n(\Omega)$ and $\mu_t^n = z_n^{-1}(e_t)_\#\eta_n$. It is easy to check condition (i), and that the support of μ_t^n is contained in the ball $\overline{B}_{2n}(0)$. Since also η_n is concentrated on curves solving the ODE $\dot\omega = v_t(\omega)$, it turns out that v_t is an admissible velocity field for μ_t^n (i.e. the continuity equation holds, see again [1, 8.2.1] for instance). We conclude that v_t is the tangent velocity fields noticing that, because of condition (i), $v_t \in \operatorname{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ implies $v_t \in \operatorname{Tan}_{\mu_t^n}(\mathscr{P}_2(\mathbb{R}^d))$. □

We can now state our approximation result.

THEOREM 6.2 (Approximation by regular curves). *Let* $\mu_t : [0,1] \to \mathscr{P}_2(\mathbb{R}^d)$ *be an absolutely continuous curve. Then there exist regular curves* $\mu_t^n : [0,1] \to \mathscr{P}_2(\mathbb{R}^d)$ *satisfying:*

(i) $\sup_t W_2(\mu_t^n, \mu_t) \to 0$ *as* $n \to +\infty$*;*
(ii) $\mu_t^n = \rho_t^n \mathscr{L}^n$*,* $\sup_t \|\rho_t^n\|_\infty < +\infty$*,* ρ_t^n *are smooth, the smooth tangent velocity fields* v_t^n *are gradients of smooth maps* $\varphi_t^n : \mathbb{R}^d \to \mathbb{R}$ *satisfying* $\sup_t \operatorname{Lip}(v_t^n) < \infty$ *and* $\{\rho_t^n > 0\}$ *is a bounded open set with a smooth boundary;*
(iii) if $v_t \in L^2_{\mu_t}$ *is the tangent field of* μ_t*, we have that* $v_t^n\mu_t^n$ *weakly converge to* $v_t\mu_t$ *and*

$$\lim_{n\to\infty} \int |v_t^n|^2 d\mu_t^n = \int |v_t|^2 d\mu_t.$$

Proof. **Step 1.** Regularization of 1-periodic solutions.

By Lemma 6.1 and a diagonal argument we can assume that the supports of μ_t are contained in a fixed compact set. By a scaling argument, we can also assume with no loss of generality that the union of these supports is a compact subset K of $(0,1)^d$. We consider the 1-periodic extension $\mu_t^{\rm per}$ of μ_t, still solving the continuity equation with the 1-periodic extension $v_t^{\rm per}$ of v_t, and the regularized densities

$$\varrho_t^n := \mu_t^{\rm per} * \chi_n,$$

still 1-periodic. Here χ_n is a family of smooth and symmetric w.r.t. 0, convolution kernels converging to δ_0 whose support has a diameter *equal* to $2\sqrt{d}$. With this choice of χ_n, we have $\inf_t \inf \varrho_t^n > 0$, and standard properties of convolution yield

$$\sup \varrho_t^n \le \sup_t \mu_t^{\rm per}([-1,2]^d) \sup \chi_n = 3^d \sup \chi_n.$$

Analogous bounds hold, of course, for all higher order derivatives of ϱ_t^n. Passing to the velocity fields, we consider as in [7, 1] this regularization:

$$w_t^n := \frac{(v_t^{\rm per}\mu_t^{\rm per}) * \chi_n}{\varrho_t^n}$$

which satisfies, thanks to the lower bound on ϱ_t^n, $\sup_t \mathrm{Lip}(w_t^n) < \infty$ (and the same holds for higher order derivatives) and preserves the validity of the continuity equation. Eventually we consider the projection $v_t^n = \nabla\varphi_t^n$ of w_t^n on periodic gradients by solving the PDE

$$\nabla \cdot (\nabla\varphi_t^n \varrho_t^n) = \nabla \cdot (w_t^n \varrho_t^n).$$

From the variational formulation of the PDE we obtain

$$\int_{(0,1)^d} |v_n^t|^2 \varrho_t^n \, dx \le \int_{(0,1)^d} |w_n|^2 \varrho_t^n dx. \tag{6.1}$$

We can use standard elliptic regularity theory to obtain that $\sup_t \mathrm{Lip}(v_t^n) < \infty$. Moreover, using Jensen's inequality as in [1, 8.1.10], we have the local estimate

$$\int_{(0,1)^d} |w_t^n|^2 \varrho_t^n dx \le \int \int_{(0,1)^d} |v_t|^2(y) \chi_n(x-y) dx d\mu_t^{\mathrm{per}}(y). \tag{6.2}$$

Step 2. Construction of the approximating sequence.

We build $\mu_t^n = \rho_t^n \mathscr{L}^d \in \mathscr{P}_2^a(\mathbb{R}^d)$, with the same velocity field v_t^n, from the periodic measures $\varrho_t^n \mathscr{L}^d$. To this aim, we shall first consider ϱ_t^n as measures in the flat d-dimensional torus $\mathbb{T}^d \sim (0,1)^d$, with velocity field v_t^n. We denote by $\mathbb{P}$ the Lebesgue measure on $\mathbb{T}^d$, by $\mathbf{X}^n(t,x) : [0,1] \times \mathbb{T}^d \to \mathbb{T}^d$ the smooth flow of v_t^n (starting from $s = 0$), and by η_n the probability measure in $C([0,1];\mathbb{T}^d)$ defined by

$$\eta_n := \mathbf{X}^n(t,\cdot)_\#(\varrho_0^n \mathbb{P}).$$

Equivalently, η_n is the law of the random variable $x \mapsto \mathbf{X}^n(\cdot,x) \in C([0,1];\mathbb{T}^d)$ under $\varrho_0^n\mathbb{P}$. Classical representation results for solutions to the continuity equation with a Lipschitz vector field ensure that $\varrho_t^n\mathbb{P} = \mathbf{X}^n(t,\cdot)_\#(\varrho_0^n\mathbb{P})$, and since $e_t \circ \mathbf{X}^n(\cdot,x) = \mathbf{X}^n(t,x)$ we obtain

$$(e_t)_\#\eta_n = \varrho_t^n \mathbb{P} \qquad \forall t \in [0,1],\ n \in \mathbb{N}. \tag{6.3}$$

From (6.1) and (6.2) we get

$$\sup_n \int_0^1 \int_{\mathbb{T}^d} |v_t^n|^2 \varrho_t^n d\mathbb{P}(x) dt < \infty$$

and this, using Prokhorov theorem as in [2, 11], gives that (η_n) is a relatively compact sequence in $\mathscr{P}(C([0,1];\mathbb{T}^d))$. It is not restrictive, extracting if necessary a subsequence, to assume that (η_n) weakly converges, in the duality with continuous and bounded functions in $C([0,1];\mathbb{T}^d)$, to some probability measure η. Passing to the limit as $n \to \infty$ in (6.3) we obtain that $(e_t)_\#\eta = \mu_t^{\mathrm{per}}$ for all $t \in [0,1]$, and this means that η-almost all the paths ω are contained in $\tilde{K}$ (here we denote by $\tilde{K}$ the image of K in $\mathbb{T}^d$ and we consider μ_t^{per} as probability measures in $\mathbb{T}^d$).

Now, let $\delta < 1$ be such that K is contained in the interior of $[\delta, 1-\delta]^d$ and define $\tilde{\eta}_n := z_n^{-1} \chi_{\Omega(\delta)} \eta_n$, where

$$\Omega(\delta) := \left\{\omega \in C([0,1];\mathbb{T}^d) :\ \omega(t) \,\mathrm{mod}(1) \in (\delta, 1-\delta)^d \ \forall t \in [0,1]\right\}, \qquad z_n := \eta_n(\Omega(\delta))$$

(in other words, we remove the trajectories that cross $\partial(\delta, 1-\delta)^d$). Since η is supported on paths contained in $\tilde{K}$, we have that $z_n \to 1$ and still $\tilde{\eta}_n$ weakly converge to $\chi_{\Omega(\delta)}\eta = \eta$. We define

$$\mu_t^n := (\tilde{e}_t)_\#\tilde{\eta}_n$$

where $\tilde e_t(\omega(t)) = \omega_t \,\mathrm{mod}(1) \in [0,1)^d$. The measures μ_t^n can also be represented by

$$\mu_t^n = z_n^{-1} \mathbf{Y}^n(t,\cdot)_\#(\chi_{E^n(\delta)} \varrho_0^n \mathscr{L}^d), \tag{6.4}$$

where $\mathbf{Y}^n(t,x) = \mathbf{X}^n(t,x) \,\mathrm{mod}(1)$ and $E^n(\delta) = \{x \in (0,1)^d : \mathbf{X}^n(\cdot,x) \in \Omega(\delta)\}$.

By construction μ_t^n are probability measures in $\mathbb{R}^d$ concentrated on $[\delta, 1-\delta]^d$. It is immediate to check that the tangent field to μ_t^n is v_t^n (because η_n is concentrated on solutions to the ODE $\dot\omega = v_t^n(\omega)$ in $\mathbb{T}^d$, and v_t^n are gradients). In particular μ_t^n are regular curves and the convergence of μ_t^n to μ_t follows at once from the convergence of $\tilde\eta_n$ to η, using the evaluation map $\tilde e_t$. Notice also that the inequality $z_n\tilde\eta_n \le \eta_n$ and the fact that the mass of their difference is infinitesimal imply

$$z_n \mu_t^n \le \varrho_t^n \mathscr{L}^d \quad \text{and} \quad \lim_{n\to\infty} (\varrho_t^n \mathscr{L}^d - z_n\mu_t^n)((0,1)^d) = 0. \tag{6.5}$$

Step 3. Convergence of velocity fields.

Notice first that (6.1) and (6.2) give

$$\begin{aligned}
\limsup_{n\to\infty} \int |v_t^n|^2 d\mu_t^n &\le \limsup_{n\to\infty} \int |v_t^n(e_t(\omega))|^2 d\eta_n(\omega) = \limsup_{n\to\infty} \int |v_t^n|^2 \varrho_t^n d\mathbb{P} \\
&\le \limsup_{n\to\infty} \int |w_t^n|^2 \varrho_t^n d\mathbb{P} \\
&\le \limsup_{n\to\infty} \int \int_{(0,1)^d} |v_t|^2(y) \chi_n(x-y) dx d\mu_t^{\mathrm{per}}(y) = \int |v_t|^2 d\mu_t.
\end{aligned}$$

Now, recall (see for instance [1, 9.4.3]) that the functional

$$G(\nu,\mu) := \begin{cases} \displaystyle\int |f|^2 d\mu & \text{if } \nu = f\mu \text{ with } f \in L^2(\mu;\mathbb{R}^d), \\ +\infty & \text{otherwise} \end{cases}$$

is jointly lower semicontinuous in $\mathscr{P}_2(\mathbb{R}^d)\times\mathscr{P}_2(\mathbb{R}^d)$ with respect to weak convergence in the duality with $C_b(\mathbb{R}^d)$, to obtain that any weak limit point σ of $v_t^n\mu_t^n$ as $n\to\infty$ has the form $\tilde v\mu_t$ for some $\tilde v \in L^2_{\mu_t}$ with $\|\tilde v\|_{\mu_t} \le \|v_t\|_{\mu_t}$. On the other hand, passing to the limit in $\nabla\cdot((w_t^n - v_t^n)\varrho_t^n) = 0$ and taking into account the weak convergence in the duality with $C_c(\mathbb{R}^d)$ of $w_t^n\varrho_t^n = (v_t^{\mathrm{per}}\mu_t^{\mathrm{per}}) * \chi_n$ to $v_t^{\mathrm{per}}\mu_t^{\mathrm{per}}$ and the convergence to 0 in $(0,1)^d$ of $\varrho_t^n\mathscr{L}^d - \mu_t^n$ (ensured, even in the strong sense, by (6.5)) we get $\nabla\cdot((\tilde v - v_t)\mu_t) = 0$. Since v_t is tangent and $\|\tilde v\|_{\mu_t} \le \|v_t\|_{\mu_t}$, it must be $\tilde v = v_t$. This proves the weak convergence of velocity fields that provides also, thanks to the lower semicontinuity of G, the lim inf inequality

$$\liminf_{n\to\infty} \int |v_t^n|^2 d\mu_t^n \ge \int |v_t|^2 d\mu_t.$$

Step 4. Eventually we can regularize the characteristic function of the set $E^n(\delta)$ in (6.4), by smooth functions χ^n such that $\{\chi_n\varrho_0^n > 0\}$ is smooth and bounded, to approximate the curve μ_t^n by curves with the same velocity field and smooth densities with respect to $\mathscr{L}^d$, with smooth supports. □

7. Additional remarks and extensions. In this section we describe in a more informal way, referring to [8] for more details, some additional construction made possible by the existence of the parallel transport along a sufficiently large class of curves.

7.1. Covariant derivative and curvature operator. It is well known that, in the classical Riemannian setting, the definition of parallel transport leads to the one of covariant derivative via the formula

$$\nabla_{\dot\gamma(t)}u(t) := \lim_{s\to t}\frac{T_s^t(u(s))-u(t)}{s-t}, \tag{7.1}$$

where $\gamma(t)$ is a smooth curve, $u(t)\in T_{\gamma(t)}M$ is a smooth vector field and T_s^t, for any s, t, is the parallel transport map from $T_{\gamma(s)}M$ to $T_{\gamma(t)}M$ along γ.

The same construction may be used in the Wasserstein setting:

DEFINITION 7.1 (Covariant derivative). *Let μ_t be a regular curve, let $v_t\in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ be its velocity vector and let $u_t\in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ be an absolutely continuous vector field along μ_t. The covariant derivative of u_t along v_t is:*

$$\nabla_{v_t}u_t := \lim_{s\to t}\frac{\mathcal{T}_s^t(u_s)-u_t}{s-t},$$

where $\mathcal{T}_s^t$ are the parallel transport maps along μ_t and the derivative takes place in $L^2_{\mu_t}$.

Using the definition of absolutely continuous vector field, it is not difficult to check that the covariant derivative exists for a.e. t and that the function $t\mapsto\|\nabla_{v_t}u_t\|_{\mu_t}$ is integrable. Indeed, inequality (5.20) implies that the covariant derivative satisfies:

$$\nabla_{v_t}u_t = \lim_{s\to t}\frac{\mathscr{P}_s^t(u_s)-u_t}{s-t} = P_{\mu_t}\left(\lim_{s\to t}\frac{u_s\circ\mathbf{T}(t,s,\cdot)-u_t}{s-t}\right). \tag{7.2}$$

If the vector field u_t is given by the gradient of smooth functions, i.e. if $u_t=\nabla_x\varphi_t(x)$ for some $\varphi_t\in C_c^\infty(\mathbb{R}^d)$ smoothly varying in time, the previous equation reads as

$$\nabla_{v_t}u_t = P_{\mu_t}\big(\partial_t\nabla\varphi_t+\nabla_x^2\varphi_t\cdot v_t\big). \tag{7.3}$$

Equation (7.3) and the analogous one (5.22) were first given in [10], although from a formal viewpoint and under stronger assumptions on the measures μ_t.

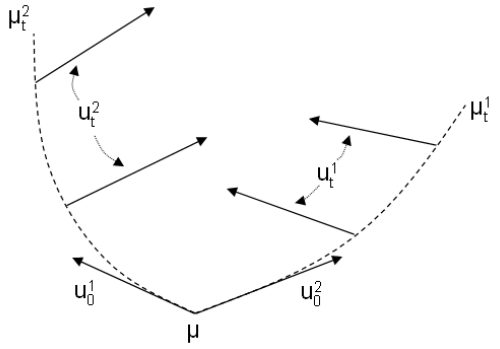
Having defined the covariant derivative, our first goal is to prove that it is the Levi-Civita connection on $(\mathscr{P}_2(\mathbb{R}^d),W_2)$. Recalling the discussion made for the classical case of Riemanniann manifolds, we need to prove that it is *compatible with the metric* and *torsion-free*. The compatibility with the metric is a simple consequence of the definition: indeed, for a given couple of absolutely continuous vector fields u_t^1, $u_t^2\in\mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ along the regular curve μ_t, we have:

$$\begin{aligned}\frac{d}{dt}\langle u_t^1,u_t^2\rangle_{\mu_t} &= \langle\frac{d}{dt}u_t^1,u_t^2\rangle_{\mu_t}+\langle u_t^1,\frac{d}{dt}u_t^2\rangle_{\mu_t}\\ &= \langle P_{\mu_t}\left(\frac{d}{dt}u_t^1\right),u_t^2\rangle_{\mu_t}+\langle u_t^1,P_{\mu_t}\left(\frac{d}{dt}u_t^2\right)\rangle_{\mu_t}\\ &= \langle\nabla_{v_t}u_t^1,u_t^2\rangle_{\mu_t}+\langle u_t^1,\nabla_{v_t}u_t^2\rangle_{\mu_t},\end{aligned} \tag{7.4}$$

having used the Leibnitz rule (5.4) and the fact that both vector fields are tangent.

To prove the torsion-free identity, we need first to understand how to calculate the Lie bracket of two vector fields. To this aim, let μ_t^i, $i=1,2$, be two regular curves such that $\mu_0^1=\mu_0^2=:\mu$ and let $u_t^i\in\mathrm{Tan}_{\mu_t^i}(\mathscr{P}_2(\mathbb{R}^d))$ be two absolutely continuous

vector fields satisfying $u_0^1 = v_0^2$, $u_0^2 = v_0^1$, where v_t^i are the tangent fields of μ_t^i. We assume that the velocity fields v_t^i of μ_t^i are continuous in time (for instance with respect to the convergence considered in Section 7.2), to be sure that (5.5) holds for *all* t with $u = u^i$ and the initial condition makes sense.



Let us consider vector fields as derivations, and the functional $\mu \mapsto F_\eta(\mu) := \int \eta d\mu$, for $\eta \in C_c^\infty(\mathbb{R}^d)$ fixed. By the continuity equation, the derivative of F_η along u_t^2 is equal to $\langle \nabla\eta, u_t^2\rangle_{\mu_t^2}$, therefore (5.5) gives:

$$\begin{aligned} u^1(u^2(F_\eta))(\mu) &= \frac{d}{dt}\langle \nabla\eta, u_t^2\rangle_{\mu_t^2}|_{t=0} = \langle \nabla^2\eta \cdot v_0^2, u_0^2\rangle_\mu + \langle \nabla\eta, \frac{d}{dt} u_t^2|_{t=0}\rangle_\mu \\ &= \langle \nabla^2\eta \cdot u_0^1, u_0^2\rangle_\mu + \langle \nabla\eta, \nabla_{v_0^2} u_t^2\rangle_\mu. \end{aligned}$$

Subtracting the analogous term $u^2(u^1(F_\eta))(\mu)$ and using the symmetry of $\nabla^2\eta$ and the identities $u_0^i = v_0^{1-i}$, $i = 0, 1$, we get

$$[u^1, u^2](F_\eta)(\mu) = \langle \nabla\eta, \nabla_{u_0^1} u_t^2 - \nabla_{u_0^2} u_t^1\rangle_\mu.$$

Given that the set $\{\nabla\eta\}_{\eta\in C_c^\infty}$ is dense in $\mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d))$, the above equation characterizes $[u^1, u^2]$ as:

$$[u^1, u^2] = \nabla_{u_0^1} u_t^2 - \nabla_{u_0^2} u_t^1, \tag{7.5}$$

which proves the torsion-free identity for the covariant derivative.

In the case of the parallel transport along a flow, considered in Remark 5.15, given that the right composition with $X(t, s, \cdot)$ is an isometry from $L^2_{\mu_t}$ to $L^2_{\mu_s}$, it holds

$$\langle u_s^1, u_s^2\rangle_{\mu_s} = \langle u_s^1 \circ X(t, s, \cdot), u_s^2 \circ X(t, s, \cdot)\rangle_{\mu_t},$$

subtracting $\langle u_t^1, u_t^2\rangle_{\mu_t}$, dividing both terms by $s - t$ and letting $s \to t$ we get that the Leibnitz rule holds even using the maps $X(t, s, \cdot)$:

$$\frac{d}{dt}\langle u_t^1, u_t^2\rangle_{\mu_t} = \langle \frac{d}{ds} u_s^1 \circ X(t, s, \cdot)_{|s=t}, u_t^2\rangle_{\mu_t} + \langle u_t^1, \frac{d}{ds} u_s^2 \circ X(t, s, \cdot)_{|s=t}\rangle_{\mu_t},$$

for any couple of vector fields u_t^i such that $t \mapsto u_t^i \circ X(s, t, \cdot)$ is absolutely continuous for $i = 1, 2$. From this formula it follows that the parallel transport along *any* flow X compatible with μ_t preserves the scalar product.

Of course, different parallel transports define different covariant derivatives $\tilde\nabla_{v_t}u_t$ via (7.2): they are expressed by

$$\tilde\nabla_{v_t}u_t := P_{\mu_t}\left(\frac{d}{ds}u_s\circ X(t,s,\cdot)_{|_{s=t}}\right).$$

Denoting by $\tilde v_t$ the velocity field of X, we get that the covariant derivative of the vector field $u_t := \nabla\varphi$, $\varphi\in C_c^\infty(\mathbb{R}^d)$, is given by $\tilde\nabla_{v_t}\nabla\varphi = P_{\mu_t}(\nabla^2\varphi\cdot\tilde v_t)$. It is easy to check that a generic covariant derivative is not torsion-free. Indeed, assume that it is and observe that in this case the two following equations hold:

$$\begin{gathered}\langle\nabla_{\nabla\varphi^1}\nabla\varphi^2,\nabla\varphi^3\rangle_\mu+\langle\nabla\varphi^2,\nabla_{\nabla\varphi^1}\nabla\varphi^3\rangle_\mu=\langle\tilde\nabla_{\nabla\varphi^1}\nabla\varphi^2,\nabla\varphi^3\rangle_\mu+\langle\nabla\varphi^2,\tilde\nabla_{\nabla\varphi^1}\nabla\varphi^3\rangle_\mu,\\ \nabla_{\nabla\varphi^1}\nabla\varphi^2-\nabla_{\nabla\varphi^2}\nabla\varphi^1=\tilde\nabla_{\nabla\varphi^1}\nabla\varphi^2-\tilde\nabla_{\nabla\varphi^2}\nabla\varphi^1,\end{gathered}$$

for any $\varphi^i\in C_c^\infty(\mathbb{R}^d)$, $i=1,2,3$. From these equalities, with some algebraic manipulations (more explicitely, by following the calculations indicated in the Koszul formula), it follows that $\langle\nabla_{\nabla\varphi^1}\nabla\varphi^2,\nabla\varphi^3\rangle_\mu=\langle\tilde\nabla_{\nabla\varphi^1}\nabla\varphi^2,\nabla\varphi^3\rangle_\mu$, so that the two covariant derivatives coincide. Furthermore, we can consider in the identity

$$P_{\mu_t}\big(\nabla^2\varphi\cdot(\tilde v_t-v_t)\big)=0\qquad\forall\varphi\in C_c^\infty(\mathbb{R}^d)$$

test functions $\varphi\in C^2(\mathbb{R}^d)$ with uniformly bounded second derivatives (by a simple approximation argument based on the finiteness of the second moments of μ_t). Choosing $\varphi(x)=|\langle x,\xi\rangle|^2$ gives

$$\int\frac{\partial\eta}{\partial\xi}\langle\tilde v_t-v_t,\xi\rangle d\mu_t=0\qquad\forall\eta\in C_c^\infty(\mathbb{R}^d),\ \xi\in\mathbb{R}^d.$$

This means the symmetric part of the distributional derivative of the vector-valued distribution $(\tilde v_t-v_t)\mu_t$ vanishes; Korn's inequality gives that the distribution is equivalent to a constant. By integrability, this constant must be 0, i.e. $\tilde v_t=v_t$ μ_t-a.e. in $\mathbb{R}^d$.

The definition of covariant derivative allows us to define the curvature tensor and to check, at least formally, that $(\mathscr{P}_2(\mathbb{R}^d),W_2)$ is positively curved by proving that its sectional curvatures are always non-negative. The spirit of the foregoing discussion and the calculations we do, are basically borrowed from Lott's work [10].

Given four vector fields $\mu\mapsto\nabla\varphi^i_\mu\in\mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d))$, $i=1,\dots,4$, the curvature tensor $\mathcal{R}$ calculated on them is defined as:

$$\begin{aligned}\langle\mathcal{R}(\nabla\varphi^1_\mu,\nabla\varphi^2_\mu)(\nabla\varphi^3_\mu),\nabla\varphi^4_\mu\rangle_\mu :=\ &\langle\nabla_{\nabla\varphi^1_\mu}(\nabla_{\nabla\varphi^2_\mu}\nabla\varphi^3_\mu),\nabla\varphi^4_\mu\rangle_\mu\\ &-\langle\nabla_{\nabla\varphi^2_\mu}(\nabla_{\nabla\varphi^1_\mu}\nabla\varphi^3_\mu),\nabla\varphi^4_\mu\rangle_\mu\\ &-\langle\nabla_{[\nabla\varphi^1_\mu,\nabla\varphi^2_\mu]}\nabla\varphi^3_\mu,\nabla\varphi^4_\mu\rangle_\mu.\end{aligned}$$

With the same calculation used in the classical Riemannian case, it is easy to check that $\mathcal{R}$ is actually a tensor, i.e. that its value at the measure μ depends only on the value of the four vector fields at μ. Therefore in order to evaluate it, we can consider the simpler vector fields $\mu\mapsto\nabla\varphi^i\in\mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d))$, $i=1,\dots,4$, where the functions φ^i do not depend on the base measure μ. This will simplify the calculations. Under this assumption we have

$$\nabla_v\nabla\varphi=P_\mu(\nabla^2\varphi\cdot v)\qquad\forall v\in\mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d)).\tag{7.6}$$

In order to give an explicit formula for $\mathcal{R}$, it is useful to introduce the function $\xi_\mu(\varphi^1,\varphi^2)\in L^2_\mu$ as

$$\xi_\mu(\varphi^1,\varphi^2):=P^\perp_\mu(\nabla^2\varphi^1\cdot\nabla\varphi^2)=\nabla^2\varphi^1\cdot\nabla\varphi^2-\nabla_{\nabla\varphi^2}\nabla\varphi^1(\mu).$$

Observe that from $\nabla^2\varphi^1\cdot\nabla\varphi^2+\nabla^2\varphi^2\cdot\nabla\varphi^1=\nabla(\langle\nabla\varphi^1,\nabla\varphi^2\rangle)\in\mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d))$ we get $\xi_\mu(\varphi^1,\varphi^2)=-\xi_\mu(\varphi^2,\varphi^1)$.

Proposition 7.2. *The curvature tensor is given by*

$$\begin{aligned}\langle\mathcal{R}(\nabla\varphi^1,\nabla\varphi^2)(\nabla\varphi^3),\nabla\varphi^4\rangle_\mu=&\langle\xi_\mu(\varphi^1,\varphi^4),\xi_\mu(\varphi^2,\varphi^3)\rangle_\mu-\langle\xi_\mu(\varphi^1,\varphi^3),\xi_\mu(\varphi^2,\varphi^4)\rangle_\mu\\&-2\langle\xi_\mu(\varphi^1,\varphi^2),\xi_\mu(\varphi^3,\varphi^4)\rangle_\mu.\end{aligned}$$

Proof. Define $\mu_t:=(Id+t\nabla\varphi^1)_\#\mu$ and $F(\nu):=\int\eta d\nu$ with $\eta:=\langle\nabla^2\varphi^3\cdot\nabla\varphi^2,\nabla\varphi^4\rangle$. Evaluate the derivative at $t=0$ of $F(\mu_t)$ to get

$$\frac{d}{dt}F(\mu_t)_{|t=0}=\frac{d}{dt}\int\eta\circ(Id+t\nabla\varphi^1)d\mu_{|t=0}=\langle\nabla\eta,\nabla\varphi^1\rangle_\mu.$$

On the other hand, using equations (7.6) and (7.4) we have

$$\begin{aligned}\frac{d}{dt}F(\mu_t)_{|t=0}&=\frac{d}{dt}\langle\nabla^2\varphi^3\cdot\nabla\varphi^2,\nabla\varphi^4\rangle_{\mu_t|t=0}\\&=\frac{d}{dt}\langle\nabla_{\nabla\varphi^2}\nabla\varphi^3(\mu_t),\nabla\varphi^4\rangle_{\mu_t|t=0}\\&=\langle\nabla_{\nabla\varphi^1}(\nabla_{\nabla\varphi^2}\nabla\varphi^3),\nabla\varphi^4\rangle_\mu+\langle\nabla_{\nabla\varphi^2}\nabla\varphi^3,\nabla_{\nabla\varphi^1}\nabla\varphi^4\rangle_\mu.\end{aligned}$$

Coupling the last two equations and then using the trivial identity $\langle P_\mu(v),P_\mu(w)\rangle_\mu=\langle v,w\rangle_\mu-\langle P^\perp_\mu(v),P^\perp_\mu(w)\rangle_\mu$, valid for any v, $w\in L^2_\mu$, we obtain the equality

$$\begin{aligned}&\langle\nabla_{\nabla\varphi^1}(\nabla_{\nabla\varphi^2}\nabla\varphi^3),\nabla\varphi^4\rangle_\mu\\=&\langle\nabla\big(\langle\nabla^2\varphi^3\cdot\nabla\varphi^2,\nabla\varphi^4\rangle\big),\nabla\varphi^1\rangle_\mu-\langle\nabla_{\nabla\varphi^2}\nabla\varphi^3,\nabla_{\nabla\varphi^1}\nabla\varphi^4\rangle_\mu\\=&\langle\nabla\big(\langle\nabla^2\varphi^3\cdot\nabla\varphi^2,\nabla\varphi^4\rangle\big),\nabla\varphi^1\rangle_\mu-\langle\nabla^2\varphi^3\cdot\nabla\varphi^2,\nabla^2\varphi^4\cdot\nabla\varphi^1\rangle_\mu\\&+\langle\xi_\mu(\varphi^3,\varphi^2),\xi_\mu(\varphi^4,\varphi^1)\rangle_\mu.\end{aligned}$$

The computation of the gradient of $\langle\nabla^2\varphi^3\cdot\nabla\varphi^2,\nabla\varphi^4\rangle$ gives

$$\begin{aligned}&\langle\nabla_{\nabla\varphi^1}(\nabla_{\nabla\varphi^2}\nabla\varphi^3),\nabla\varphi^4\rangle_\mu\\=&\int\nabla^3\varphi^3(\nabla\varphi^2,\nabla\varphi^4,\nabla\varphi^1)d\mu+\langle\nabla^2\varphi^3\cdot\nabla\varphi^4,\nabla^2\varphi^2\cdot\nabla\varphi^1\rangle_\mu\\&+\langle\xi_\mu(\varphi^3,\varphi^2),\xi_\mu(\varphi^4,\varphi^1)\rangle_\mu.\end{aligned}\tag{7.7}$$

Analogously, it holds:

$$\begin{aligned}&\langle\nabla_{\nabla\varphi^2}(\nabla_{\nabla\varphi^1}\nabla\varphi^3),\nabla\varphi^4\rangle_\mu\\=&\int\nabla^3\varphi^3(\nabla\varphi^1,\nabla\varphi^4,\nabla\varphi^2)d\mu+\langle\nabla^2\varphi^3\cdot\nabla\varphi^4,\nabla^2\varphi^1\cdot\nabla\varphi^2\rangle_\mu\\&+\langle\xi_\mu(\varphi^3,\varphi^1),\xi_\mu(\varphi^4,\varphi^2)\rangle_\mu,\end{aligned}\tag{7.8}$$

so that, subtracting (7.8) from (7.7), the symmetry of $\nabla^3\varphi^3$ gives

$$\begin{aligned}&\langle\nabla_{\nabla\varphi^1}(\nabla_{\nabla\varphi^2}\nabla\varphi^3),\nabla\varphi^4\rangle_\mu-\langle\nabla_{\nabla\varphi^2}(\nabla_{\nabla\varphi^1}\nabla\varphi^3),\nabla\varphi^4\rangle_\mu\\ =&\langle\nabla^2\varphi^3\cdot\nabla\varphi^4,\nabla^2\varphi^2\cdot\nabla\varphi^1\rangle_\mu-\langle\nabla^2\varphi^3\cdot\nabla\varphi^4,\nabla^2\varphi^1\cdot\nabla\varphi^2\rangle_\mu\\ &+\langle\xi_\mu(\varphi^3,\varphi^2),\xi_\mu(\varphi^4,\varphi^1)\rangle_\mu-\langle\xi_\mu(\varphi^3,\varphi^1),\xi_\mu(\varphi^4,\varphi^2)\rangle_\mu.\end{aligned}\tag{7.9}$$

Recalling equation (7.5) we get

$$\begin{aligned}\langle\nabla_{[\nabla\varphi^1,\nabla\varphi^2]}\nabla\varphi^3,\nabla\varphi^4\rangle_\mu&=\langle\nabla^2\varphi^3\cdot P_\mu(\nabla^2\varphi^2\cdot\nabla\varphi^1-\nabla^2\varphi^1\cdot\nabla\varphi^2),\nabla\varphi^4\rangle_\mu\\ &=\langle P_\mu(\nabla^2\varphi^2\cdot\nabla\varphi^1-\nabla^2\varphi^1\cdot\nabla\varphi^2),\nabla^2\varphi^3\cdot\nabla\varphi^4\rangle_\mu\\ &=\langle\nabla^2\varphi^3\cdot\nabla\varphi^4,\nabla^2\varphi^2\cdot\nabla\varphi^1-\nabla^2\varphi^1\cdot\nabla\varphi^2\rangle_\mu\\ &-\langle\xi_\mu(\varphi^2,\varphi^1),\xi_\mu(\varphi^3,\varphi^4)\rangle_\mu+\langle\xi_\mu(\varphi^1,\varphi^2),\xi_\mu(\varphi^3,\varphi^4)\rangle_\mu.\end{aligned}$$

Subtracting the last equations from (7.9), all the terms except those involving the functions ξ_μ cancel, and the thesis follows. □

From the representation formula of the curvature tensor, it follows immediately that the sectional curvatures of $\mathscr{P}_2(\mathbb{R}^d)$ are non-negative (for the definition see [6], Chapter 4, section 3). Indeed, it holds:

$$K(\nabla\varphi,\nabla\psi)(\mu)=\frac{\langle\mathcal{R}(\nabla\varphi,\nabla\psi)\nabla\psi,\nabla\varphi\rangle_\mu}{\|\nabla\varphi\|_\mu^2\|\nabla\psi\|_\mu^2-\langle\nabla\varphi,\nabla\psi\rangle_\mu^2}=\frac{3\|\xi_\mu(\varphi^1,\varphi^2)\|_\mu^2}{||\nabla\varphi||_\mu^2\|\nabla\psi\|_\mu^2-\langle\nabla\varphi,\nabla\psi\rangle_\mu^2}\geq 0.$$

7.2. A distance on the tangent bundle. Recall that, for a Riemannian manifold M, it is possible to endow the tangent bundle TM with a natural Riemannian metric, the so-called Sasaki metric, in the following way (see also [6], Chapter 3, exercise 2).

Fix a point $(p,u)\in TM$ and choose two regular curves $[0,1]\ni t\to\alpha^i(t)\in TM$, $i=1,2$, such that $\alpha^1(0)=\alpha^2(0)=(p,u)$. Let $(p^i(t),u^i(t)):=\alpha^i(t)$ and $v^i(t):=(p^i(t))'$, $i=1,2$. Clearly $V^i:=(\alpha^i)'(0)\in T_{(p,u)}(TM)$, $i=1,2$. The scalar product $\langle\cdot,\cdot\rangle^*$ between V^1 and V^2 is defined as

$$\langle V^1,V^2\rangle^*:=\langle v^1(0),v^2(0)\rangle+\langle\nabla_{v^1}u^1(0),\nabla_{v^2}u^2(0)\rangle.$$

It is possible to show that this is a good definition, that is, it depends only on V^1, V^2 and not on the particular curves $\alpha^1(t)$, $\alpha^2(t)$ chosen, therefore it defines a metric tensor on TM. It is then easy to see that the distance d on TM induced by this metric tensor is given by

$$d^2\Big((p^1,u^1),(p^2,u^2)\Big)=\inf_\gamma\big(\mathcal{L}(\gamma)\big)^2+|\mathcal{T}(u^1)-u^2|^2,\tag{7.10}$$

where the infimum is taken among all the smooth curves $\gamma(t)$ in M connecting p^1 to p^2, $\mathcal{L}(\gamma)$ is the length of γ and $\mathcal{T}(u^1)$ is the parallel transport of u^1 along γ to the point p^2.

Now turn back to the space $(\mathscr{P}_2(\mathbb{R}^d),W_2)$. Define the tangent bundle as

$$\mathrm{Tan}\,(\mathscr{P}_2(\mathbb{R}^d)):=\big\{(\mu,u)\ :\ \mu\in\mathscr{P}_2(\mathbb{R}^d),\ u\in\mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d))\big\},$$

and say that a sequence (μ_n, u_n) converges to (μ, u) if:

$$\begin{aligned} &\lim_{n\to\infty} W(\mu_n,\mu) = 0,\\ &\lim_{n\to\infty} \|u_n\|_{\mu_n} = \|u\|_{\mu},\\ &\lim_{n\to\infty} \langle u_n, \nabla\varphi\rangle_{\mu_n} = \langle u, \nabla\varphi\rangle_{\mu} \quad \forall \varphi \in C_c^\infty(\mathbb{R}^d). \end{aligned} \tag{7.11}$$

Even if we do not have a differential structure on the tangent bundle, and therefore we cannot mimick directly the definition of the Riemannian distance on it, equation (7.10) suggests the introduction of the following function on $[\mathrm{Tan}\,(\mathscr{P}_2(\mathbb{R}^d))]^2$:

$$d^2\big((\mu,u),(\nu,v)\big) := \inf_{\mu_t}\left\{(\mathcal{L}(\mu_t))^2 + \|v - \mathcal{T}_0^1(u)\|_\nu^2\right\},$$

where the infimum is taken on the set of regular curves $\mu_t : [0,1] \to \mathscr{P}_2(\mathbb{R}^d)$ such that $\mu_0 = \mu$ and $\mu_1 = \nu$, $\mathcal{L}(\mu_t)$ is the length of μ_t and $\mathcal{T}_t^s$ are the parallel transport maps along μ_t. In particular we define $d((\mu,u),(\nu,v)) := +\infty$ if there is no regular curve connecting μ to ν.

The function d behaves like a distance on $\mathrm{Tan}\,(\mathscr{P}_2(\mathbb{R}^d))$, the only problem being that it is not real valued. Given that regular curves are dense in the set of absolutely continuous curves, a natural candidate for a relaxation of d is its lower semicontinuous envelope d^*, defined by:

$$\begin{aligned} &d^*\big((\mu,u),(\nu,v)\big)\\ &:= \inf\left\{\liminf_{n\to\infty} d\big((\mu_n,u_n),(\nu_n,v_n)\big) \,:\, (\mu_n,u_n)\to(\mu,u),\ (\nu_n,v_n)\to(\nu,v)\right\}. \end{aligned}$$

However, it is not clear to us whether the function d^* is sufficienly well-behaved, for instance,whether the triangle inequality holds. Therefore we modify a bit the definition, and we introduce the function $\mathcal{D}$ as:

$$\begin{aligned} &\mathcal{D}\big((\mu,u),(\nu,v)\big)\\ &:= \inf\left\{d^*\big((\mu,\nabla\varphi),(\nu,\nabla\psi)\big) + \|u-\nabla\varphi\|_\mu + \|v-\nabla\psi\|_\nu \,:\, \varphi,\,\psi\in C_c^\infty(\mathbb{R}^d)\right\}. \end{aligned}$$

With the introduction of $\mathcal{D}$ we are allowed to regularize the vectors u, v, provided we pay the L^2 difference between the regularizations and the vectors themselves.

The following result is proved in Chapter 6 of [8].

PROPOSITION 7.3. *$\mathcal{D}$ is a distance and metrizes the convergence in* $\mathrm{Tan}\,(\mathscr{P}_2(\mathbb{R}^d))$ *in the sense of equations* (7.11). *Furthermore, for any absolutely continuous curve* (μ_t, u_t) *in* $\mathrm{Tan}\,(\mathscr{P}_2(\mathbb{R}^d))$ *the curve* μ_t *is absolutely continuous in* $(\mathscr{P}_2(\mathbb{R}^d), W_2)$ *and:*

$$\lim_{s\to t} \frac{\mathcal{D}\big((\mu_s,u_s),(\mu_t,u_t)\big)}{|s-t|} \geq \|v_t\|_{\mu_t} \qquad \text{for a.e. } t, \tag{7.12}$$

where v_t is the tangent field of μ_t.
Conversely, if μ_t is a regular curve and u_t is a parallel transport along it, the map (μ_t, u_t) is absolutely continuous in $(\mathrm{Tan}\,(\mathscr{P}_2(\mathbb{R}^d)), \mathcal{D})$ and equality holds a.e. in (7.12).

This proposition suggests a definition of *weak parallel transport* along an absolutely continuous curve μ_t: $u_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(\mathbb{R}^d))$ is a weak parallel transport along

μ_t if equality in (7.12) holds for a.e. t. For a weak parallel transport it is possible to show that the map $t \mapsto \langle u_t, \nabla\eta\rangle_{\mu_t}$ is absolutely continuous and that its derivative is given by (5.8).

However, it is not clear to the authors whether the weak parallel transport preserves the scalar product, or whether the parallel transport is unique. Furthermore, the density of regular curves is not enough to gain existence of weak parallel transport through an approximation argument. The key problem is that the space $(\mathrm{Tan}\,(\mathscr{P}_2(\mathbb{R}^d)), \mathcal{D})$ is *not* complete, as it may happen for a sequence $((\mu_n, u_n)) \subset \mathrm{Tan}\,(\mathscr{P}_2(\mathbb{R}^d))$ to converge to some (μ, u) with $u \notin \mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d))$: precisely, it may happen that $W(\mu_n, \mu) \to 0$, $u_n\mu_n \to u\mu$ in duality with $C_c(\mathbb{R}^d)$ and $\|v_n\|_{\mu_n} \to \|v\|_\mu$. Example 5.20 shows that it might be impossible to extend a (weak) parallel transport "backward" to the initial point of a geodesic.

REFERENCES

[1] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in spaces of probability measures*, Birkhäuser, 2005.

[2] L. Ambrosio, S. Lisini, and G. Savaré, *Stability of flows associated to gradient vector fields and convergence of iterated transport maps*, Manuscripta Math., 121 (2006), pp. 1–50.

[3] L. Ambrosio, G. Savaré, and L. Zambotti, *Existence and stability for Fokker-Planck equations with log-concave reference measure*, ArXiv Mathematics e-prints, (2007).

[4] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math., 84 (2000), pp. 375–393.

[5] J. A. Carrillo, R. J. McCann, and C. Villani, *Contractions in the 2-Wasserstein length space and thermalization of granular media*, Arch. Ration. Mech. Anal., 179 (2006), pp. 217–263.

[6] M. P. do Carmo, *Riemannian geometry*, Mathematics: Theory & Applications, Birkhäuser Boston Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

[7] L. C. Evans and W. Gangbo, *Differential equations methods for the Monge-Kantorovich mass transfer problem*, Mem. Amer. Math. Soc., 137 (1999), pp. viii+66.

[8] N. Gigli, *On the geometry of the space of measures in $\mathbb{R}^d$ endowed with the quadratic optimal transportation distance*, Scuola Normale Superiore, Pisa, 2008. (Paper can be downloaded from "http://cvgmt.sns.it")

[9] 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).

[10] J. Lott, *Some geometric calculations om Wasserstein space,* Comm. Math. Phys., 272 (2008), pp. 423–437.

[11] J. Lott and C. Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. Math., (to appear).

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

[13] S.-i. Ohta, *Gradient flows on wasserstein spaces over compact alexandrov spaces*, tech. rep., Universität Bonn, 2007.

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

[15] G. Savaré, *Gradient flows and diffusion semigroups in metric spaces under lower curvature bounds*, C. R. Math. Acad. Sci. Paris, 345 (2007), pp. 151–154.

[16] C. Villani, *Optimal transport, old and new*, Springer Verlag, 2008.