Second order analysis on $(\mathscr{P}_2(M), W_2)$

Nicola Gigli*
University of Bordeaux

May 2009

Abstract

We develop a rigorous second order analysis on the space of probability measures on a Riemannian manifold M endowed with the quadratic optimal transport distance W_2. Our discussion comprehends: definition of covariant derivative, discussion of the problem of existence of parallel transport, calculus of the Riemannian curvature tensor, differentiability of the exponential map and existence of Jacobi fields. This approach does not require any smoothness assumption on the measures considered.

Contents

*nicolagigli@googlemail.com

Introduction

The aim of this work is to build a solid theory of second order analysis in the quadratic Wasserstein space over a Riemannian manifold. To our knowledge, this topic has been investigated only by few authors: apart from the PhD thesis of the author [10] and his paper with Ambrosio [1] (both of these works were concerned with the case $M = \mathbb{R}^d$), the only other work on the topic of which we are aware is of Lott [14] (who considered generic compact Riemannian manifolds). These works, written independently at the same time, attack the problem from quite different viewpoints: Lott was more concerned with the description of the second order analysis, rather than with its construction, while the author and Ambrosio were more interested in proving existence theorems and to find the minimal regularity assumptions needed by the theory to work. This means that Lott assumed all of the objects he was working with to be smooth enough to justify his calculations: with this approach he was able to find out how the formulas for covariant derivative and curvature tensor look like (he also described the Poisson structure of the space $\mathscr{P}_2(M)$ when M has a Poisson structure itself - this is outside the scope of this paper). On the other hand, the author and Ambrosio worked, in the case $M = \mathbb{R}^d$, without regularity assumptions and were able to prove the existence of the parallel transport along a certain dense class of curves; once this was done, they turned to the definition of covariant derivative and curvature tensor and recovered the formulas found also by Lott for the case $M = \mathbb{R}^d$.

In this work we use the same approach used in [1, 10] to build the theory of second order analysis on $\mathscr{P}_2(M)$: on one hand, we replicate most of the results valid in $\mathbb{R}^d$ to this more general case, and we recover Lott's formulas in a more precise contest which allow us to describe the minimal regularity assumptions needed by the objects to be well defined. On the other, we push the investigation further, showing, for instance, that the problem of Jacobi fields is well posed, and identifying the solutions of the Jacobi equation with the differential of the exponential map.

The theory we develop works without any regularity assumption on the measures involved: actually, we will see that what often matters more, is some Lipschitz property of the vector fields involved.

Regarding the manifold, we assume that it is C^∞, connected, complete and without boundary.

The paper is structured as follows. In the first Chapter we recall the basic facts of the theory. Although the material presented here is now standard among specialists, we preferred to spend a bit of time in the introduction, mainly to fix the notation and the terminology we will use in the work. In particular, we recall the definition of tangent space $\mathrm{Tan}_\mu(\mathscr{P}_2(M)) \subset L^2_\mu$ (where L^2_μ is the Hilbert space of Borel tangent vector fields whose squared norm is μ-integrable):

$$\mathrm{Tan}_\mu(\mathscr{P}_2(M)) := \overline{\left\{\nabla\varphi \ : \ \varphi \in C^\infty_c(M)\right\}}^{L^2_\mu},$$

and the fact that for a given absolutely continuous curve (μ_t) in $(\mathscr{P}_2(M), W_2)$ there exists a unique choice - up to a negligible set of times - of $v_t \in L^2_{\mu_t}$ such that

$$\frac{d}{dt}\mu_t + \nabla\cdot(v_t\mu_t) = 0 \quad \text{in the sense of distributions}$$

and

$$\|v_t\|_{\mu_t} = \lim_{h\to 0}\frac{W_2(\mu_{t+h}, \mu_t)}{|h|}, \qquad a.e.\ t,$$

where $\|u\|_\mu$ is the norm of $u \in L^2_\mu$. For such a choice it always holds

$$v_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M)), \qquad a.e.\ t,$$

and we will call this vector field the *velocity vector field* of the curve (μ_t).

In the second Chapter we introduce the fundamental notion of *regular curve* which will be the curves along which we are able to define and study the regularity of vector fields. In order to introduce them, we first recall the definition of Lipschitz constant of a vector field $v \in L^2_\mu$:

$$\mathrm{L}(v) := \inf \varliminf_{n\to\infty} \sup_{x\in M} \|\nabla\xi^n(x)\|_{\mathrm{op}},$$

where the infimum is taken w.r.t. all sequences $n \mapsto \xi^n$ of smooth vector fields converging to v in L^2_μ and $\|\cdot\|_{\mathrm{op}}$ is the operator norm. Then we say that a given absolutely continuous curve (μ_t) is regular if its velocity vector field (v_t) satisfies

$$\int_0^1 \mathrm{L}(v_t)dt < \infty,$$
$$\int_0^1 \|v_t\|^2_{\mu_t} dt < \infty.$$

Under this assumptions and using the Cauchy-Lipschitz theorem it is immediate to verify the existence and uniqueness of maps $\mathbf{T}(t,s,x)$, which we will call *flow maps* satisfying

$$T(t,t,x) = x, \qquad \forall t, x \in \mathrm{supp}(\mu_t)$$
$$\frac{d}{ds}\mathbf{T}(t,s,x) = v_s\big(\mathbf{T}(t,s,x)\big), \qquad \forall t, x \in \mathrm{supp}(\mu_t) \text{ } a.e. \text{ } s.$$

The importance of regular curves is twofold: both geometric and algebraic.

From an algebraic point of view, the existence of the flow maps allows the *translation* of vector fields along a regular curve. This means the following. Suppose we have $u \in L^2_{\mu_t}$ for some t. Then we can define the translated vector field $\tau_t^s(u) \in L^2_{\mu_s}$ in the following way:

$$\tau_t^s(u)(x) := \begin{cases} \text{the parallel transport of } u(\mathbf{T}(s,t,x)) \text{ along the curve} \\ r \mapsto \mathbf{T}(s,r,x) \text{ from } r = t \text{ to } r = s. \end{cases}$$

Since the parallel transport is always an isometry, it is not hard to check that $\tau_t^s : L^2_{\mu_t} \to L^2_{\mu_s}$ is actually an isometry. Also, from the group property of the flow maps it easily follows that $\tau_r^s \circ \tau_t^r = \tau_t^s$.

From a geometric point of view, the importance of regular curves is due to the fact that the *angle between tangent spaces* varies smoothly along these curves. This means that if we have $u \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$ and we translate it to obtain $\tau_t^s(u) \in L^2_{\mu_s}$ (in general the result of this translation is no more a tangent vector), then $\tau_t^s(u)$ is 'almost' tangent for s close to t, in the following quantitative sense:

$$\left\| \tau_t^s(v) - \mathrm{P}_{\mu_s}\big(\tau_t^s(v)\big) \right\|_{\mu_s} \leq \|v\|_{\mu_t} \left(e^{\left| \int_t^s \mathrm{L}(v_r)dr \right|} - 1 \right), \tag{0.1}$$

where $\mathrm{P}_\mu : L^2_\mu \to \mathrm{Tan}_\mu(\mathscr{P}_2(M))$ is the orthogonal projection. As we will see, this fact (explained and proven in theorem 2.13), will be the key enabler for the proof of existence of parallel transport.

We conclude the Chapter with the study of the problem of regularity of geodesics. Here we will spend some time to improve the known result due to Fathi ([8]), that if (v_t) is the velocity vector field of a geodesic between measures with compact support, then $\mathrm{L}(v_t) \leq \frac{C}{\min\{t,1-t\}}$, for some c.

What we prove, is that under the same assumptions, there exist functions ϕ_t which are both $-\frac{C}{\min\{t,1-t\}}$-concave and $-\frac{C}{\min\{t,1-t\}}$-convex such that $v_t = \nabla\phi_t$ on $\mathrm{supp}(\mu_t)$ (as we will discuss, this fact is not a direct consequence of Fathi's result couped with the tangency of the v_t's).

In the third Chapter, we define the notion of absolutely continuous vector fields and study their first properties. The idea is the following: given a vector field (u_t) along (μ_t), i.e. a map $t \mapsto u_t \in L^2_{\mu_t}$, we can say that it is absolutely continuous whenever the map

$$t \mapsto \tau_t^{t_0}(u_t) \in L^2_{\mu_{t_0}},$$

is absolutely continuous for any t_0. Observe the key role played by the translation maps: using them we are able to carry the problem of time regularity of a vector field defined on different L^2 spaces for different times, into a question on the regularity of a curve with values in the fixed Hilbert space $L^2_{\mu_{t_0}}$. Also, by the group property of the translation maps, it is immediate to check that $t \mapsto \tau_t^{t_0}(u_t) \in L^2_{\mu_{t_0}}$ is absolutely continuous for any t_0 if and only if it is for some t_0.

The definition of (total) derivative of an absolutely continuous vector field now comes pretty naturally: it is sufficient to set

$$\frac{\boldsymbol{d}}{dt} u_t := \tau_0^t \left(\frac{d}{dt} \tau_t^0(u_t) \right) \in L^2_{\mu_t}, \qquad \forall t_0,$$

and we will prove, among other properties of such derivation, the important Leibniz rule:

$$\frac{d}{dt} \left\langle u_t^1, u_t^2 \right\rangle_{\mu_t} = \left\langle \frac{\boldsymbol{d}}{dt} u_t^1, u_t^2 \right\rangle_{\mu_t} + \left\langle u_t^1, \frac{\boldsymbol{d}}{dt} u_t^2 \right\rangle_{\mu_t}, \qquad a.e.\ t,$$

where $\langle \cdot, \cdot \rangle_\mu$ is the scalar product in L^2_μ. The same idea of translating the vector field, works also for higher order of regularity: for instance, we can say that (u_t) is C^∞ whenever $t \mapsto \tau_t^{t_0}(u_t) \in L^2_{\mu_{t_0}}$ is C^∞ for some, and thus any, t_0. We will discuss the topic by also giving some concrete examples.

We will spend the rest of the Chapter in analyzing the main properties of absolutely continuous vector fields. Among these, we prove that 'smooth vector fields are dense in the class of absolutely continuous vector fields'. This means that given an absolutely continuous vector field (u_t), we can find a sequence $n \mapsto (\xi_t^n)$ of smooth vector fields $(x,t) \mapsto \xi_t^n(x) \in T_x M$ such that

$$\begin{aligned} \xi_t^n &\to u_t, && \text{in } L^2_{\mu_t} \text{ as } n \to \infty \text{ uniformly on } t \\ \frac{\boldsymbol{d}}{dt} \xi_t^n &\to \frac{\boldsymbol{d}}{dt} u_t, && \text{in } L^2_{\mu_t} \text{ as } n \to \infty \text{ for a.e. } t, \\ \int_0^1 \left\| \frac{\boldsymbol{d}}{dt} \xi_t^n \right\|_{\mu_t} dt &\to \int_0^1 \left\| \frac{\boldsymbol{d}}{dt} u_t \right\|_{\mu_t} dt, && \text{as } n \to \infty. \end{aligned}$$

The fourth Chapter is dedicated to the problem of parallel transport: we show that it exists along regular curves, and that in general it may fail to exists if the curve is not regular. This part of the work is very similar to the analogous part appeared in [1, 10].

In the fifth Chapter we introduce the covariant derivative. To do so, we imitate the classical definition of covariant derivative via parallel transport and define:

$$\frac{\boldsymbol{D}}{dt} u_t := \lim_{h \to 0} \frac{\mathcal{T}_{t+h}^t(u_{t+h}) - u_t}{h},$$

where $\mathcal{T}_t^s$ are the parallel transport maps along (μ_t), (u_t) is a tangent and absolutely continuous vector field and the limit is intended in $L^2_{\mu_t}$. It turns out that the covariant derivative of a tangent and absolutely continuous vector field (u_t) exists for a.e. t and is given by

$$\frac{\boldsymbol{D}}{dt}u_t = \mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t\right).$$

A major result of this Chapter (which is new both respect to [14] and to [1, 10]) is that if (u_t) is an absolutely continuous vector field, not necessarily tangent, then the vector field $(\mathrm{P}_{\mu_t}(u_t))$ is absolutely continuous as well and its covariant derivative is given by

$$\frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(u_t) = \mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t - \nabla v_t \cdot \mathrm{P}^{\perp}_{\mu_t}(u_t)\right),$$

where it is part of the result the fact that the above expression makes sense for a.e. t along a regular curve (note that if v_t is only Lipschitz and μ_t arbitrary, the gradient of v_t may very well be not defined a.e. w.r.t. μ_t). Starting from this formula, we will be able to compute total and covariant derivatives of some basic kind of vector fields. We will conclude the section with the study of smoothness in time of operators. A regularity result that we obtain is that along a geodesic in $\mathcal{P}_c(M)$, the projection operator P_{μ_t} is a C^∞ operator: this means that if (u_t) is C^∞ (in the sense described above, i.e. the curve $t \mapsto \tau_t^0(u_t) \in L^2_{\mu_0}$ is C^∞), then $\mathrm{P}_{\mu_t}(u_t)$ is C^∞ as well.

In the sixth Chapter, we study the curvature tensor. We start by computing it for vector fields of very special kind (this will be done by imitating the analogous calculations done by Lott in [14]), then, using the techniques developed in the previous Chapters, we will be able to discuss precisely the minimal regularity assumption needed to be sure that this formula makes sense.

The seventh Chapter is devoted to study the differentiability properties of the exponential map. The problem can be settled in the following terms: given a regular curve (μ_t) and an absolutely continuous vector field (u_t) along it, is that true that the curve $t \mapsto \nu_t := \mathbf{exp}_{\mu_t}(u_t) = (\exp(u_t))_\#\mu_t$ is absolutely continuous? If the answer is yes, who is its velocity vector field? Under mild assumptions on (μ_t) and (u_t) we will see that actually the curve (ν_t) is absolutely continuous, and that its velocity vector field may be identified via the use of Jacobi fields on M in the following way. Assume (just for simplicity, this assumption is actually unneeded) that the map $\exp(u_t)$ is μ_t-essentially invertible for any t, and for fixed t and given $\tilde u^1, \tilde u^2 \in L^2_{\mu_t}$, define the vector field $j_{\mu_t,u_t}(\tilde u^1, \tilde u^2)$ by:

$$j_{\mu_t,u_t}(\tilde u^1, \tilde u^2)(\exp_x(u_t(x))) := \begin{cases} \text{the value at } s=1 \text{ of the Jacobi field } s \mapsto j_s \in T_{\exp_x(su_t(x))}M \\ \text{along the geodesic } s \mapsto \exp_x(su_t(x)) \\ \text{having initial conditions } j_0(x) = \tilde u^1(x),\ j_0'(x) = \tilde u^2(x). \end{cases}$$

Then the vector field $j_{\mu_t,u_t}(\tilde u^1, \tilde u^2)$ is well defined ν_t-a.e. and, with some assumptions on μ_t, u_t actually belongs to $L^2_{\nu_t}$. It is then not hard to verify (see in particular the first section of the

Chapter) that the curve (ν_t) satisfies the continuity equation

$$\frac{d}{dt}\nu_t + \nabla \cdot \Big(j_{\mu_t,u_t}\Big(v_t, \frac{\boldsymbol{d}}{dt}u_t\Big)\nu_t\Big) = 0,$$

so that its velocity vector field is given by

$$\mathrm{P}_{\nu_t}\left(j_{\mu_t,u_t}\Big(v_t, \frac{\boldsymbol{d}}{dt}u_t\Big)\right).$$

In the last Chapter we show that the curvature operator and the differential of the exponential map are linked together via Jacobi fields, that is: on one hand, since we have the curvature tensor, we can write down the Jacobi equation, and on the other, since we know how to differentiate the exponential map, we have a reasonable guess on the solution of the Jacobi equation itself. What we prove here is that - as expected - this guess is true: the differential of the exponential map provides the unique solution of the Jacobi equation with prescribed initial conditions.

As a first application of the existence of Jacobi fields, we prove an estimate from above of the distance between optimal transport maps in terms of Wasserstein distance between their push forwards. To be more clear, let us assume that $M = \mathbb{R}^d$ and recall that in this case it holds the trivial inequality

$$W_2(T_\#\mu, S_\#\mu) \leq \|T - S\|_\mu.$$

In general, this inequality cannot be reversed. However, in some special situations, something can be said. Let $\nu := T_\#\mu$ and assume that T is optimal. Also assume that the geodetic from μ to ν induced by T can be extended in both directions for 'sufficiently long time' (the precise quantification is given in section 8.3), and let $S : \mathrm{supp}(\mu) \to \mathbb{R}^d$ be another optimal transport map. Then we know that $T_\varepsilon := T + \varepsilon S$ is optimal as well for any $\varepsilon > 0$. Let $\nu_\varepsilon := (T_\varepsilon)_\#\mu$. The inequality that we prove is:

$$\overline{\lim_{\varepsilon\to 0}} \frac{\|T_\varepsilon - T\|_\mu}{W_2(\nu_\varepsilon, \nu)} \leq C < \infty,$$

giving also some informations on the value of C.

Finally, in the appendix we study some related problems regarding the geometry of the space $(\mathscr{P}_2(M), W_2)$. In the first, we prove that any absolutely continuous curve may be approximated by regular curves made by, roughly said, 'smooth measures with smooth velocities'.

In the second we show how the Riemannian structure of $(\mathscr{P}_2(M), W_2)$ leads to the notion of C^1 curves: we discuss the topic by also giving some examples.

Finally, in the last appendix we discuss the possibility of defining a good notion of absolute continuity for vector fields defined along general absolutely continuous curves (and not just regular ones). The results here are not conclusive. In terms of applicability, it seems that the definition we propose is better than the one proposed in [1] and [10]: at least it allows to do a first - very basic - step into the world of Lagrangian over $(\mathscr{P}_2(M), W_2)$. As we discovered when the work on this paper was concluded, the question of Lagrangian over $(\mathscr{P}_2(M), W_2)$ was already exploited, from a different point of view, in the recent work [9].

1 Preliminaries and notation

Notation

(μ_t), (u_t)...	Maps from a real interval to the set of measures, vector fields...
$C_c^\infty(M)$	Set of smooth real valued functions on M with compact support
d	Riemannian distance on M,
$\mathcal{C}(K)$	Bound from above of the curvature of M on the compact set $K \subset M$, see first section below
$\mathcal{H}(K)$	Bound on the Hessian of the squared distance on K, see end of first section below,
$\mathbf{exp}_\mu$	Exponential map on $\mathscr{P}_2(M)$ defined as $\mathbf{exp}_\mu(u) := (\exp(u))_\# \mu$ for any $u \in L^2_\mu$, see also 1.7,
$\dfrac{\boldsymbol{d}}{dt} u_t$	Total derivative of the absolutely continuous vector field (u_t), see 3.6,
$\dfrac{\boldsymbol{D}}{dt} u_t$	Covariant derivative of the tangent and absolutely continuous vector field (u_t), see 5.1,
$\mathrm{L}(u)$	Lipschitz constant of the vector field $u \in L^2_\mu$, see 2.1,
$\mathrm{S}(u)$	Supremum of $\lvert u(x) \rvert$ over $\mathrm{supp}(\mu)$ for a given Lipschitz vector $u \in L^2_\mu$, see 2.2
LNL_μ	Lipschitz-non-Lipschitz set. It is the subset of $[L^2_\mu]^2$ which is the natural domain of the tensor $\mathcal{N}_\mu$. It is endowed with the topology defined in 5.7,
$\mathcal{N}_\mu$	Tensor describing the infinitesimal variation of tangent spaces, see 5.9,
$\mathcal{O}_v(\cdot)$	Map from L^2_μ into itself defined - for any given Lipschitz vector $v \in L^2_\mu$ - as $\mathcal{O}_v(u) := \mathcal{N}_\mu(v, u)$, see 5.11,
$\mathcal{O}_v^*(\cdot)$	Adjoint of $\mathcal{O}_v(\cdot)$, see 5.11,
$\mathrm{P}_\mu, \mathrm{P}_\mu^\perp$	Orthogonal projections from L^2_μ to the tangent (resp. normal) space,
$\mathbf{R}$	Curvature tensor in $(\mathscr{P}_2(M), W_2)$,
$\mathbf{T}(t, s, \cdot)$	Flow maps of a given regular curve, or, more generally, flow maps associated to a transport couple (μ_t, v_t) satisfying $\int_0^1 \mathrm{L}(v_t) < \infty$, see 2.6,
τ_t^s	Translation maps along a given regular curve, see 2.11
$\mathcal{T}_t^s$	Parallel transport maps along a given regular curve, see Chapter 4,
$\mathcal{V}(M)$	Set of maps $\xi(\cdot) : M \to TM$ such that $\xi(x) \in T_xM$ for every $x \in M$, which are smooth and with compact support
$\mathcal{V}(M \times [0,1])$	Set of maps $\xi_\cdot(\cdot) : [0,1] \times M \to TM$ such that $\xi_t(x) \in T_xM$ for every $x \in M$, $t \in [0,1]$, which are smooth and with compact support,
W_2	Quadratic Wasserstein distance. Typically the underlying space is $\mathscr{P}_2(M)$, but sometime we will deal with $(\mathscr{P}_2(X), W_2)$, where $X = (X, d)$ is a metric space different from (M, d) (e.g. $X = TM$ endowed with the Sasaki metric),

1.1 Riemannian manifolds

In this work M is a connected complete C^∞ Riemannian manifold, TM is the tangent bundle and T_xM the tangent space at $x \in M$. The scalar product on T_xM will be denoted by $\langle \cdot, \cdot \rangle$. We will denote by $C_c^\infty(M)$ the set of C^∞, compactly supported, real valued functions, by $\mathcal{V}(M)$ the set of C^∞, compactly supported tangent vector fields. In the following we will often use the letters φ, ψ for generic elements of $C_c^\infty(M)$ and ξ, η for those of $\mathcal{V}(M)$. For example, with this notation we have $\langle \xi, \eta \rangle \in C_c^\infty(M)$. The gradient of a function $\varphi \in C_c^\infty(M)$ is a well defined element of $\mathcal{V}(M)$ and will be denoted by $\nabla\varphi$. The set $\mathcal{V}(M \times [0,1])$ is the set of tangent vector fields, smooth and compactly supported, that is, of the maps $(t,x) \mapsto \xi_t(x) \in T_xM$ which are C^∞ in both the variables and such that $\xi_t(x) = 0$ for every $t \in [0,1]$ if $x \notin K$ for some compact set $K \subset M$.

For a given C^1 curve γ from an interval $I \subset \mathbb{R}$ onto M, we will write γ' for its derivative, so that we have $\gamma'(t) \in T_{\gamma(t)}M$ for every $t \in I$.

We endow M with the metric d induced by the Riemannian structure, i.e.: for any couple $x, y \in M$ we define

$$\mathrm{d}(x,y) := \inf \int_0^1 |\gamma'(t)| dt,$$

where the infimum is taken on the set of C^1 curves on $[0,1]$ satisfying $\gamma(0) = x$ and $\gamma(1) = y$.

We say that a curve γ is a geodesic if it has constant speed and is globally minimizing in the sense of the above definition.

For two given vector fields $\xi, \eta \in \mathcal{V}(M)$, their Lie bracket will be denoted by $[\xi, \eta]$, i.e.:

$$[\xi,\eta](\varphi) := \xi(\eta(\varphi)) - \eta(\xi(\varphi)), \qquad \forall \varphi \in C_c^\infty(M).$$

The Levi-Civita connection on M (or covariant derivative) will be denoted by ∇. Recall that it is well defined the vector $\nabla_v \xi \in T_xM$ for any $x \in M$, $v \in T_xM$ and $\xi \in \mathcal{V}(M)$ and that the covariant derivative is identified in the set of all connections by the identities:

$$\begin{aligned} \nabla \left\langle \xi^2, \xi^3 \right\rangle \cdot \xi^1 &= \left\langle \nabla_{\xi^1} \xi^2, \xi^3 \right\rangle + \left\langle \xi^2, \nabla_{\xi^1} \xi^3 \right\rangle, \\ [\xi^1, \xi^2] &= \nabla_{\xi^1} \xi^2 - \nabla_{\xi^2} \xi^1, \end{aligned} \tag{1.1}$$

called *compatibility with the metric* and *torsion free identity*, respectively.

For given $\xi \in \mathcal{V}(M)$ and $x \in M$, the covariant derivative defines a linear map, called gradient of ξ and denoted by $\nabla\xi(x)$, from T_xM into itself given by $v \mapsto \nabla_v \xi$:

$$\nabla\xi(x) \cdot v := \nabla_v \xi(x). \tag{1.2}$$

The divergence $\nabla \cdot \xi$ of a vector field ξ is the trace of its gradient. If ξ is the gradient of a smooth function $\varphi \in C_c^\infty(M)$, we will denote its gradient by $\nabla^2\varphi$, and a simple consequence of equations (1.1) is

$$\left\langle \nabla^2\varphi \cdot \xi^1, \xi^2 \right\rangle = \left\langle \nabla_{\xi^1} \nabla\varphi, \xi^2 \right\rangle = \left\langle \nabla_{\xi^2} \nabla\varphi, \xi^1 \right\rangle = \left\langle \nabla^2\varphi \cdot \xi^2, \xi^1 \right\rangle, \quad \forall \varphi \in C_c^\infty(M), \xi^1, \xi^2 \in \mathcal{V}(M), \tag{1.3}$$

which shows that $\nabla^2\varphi$ is a symmetric operator.

Given $\xi \in \mathcal{V}(M)$, $\nabla^2\xi(x)$ is a tensor with 3 indexes, which we will always think as the map from $[T_xM]^2$ into T_xM defined by

$$\nabla^2\xi(\eta^1,\eta^2) := \big(\nabla(\nabla\xi\cdot\eta^1)\big)\cdot\eta^2 - \nabla\xi\cdot(\nabla\eta^1\cdot\eta^2), \tag{1.4}$$

where both the sides depend on $x \in M$. Using the definition (1.2) of gradient of a vector field, it is easy to check that $\nabla^2\xi$ is a tensor on η^1 and η^2, thus the above is a good definition.

Given two vector fields ξ^1, ξ^2, the curvature operator $R(\xi^1,\xi^2)$ is a family of maps from T_xM into itself, defined for any $x \in M$ as

$$R(\xi^1,\xi^2)\xi^3 := \nabla_{\xi^2}\nabla_{\xi^1}\xi^3 - \nabla_{\xi^1}\nabla_{\xi^2}\xi^3 + \nabla_{[\xi^1,\xi^2]}\xi^3,$$

where both sides depend on $x \in M$ (some textbooks define this tensor with the opposite sign - here we are aligned to [6]). Applying the definition of $\nabla\xi$ and $\nabla^2\xi$ it is easy to check that $R(\xi^1,\xi^2)\xi^3$ is a tensor on the three vector fields involved ad that it holds the formula

$$R(\xi^1,\xi^2)\xi^3 = \nabla^2\xi^3(\xi^1,\xi^2) - \nabla^2\xi^3(\xi^2,\xi^1), \tag{1.5}$$

it is for this reason that it is said that the curvature measures the non commutativity of the derivatives.

For every compact set $K \subset M$, we will denote by $\mathcal{C}(K) \in \mathbb{R}$ the bound on the curvature on K, that is: $\mathcal{C}(K)$ is the smallest constant $\mathcal{C}$ for which is true:

$$\big|\left\langle R(u^1,u^2)u^3,u^4\right\rangle_x\big| \le \mathcal{C}|u^1|_x|u^2|_x|u^3|_x|u^4|_x \qquad \forall x \in K, u^1,u^2,u^3,u^4 \in T_xM.$$

Recall that given the Riemannian metric on M, there is a natural Riemannian metric induced on TM: the so-called Sasaki metric (see e.g. [6], Chapter 3, exercise 2). To describe it, fix a point $(x,u) \in TM$ and choose two smooth curves $[0,1] \ni t \to \alpha^i(t) \in TM$, $i = 1,2$, such that $\alpha^1(0) = \alpha^2(0) = (x,u)$. Let $(x^i(t),u^i(t)) := \alpha^i(t)$ and $v^i(t) := (x^i(t))'$, $i = 1,2$. 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

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

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 $\alpha^1(t)$, $\alpha^2(t)$, 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

$$\mathrm{D}^2\Big((x^1,u^1),(x^2,u^2)\Big) = \inf_\gamma\Big\{\big(\mathcal{L}(\gamma)\big)^2 + |\mathcal{T}(u^1)-u^2|^2\Big\}, \tag{1.6}$$

where the infimum is taken among all the smooth curves $\gamma(t)$ in M connecting x^1 to x^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 x^2. In particular, we have the following bound on D:

$$\mathrm{D}^2\Big((x^1,u^1),(x^2,u^2)\Big) \le \mathrm{d}^2(x^1,x^2) + |\mathcal{T}(u^1)-u^2|^2, \tag{1.7}$$

where here $\mathcal{T}(u^1)$ is the parallel transport of u^1 into $T_{x^2}M$ along any geodesic connecting x^1 to x^2. In the following we will sometime write $\mathrm{D}(u^1,u^2)$ instead than $\mathrm{D}\Big((x^1,u^1),(x^2,u^2)\Big)$, when it is clear who are the base points x^1, x^2.

A notion of particular importance on Riemannian manifolds, is that of Jacobi field, which identifies the gradient of the exponential map. Recall that the problem which the Jacobi fields solve is: given a smooth path $s\mapsto(x_s,v_s)\in TM$, what is the derivative at $s=0$ of $\exp_{x_s} v_s$? Let us briefly recall how the Jacobi equation comes out. Consider the smooth map $f:[0,1]^2\to M$ given by $f(t,s):=\exp_{x_s}(tv_s)$. From the equality $\frac{D}{dt}\frac{d}{ds}f=\frac{D}{ds}\frac{d}{dt}f$, which may be written as $[\frac{d}{dt}f,\frac{d}{ds}f]=0$, and the fact that $\frac{D}{dt}\frac{D}{dt}f=0$ it follows

$$R\left(\frac{d}{dt}f,\frac{d}{ds}f\right)\frac{d}{dt}f=\frac{D}{ds}\frac{D}{dt}\frac{d}{dt}f-\frac{D}{dt}\frac{D}{ds}\frac{d}{dt}f=-\frac{D}{dt}\frac{D}{dt}\frac{d}{ds}f.$$

Letting $J_t:=\frac{d}{ds}f_{|s=0}$, $\gamma(t):=f(t,0)$ and evaluating the above equation for $s=0$, we get

$$J_t''+R(\gamma'(t),J_t)\gamma'(t)=0,$$

which is the Jacobi equation (here J_t'' stands for the second covariant derivative along γ). Thus the searched value of $\frac{d}{ds}\exp_{x_s}(v_s)_{|s=0}$ is given by J_1, where (J_t) solves the Jacobi equation with the initial conditions

$$J_0=x_0',$$
$$J_0'=\lim_{h\to 0}\frac{\tau_s^0(v_s)-v_0}{h},$$

where $\tau_s^0(v_s)$ is the parallel transport of v_s from $T_{x_s}M$ to $T_{x_0}M$ along the curve $s\mapsto x_s$ (if $x_0'\neq 0$ the formula reads as $J_0'=\nabla_{x_0'}v_s$). It is just a matter of standard comparison arguments based on the inequality

$$-\mathcal{C}(K)|v_0|^2|J_t|\leq|J_t''|\leq\mathcal{C}(K)|v_0|^2|J_t|,$$

to see that for the norm of J_t they hold the bounds:

$$|J_t|\geq|J_0|\cos\left(|v_0|\sqrt{\mathcal{C}(K)}t\right)-\frac{|J_0'|}{|v_0|\sqrt{\mathcal{C}(K)}}\sin\left(|v_0|\sqrt{\mathcal{C}(K)}t\right) \tag{1.8a}$$

$$|J_t|\leq|J_0|\cosh\left(|v_0|\sqrt{\mathcal{C}(K)}t\right)+\frac{|J_0'|}{|v_0|\sqrt{\mathcal{C}(K)}}\sinh\left(|v_0|\sqrt{\mathcal{C}(K)}t\right), \tag{1.8b}$$

where K is the range of $t\mapsto\exp_{x_0}(tv_0)$. If M has non negative sectional curvature, much more is true: there is no exponential behavior on the norm of the Jacobi field, that is

$$|J_t|\leq|J_0|+t|J_0'|. \tag{1.9}$$

From (1.8b) we get the inequalities

$$\mathrm{d}\big(\exp_x(v_1),\exp_x(v_2)\big)\leq\frac{|v_1-v_2|}{\max\{|v_1|,|v_2|\}\sqrt{\mathcal{C}(K)}}\sinh\left(\max\{|v_1|,|v_2|\}\sqrt{\mathcal{C}(K)}\right), \tag{1.10}$$

valid for any $x \in M$, $v_1, v_2 \in T_xM$, and $K = \{\exp_x(tv_1 + sv_2)\}_{t,s\in[0,1]}$, and

$$\mathrm{d}\big(\exp_x(v), \exp_y(w)\big) \leq \mathrm{D}\big((x,v),(y,w)\big)\left(\cosh\left(V\sqrt{\mathcal{C}(K)}\right) + \frac{1}{V\sqrt{\mathcal{C}(K)}}\sinh\left(V\sqrt{\mathcal{C}(K)}\right)\right), \tag{1.11}$$

valid for any $x, y \in M$ and $v \in T_xM$, $w \in T_yM$, where $V := \max\{|v|_x, |w|_y\}$ and $K = \{\exp_{x_s}(v_t)\}_{t,s\in[0,1]}$ and (x_s, v_s) is a geodesic in TM connecting (x,v) to (y,w). The first follows by considering the curve $\gamma(t) := \exp_x((1-t)v_1 + tv_2)$ and observing that $\mathrm{d}(\exp_x(v_1), \exp_x(v_2)) \leq \int_0^1 |\gamma'(t)|dt$ (observe that $r \mapsto \sinh(r)/r$ is increasing on $[0, +\infty)$). The second follows by considering a constant speed geodesic $t \mapsto (x_t, v_t)$ from (x,v) to (y,w) in (TM, D), defining the curve $\gamma(t) := \exp_{x_t}(v_t)$ and using the fact that $\mathrm{d}\big(\exp_x(v), \exp_y(w)\big) \leq \int_0^1 |\gamma'(t)|dt$.

Definition 1.1 ($\lambda-$convexity and $\lambda-$concavity) *We say that a function $\varphi : M \to \mathbb{R} \cup \{\pm\infty\}$ is $\lambda-$convex on a certain open set $\Omega \subset M$ if the distributional Hessian of φ is bigger or equal than λId on Ω. Similarly for $\lambda-$concavity.*

We say that a function is semiconvex (semiconcave) if it is λ convex (concave) for some $\lambda \in \mathbb{R}$.

It is possible to check that the supremum of a family of $\lambda-$convex functions on Ω is still $\lambda-$convex, and that the infimum of a family of $\lambda-$concave functions if $\lambda-$concave.

For a function $\varphi : M \to \mathbb{R} \cup \{\pm\infty\}$ we indicate with $\partial^-\varphi(x)$ the subdifferential of φ at x, defined as the set of vectors $v \in T_xM$ such that it holds

$$\varphi(y) \geq \varphi(x) + \langle v, \exp_x^{-1}(y)\rangle + o(\mathrm{d}(x,y)),$$

and similarly for the superdifferential $\partial^+\varphi$. A λ-convex function has non empty subdifferential at all points in the interior of the set where it is finite.

We conclude recalling that the distance function $x \mapsto \frac{\mathrm{d}^2(x,y)}{2}$ is locally semiconcave. In particular, for every compact set $K \subset M$, there exists a least constant $\mathcal{H}(K) > 0$ such that $x \mapsto \mathrm{d}^2(x,y)$ is $-\mathcal{H}(K)$-concave on some neighborhood of K for every $y \in K$. If the manifold has non-negative sectional curvatures, then $x \mapsto \frac{\mathrm{d}^2(x,y)}{2}$ is -1-concave on the whole M for any $y \in M$.

1.2 The distance W_2

The natural set to endow with the Wasserstein distance is the set $\mathscr{P}_2(M)$ of probability measures with bounded second moment:

$$\mathscr{P}_2(M) := \left\{\mu \in \mathscr{P}(M) \ : \ \int \mathrm{d}^2(x, x_0)d\mu(x) < \infty \quad \forall x_0 \in M\right\}.$$

It is easy to check that $\int \mathrm{d}^2(x, x_0)d\mu(x) < \infty$ for any $x_0 \in M$ if and only if it is finite for some $x_0 \in M$.

Recall that for any couple of topological spaces X, Y, any Borel probability measure μ on X and any Borel map $f : X \to Y$, the push forward $f_\#\mu$ of μ through f is the Borel probability measure on Y defined by

$$f_\#\mu(E) := \mu\big(f^{-1}(E)\big), \qquad \forall \text{ Borel sets E} \subset \text{Y}.$$

The Wasserstein distance W_2 on $\mathscr{P}_2(M)$ is defined by

$$W_2(\mu,\nu) := \sqrt{\inf \int \mathrm{d}^2(x,y)d\boldsymbol{\gamma}(x,y)},$$

where the infimum is taken in the set $\mathcal{Adm}(\mu,\nu)$ of *admissible plans* $\boldsymbol{\gamma}$ from μ to ν, i.e. among all the probability measures on M^2 satisfying $\pi^1_\#\boldsymbol{\gamma} = \mu$ and $\pi^2_\#\boldsymbol{\gamma} = \nu$, where π^1 and π^2 are the projections onto the first and second coordinate respectively. The quantity $\int \mathrm{d}^2(x,y)d\boldsymbol{\gamma}(x,y)$ is called the *cost* of the plan $\boldsymbol{\gamma}$. A plan which realizes the infimum is called *optimal* and the set of optimal plans for a given couple (μ,ν) of measures will be indicated by $\mathcal{Opt}(\mu,\nu)$. A plan is said to be *induced by a map*, if it is of the form $(Id,T)_\#\mu$ for some measurable map T. This is the same as to say that $\boldsymbol{\gamma}$ is concentrated in the graph of T.

It is well known that the function W_2 is a distance on $\mathscr{P}_2(M)$, we skip the proof this fact: the interested reader may study the question in detail on, for instance, [25].

For a good understanding of the theory of optimal transport, it is of fundamental importance the notion of cyclical monotonicity.

Definition 1.2 (d^2-cyclical monotonicity) *A subset K of $M \times M$ is d^2-cyclically monotone if for every $n \in \mathbb{N}$, every $(x_i, y_i) \in K$, $i = 0, \ldots, n-1$, and every permutation σ of $\{0, \ldots, n-1\}$ it holds:*

$$\sum_{i=0}^{n-1} \mathrm{d}^2(x_i, y_i) \leq \sum_{i=0}^{n-1} \mathrm{d}^2(x_i, y_{\sigma(i)}),$$

where $y_n := y_0$.

The importance of the above definition is due to the following well-known theorem, for which we give only a sketch of the proof. By support (supp) of a measure, we intend the smallest closed set on which the measure is concentrated.

Theorem 1.3 *Let $\mu, \nu \in \mathscr{P}_2(M)$. A plan $\boldsymbol{\gamma} \in \mathcal{Adm}(\mu,\nu)$ is optimal if and only if its support* $\mathrm{supp}(\boldsymbol{\gamma})$ *is a d^2-cyclically monotone set.*

Idea of the proof.

(*Optimality $\Rightarrow$ d^2-cyclical monotonicity*). Argue by contradiction. Suppose that γ is an optimal plan which whose support is not d^2-cyclically monotone. Then there exist $n \in \mathbb{N}$, $(x_i, y_i) \in \mathrm{supp}(\boldsymbol{\gamma})$, $i = 0, \ldots, n-1$, and a permutation σ of $\{0, \ldots, n-1\}$ such that

$$\sum_{i=0}^{n-1} \mathrm{d}^2(x_i, y_i) > \sum_{i=0}^{n-1} \mathrm{d}^2(x_i, y_{\sigma(i)}).$$

This inequality implies that the cost of the plan γ may be reduced by 'redefining it on the points (x_i, y_i)'. Indeed, instead than moving some mass from x_i to y_i, $i = 0, \ldots, n-1$, it is strictly better to move the same mass from x_i to $y_{\sigma(i)}$. This contradicts the optimality, and the implication follows.

(d^2-*cyclical monotonicity* $\Rightarrow$ *Optimality*). Roughly said, every transport plan from μ to ν can be obtained from another transport plan by a 'reshuffling', therefore the previous argument can be reversed. To see this, assume that both μ and ν are concentrated on N points, say $\mu = \frac{1}{N}\sum_1^N \delta_{x_i}$ and $\nu = \frac{1}{N}\sum_1^N \delta_{y_i}$ (by a density argument it can be shown that it is sufficient to deal with this kind of measures). Then it is known that in this situation the extremal points of the closed and convex set $\mathscr{Adm}((,\mu),\nu)$ are plans induced by invertible maps between the supports of the measures. Since the cost of a plan is a linear map from $\mathscr{Adm}((,\mu),\nu)$ to $\mathbb{R}$, we can assume that the plans we are dealing with are induced by maps. Thus let $(Id, T)_\#\mu$ be a d^2-cyclically monotone plan: we want to prove that it is optimal. To this aim, pick an optimal plan $(Id, S)_\#\mu$ and consider the map $R := S^{-1} \circ T$. The map R is clearly a permutation of the set $\{x_1, \ldots, x_N\}$. Consider any cycle of this permutation, say - up to a relabeling - $x_1, x_2, \ldots, x_k$, so that $R(x_i) = x_{i+1}$ for $1 \le i < k$ and $R(x_k) = x_1$. Up to relabeling also the points in the support of ν, we may assume that $T(x_i) = y_i$ for $1 \le i \le k$, so that $S(x_i) = y_{i-1}$ for $1 < i \le k$ and $S(x_1) = y_k$. Since $(Id, T)_\#\mu$ is d^2-cyclically monotone by hypothesis, we have

$$\sum_{i=1}^{k} \mathrm{d}^2(x_i, y_i) \le \sum_{i=1}^{k} \mathrm{d}^2(x_i, y_{i-1}).$$

On the other side, since $(Id, S)_\#\mu$ is optimal, by the previous step we know that it is d^2-cyclically monotone, therefore it also holds the opposite inequality. Thus we have

$$\sum_{i=1}^{k} \mathrm{d}^2(x_i, y_i) = \sum_{i=1}^{k} \mathrm{d}^2(x_i, y_{i-1}).$$

Adding up this equality over all the cycles of R we obtain that the cost of $(Id, T)_\#\mu$ is equal to the cost of $(Id, S)_\#$
mu, and therefore it is optimal. □

We will denote by L^2_μ the set of measurable vector fields v on M such that $\int |v(x)|_x^2 d\mu(x) < \infty$ endowed with the scalar product $\langle v, w\rangle_\mu := \int \langle v(x), w(x)\rangle_x \, d\mu(x)$ where we identify two vector fields which differs only on a $\mu-$negligible set. The norm of a vector $v \in \mathrm{L}^2_\mu$ will be denoted by $\|v\|_\mu$. For a real valued function φ we will write $\|\varphi\|_\mu$ for $(\int |\varphi(x)|^2 d\mu(x))^{1/2}$.

Definition 1.4 (Optimal tangent plans) *Given $\mu, \nu \in \mathscr{P}_2(M)$, we will denote by $\mathscr{OptTan}(\mu,\nu)$ the set of plans $\gamma \in \mathscr{P}(TM)$ satisfying*

$$(\pi^M, \exp)_\#\gamma \in \mathscr{Opt}(\mu,\nu),$$
$$\int |\mathrm{v}|^2 d\gamma(x, \mathrm{v}) = W_2^2(\mu,\nu).$$

Thus a plan $\boldsymbol{\gamma} \in \mathit{OptTan}(\mu, \nu)$ describes in which direction the mass of μ should move to reach ν in the fastest way. Observe that plans in $\mathit{OptTan}(\mu, \nu)$ carry more information about optimal transport, than those in $\mathit{Opt}(\mu, \nu)$, as the latter is only an optimal coupling between the initial and final mass , while the former describe also which geodesic path we are choosing from the starting point to the final one (if the starting measure is sufficiently regular - absolutely continuous w.r.t. the volume measure is enough - then there is no difference in the two descriptions, since the optimal map 'almost never hits the cut locus' - see e.g. [25], bibliographical notes of Chapter 13 - In this case it is well known that both the sets contain only one element, which is induced by a map, see also theorem 1.17).

Definition 1.5 (Rescalation of plans) *For a general measure $\boldsymbol{\gamma} \in \mathscr{P}(TM)$ and a real number λ we define the* rescaled *measure $\lambda \cdot \boldsymbol{\gamma} \in \mathscr{P}(TM)$ as*

$$\lambda \cdot \boldsymbol{\gamma} := (\mathrm{Omot}(\lambda))_{\#}\boldsymbol{\gamma},$$

where $\mathrm{Omot}(\lambda) : TM \to TM$ *is defined by* $\mathrm{Omot}(\lambda)(x, v) := (x, \lambda v)$.

It is well known that geodesics connecting two measures μ and ν in $\mathscr{P}_2(M)$ are in correspondence with plans in $\mathit{OptTan}(\mu, \nu)$:

Proposition 1.6 *Let $\mu, \nu \in \mathscr{P}_2(M)$ and $[0, 1] \ni t \mapsto \mu_t \in \mathscr{P}_2(M)$ a curve. Then $t \mapsto \mu_t$ is a geodesic connecting μ to ν if and only if there exists $\boldsymbol{\gamma} \in \mathit{OptTan}(\mu, \nu)$ such that*

$$\mu_t = (\exp)_{\#}(t \cdot \boldsymbol{\gamma}), \qquad \forall t \in [0, 1].$$

This proposition motivates the following definition:

Definition 1.7 (Exponential map in $(\mathscr{P}_2(M), W_2)$) *Let $\boldsymbol{\gamma} \in \mathscr{P}_2(TM)$, i.e. a measure in $\mathscr{P}(TM)$ satisfying $\pi^M_{\#}\boldsymbol{\gamma} \in \mathscr{P}_2(M)$ and $\int |\mathrm{v}|^2 d\boldsymbol{\gamma}(x, \mathrm{v}) < \infty$. Then the exponential $\mathbf{exp}(\boldsymbol{\gamma}) \in \mathscr{P}_2(M)$ is defined by*

$$\mathbf{exp}(\boldsymbol{\gamma}) := (\exp)_{\#}\boldsymbol{\gamma}.$$

Similarly, if $\mu \in \mathscr{P}_2(M)$ and $v \in L^2_\mu$, the exponential $\mathbf{exp}_\mu(v) \in \mathscr{P}_2(M)$ of the vector v with base measure μ is:

$$\mathbf{exp}_\mu(v) := \big(\exp(v)\big)_{\#}\mu.$$

We conclude with the definition of convergence of vector fields belonging to different L^2 spaces.

Definition 1.8 (Convergence of maps) *Let $\mu^n, \mu \in \mathscr{P}_2(M)$, $v^n \in \mathrm{L}^2_{\mu^n}$, and $v \in \mathrm{L}^2_\mu$, $n \in \mathbb{N}$, and assume that $W_2(\mu^n, \mu) \to 0$ as $n \to \infty$. We say that the sequence of maps (v^n) weakly converges to v if the following two conditions hold:*

$$\begin{aligned} \sup_{n\in\mathbb{N}} \|v^n\|_{\mu^n} &< \infty && \textit{as } n \to \infty, \\ \langle v^n, \xi \rangle_{\mu^n} &\to \langle v, \xi \rangle_\mu && \textit{as } n \to \infty \ \forall \xi \in \mathcal{V}(M). \end{aligned}$$

It is possible to check that in this situation it holds $\|v\|_\mu \leq \underline{\lim}_{n\to\infty} \|v^n\|_{\mu^n}$. We say that (v^n) strongly converges to v (or simply converges) if it converges weakly and $\lim_{n\to\infty} \|v^n\|_{\mu^n} = \|v\|_\mu$.

1.3 Kantorovich's dual problem

Here we recall briefly the dual problem, introduced by Kantorovich, of the optimal transport. Again, everything works in a level of generality much bigger than the one described here: for a complete treatment of the problem, see Chapter 5 of [25].

Just a word on the notation used. There is not yet a universal alignment regarding the signs to use in the definition of $c-$convexity, $c-$concavity and $c-$transforms, here we use the same notation of [25], with just one exception: we distinguish between the two notions of $c-$transforms which arise.

Definition 1.9 (c_+ and c_- transforms) *Let $\varphi : M \to \mathbb{R} \cup \{\pm\infty\}$. The functions φ^{c+}, φ^{c-} are defined as*

$$\varphi^{c+}(x) := \inf_{y \in M} \left(c(x,y) - \varphi(y)\right),$$
$$\varphi^{c-}(x) := \sup_{y \in M} \left(-c(x,y) - \varphi(y)\right).$$

The relation between c^+ and c^- transform is:

$$\varphi^{c+} = -\left(-\varphi\right)^{c-}$$

Observe that we have the following two trivial inequalities:

$$\varphi(x) + \varphi^{c-}(y) \geq -c(x,y), \qquad \forall x, y \in M,$$
$$\varphi(x) + \varphi^{c+}(y) \leq c(x,y), \qquad \forall x, y \in M,$$

where in the first we adopted the convention $+\infty + (-\infty) = +\infty$ and in the second $-\infty + \infty = -\infty$.

Definition 1.10 ($c-$convexity and $c-$concavity) *We say that $\varphi : M \to \mathbb{R} \cup \{-\infty\}$ is $c-$concave if it is not identically $-\infty$ and there exists $\psi : M \to \mathbb{R} \cup \{\pm\infty\}$ such that*

$$\varphi = \psi^{c+}.$$

Similarly, we say that $\varphi : M \to \mathbb{R} \cup \{+\infty\}$ is $c-$convex if it is not identically $+\infty$ and there exists $\psi : M \to \mathbb{R} \cup \{\pm\infty\}$ such that

$$\varphi(x) = \psi^{c-}.$$

Clearly the opposite of a $c-$convex function is $c-$concave and viceversa.

Definition 1.11 ($c-$subdifferential and $c-$superdifferential) *Let $\varphi : M \to \mathbb{R} \cup \{+\infty\}$ be a $c-$convex function. Its $c-$subdifferential $\partial^{c-}\varphi \subset M^2$ is defined as*

$$\partial^{c-}\varphi := \{(x,y) \ : \ \varphi(x) + \varphi^{c-}(y) = -c(x,y)\},$$

and the $c-$subdifferential $\partial^{c-}\varphi(x)$ at a point $x \in M$ is the set of y such that $(x,y) \in \partial^{c-}\varphi$. Analogously, if $\varphi : M \to \mathbb{R} \cup \{-\infty\}$ is a $c-$concave function, its $c-$superdifferential $\partial^{c+}\varphi \subset M^2$ is defined as

$$\partial^{c+}\varphi := \{(x,y) \ : \ \varphi(x) + \varphi^{c+} = c(x,y)\},$$

and the $c-$superdifferential $\partial^{c+}\varphi(x)$ at a point $x \in M$ is the set of y such that $(x,y) \in \partial^{c+}\varphi$.

We recall the statement of Kantorovich's duality result.

Theorem 1.12 (Kantorovich's duality) *Let $\mu, \nu \in \mathscr{P}_2(M)$. Then there exists a $c-$concave function ψ such that for every $\gamma \in \mathit{Opt}(\mu, \nu)$ it holds $\mathrm{supp}(\gamma) \subset \partial^{c+}\psi$. For any such ψ it holds $\psi \in L^1_\mu$, $\psi^{c+} \in L^1_\nu$ and*

$$W_2^2(\mu, \nu) = \int \psi d\mu + \int \psi^{c-} d\nu. \tag{1.12}$$

A similar statement holds with $c-$convex functions.

Given a couple $\mu, \nu \in \mathscr{P}_2(M)$, we will call any $c-$concave function ψ satisfying equation (1.12) a *$c-$concave potential*. Similarly, if ψ is $c-$convex and satisfies

$$W_2^2(\mu, \nu) = -\int \psi \, d\mu - \int \psi^{c-} d\nu,$$

we will say that it is a *$c-$convex potential*. The theorem we just proved ensures the existence of both $c-$concave and $c-$convex potentials.

Remark 1.13 Suppose that the manifold M has non negative sectional curvature. Then the squared distance function is -1-concave. Thus in this situation any c-convex potential is -1-convex and any c-concave potential is -1-concave. ■

Remark 1.14 In case the two given measures μ, ν have compact support, there exists a $c-$concave Kantorovich potential ψ of the form

$$\psi(x) = \inf_{y \in K} c(x, y) - f(y),$$

for some function $f : M \to \mathbb{R} \cup \{\pm\infty\}$, where K is a compact set whose interior contains the supports of μ and ν.

In particular, by the bound on the Hessian of the squared distance function, this potential is $-\mathcal{H}_K-$concave in the interior of K. ■

It is important to underline that the $c-$subdifferential (as well as the $c-$superdifferential), at a certain point x is made of points on the manifold, and not of tangent vectors. However there is a strict link between the $c-$subdifferential and the usual subdifferential, as the following proposition shows: this link was exploited in the setting of optimal transport by McCann in [18] (the same argument used by McCann was already known to Cabré which used it in an earlier work on elliptic equation on manifolds [4]).

Proposition 1.15 (Cabré-McCann) *Let $\varphi : M \to \mathbb{R} \cup \{+\infty\}$ be a c-convex function and $x, y \in M$ such that $y \in \partial^{c-}\varphi(x)$. Then $\exp_x^{-1}(y) \subset \partial^-\varphi(x)$. Conversely, if φ is differentiable at x and $\nabla\varphi(x) = v$, then $y := \exp_x(v)$ is the unique point in $\partial^{c-}\varphi(x)$. Similarly for $c-$concave functions.*

Proof Choose $v \in T_xM$ such that $\exp_x(v) = y$ and recall that the superdifferential of $\mathrm{d}^2(\cdot, y)$ contains v, that is:

$$\frac{\mathrm{d}^2(z,y)}{2} \le \frac{\mathrm{d}^2(x,y)}{2} + \langle v, \exp_x^{-1}(z)\rangle + o(\mathrm{d}(x,z)).$$

Now observe that the condition $y \in \partial^{c-}\varphi(x)$ may be written as

$$-\varphi(z) - \frac{\mathrm{d}^2(z,y)}{2} \le -\varphi(x) - \frac{\mathrm{d}^2(x,y)}{2}, \qquad \forall z \in M,$$

from which it follows

$$\varphi(z) - \varphi(x) \ge -\frac{\mathrm{d}^2(z,y)}{2} + \frac{\mathrm{d}^2(x,y)}{2} \ge \langle v, \exp_x^{-1}(z)\rangle + o(\mathrm{d}(x,z)),$$

which is the conclusion.

To prove the converse, is enough to show that the $c-$subdifferential of φ at x is non empty, as then from what we just proved the (only) point in it must coincide with $\exp_x(v)$. From the definition of $c-$convexity we know that there exists sequences $n \mapsto y_n \in M$ and $n \mapsto a_n \in \mathbb{R}$ such that for every $n \in \mathbb{N}$ it holds

$$\begin{aligned} \varphi(x) &\le -\frac{\mathrm{d}^2(x,y_n)}{2} + a_n + \varepsilon_n, \\ \varphi(\tilde{x}) &\ge -\frac{\mathrm{d}^2(x,y_n)}{2} + a_n, \qquad \forall \tilde{x} \in M, \end{aligned} \tag{1.13}$$

for some $\varepsilon_n \downarrow 0$. To conclude it is sufficient to show that $\{y_n\}_{n\in\mathbb{N}}$ is bounded, as then any limit of (y_n) belongs to $\partial^{c-}\varphi(x)$. Argue by contradiction and assume that $\mathrm{d}(x, y_n) =: d_n \to \infty$ as $n \to \infty$. Now for every $n \in \mathbb{N}$ let γ_n be a constant speed geodesic on $[0,1]$ from x to y_n and define

$$x_n := \gamma_n\left(d_n^{-\frac{3}{2}}\right).$$

Note that $\mathrm{d}(x, x_n) = d_n^{-\frac{1}{2}} \to 0$. The second equation in (1.13) gives

$$\varphi(x_n) \ge -\frac{\mathrm{d}^2(x_n,y_n)}{2} + a_n = -\frac{1}{2}\left(d_n - \frac{1}{\sqrt{d_n}}\right)^2 + a_n \ge \varphi(x) - \varepsilon_n + \sqrt{d_n} - \frac{1}{2d_n} \to +\infty,$$

which contradicts the continuity of φ in x. □

Remark 1.16 The converse implication in this theorem is false if one doesn't assume φ to be differentiable at x: i.e. it is *not* true in general that $v \in \partial^-\varphi(x)$ implies $\exp_x(v) \in \partial^{c-}\varphi(x)$. (We thank Figalli for having helped us understanding this point). The question is related to the so called *regularity* of the cost function. A sufficient condition for this regularity is the satisfaction of the Ma-Trudinger-Wang condition (see [16]). We won't stress this point further, the interested reader may look at [25], chapter 12. ■

The Kantorovich duality result and the above proposition allow to understand when the optimal plan is unique and induced by a map and to characterize this map. Quite surprisingly, we won't need this result in this work (we chose to include it because of its fundamental importance in the theory of mass transportation).

Theorem 1.17 (Brenier-McCann) *Let $\mu, \nu \in \mathcal{P}_c(M)$ and assume that μ is absolutely continuous. Then there exists a unique optimal plan from μ to ν and this plan is induced by the map* $\exp(\nabla\varphi)$*, where φ is a Kantorovich c-convex potential for μ, ν.*

Proof By remark 1.14 we know that there exists a Kantorovich potential φ which is semi-convex in some open set Ω containing the supports of both μ and ν. By a classical result of convex analysis, φ is a.e. differentiable in Ω w.r.t. the volume measure. Thus, by the hypothesis on μ, it is also $\mu-$a.e. differentiable. By theorem 1.12 we know that every optimal plan $\boldsymbol{\gamma}$ from μ to ν must be concentrated on $\partial^{c-}\varphi$. By proposition 1.15 and what we said on the differentiability of φ we get that for μ-a.e. x there is only on $y \in M$ such that $(x, y) \in \mathrm{supp}(\boldsymbol{\gamma})$, and that this y is identified by $y = \exp_x(\nabla\varphi(x))$. Which is the thesis. □

In the previous section we pointed out the structure of geodesics in $\mathscr{P}_2(M)$; a natural question which arises is then: how are $c-$convex potential interpolated along geodesics? The answers is: through the Hopf-Lax formula. We recall below the basic facts and proofs, the exposition is mainly burrowed from [25].

Definition 1.18 (Interpolation of supports) *Let $\mu, \nu \in \mathscr{P}_2(M)$ and let $A_0 := \mathrm{supp}(\mu)$, $A_1 := \mathrm{supp}(\nu)$. The sets A_t, $t \in [0, 1]$, are defined as:*

$$A_t := \Big\{\gamma(t) \ : \ \gamma(\cdot) \text{ is a minimizing geodesic, } \gamma(0) \in A_0, \ \gamma(1) \in A_1\Big\}.$$

From proposition 1.6 it follows immediately that $\mathrm{supp}(\mu_t) \subset A_t$ for every t, when (μ_t) is a geodesic connecting μ_0 to μ_1.

Definition 1.19 (Interpolated costs) *For any $t < s \in [0, 1]$ we define the function $c^{t,s} : M^2 \to \mathbb{R}$ as*

$$c^{t,s}(x, y) := \frac{\mathrm{d}^2(x, y)}{s - t}$$

Proposition 1.20 *Let $t_1 < t_2 < t_3 \in [0, 1]$. Then it holds*

$$c^{t_1,t_3}(x, y) = \inf_{z \in M} c^{t_1,t_2}(x, z) + c^{t_2,t_3}(z, y). \tag{1.14}$$

Proof From the inequality

$$(a + b)^2 \leq \frac{a^2}{t} + \frac{b^2}{1 - t},$$

valid for any $a, b \in \mathbb{R}$ and $t \in (0, 1)$ we get that the inequality $\leq$ holds in (1.14). To gain that equality can be reached, consider a constant speed geodesic $\gamma : [t_1, t_3] \to M$ connecting x to y and choose $z = \gamma(t_2)$. □

From the definition of the interpolated costs, they come out naturally the definitions of $c^{t,s}$–concavity, $c^{t,s}$–convexity, $c_{\pm}^{t,s}$ transforms and $c_{\pm}^{t,s}$ differentials by imitating the analogous definition given for the cost $c = c^{0,1}$, we omit the details.

Fix $\mu_0, \mu_1 \in \mathscr{P}_2(M)$.

Definition 1.21 (Hopf-Lax formula) *Let $\phi : M \to \mathbb{R} \cup \{\pm\infty\}$ be a function and $t, s \in [0,1]$. Define the function $H_t^s(\phi)$ as*

$$H_t^s(\phi)(x) := \begin{cases} \inf_{y \in A_s} \big\{ c^{t,s}(x,y) + \phi(y) \big\}, & \text{if } t < s, \\ \phi(x) & \text{if } t = s, \\ \sup_{y \in A_t} \big\{ -c^{t,s}(x,y) + \phi(y) \big\}, & \text{if } t > s. \end{cases}$$

Proposition 1.22 (Basic properties of the Hopf-Lax formula) *We have the following three properties:*

i) For any $t, s \in [0,1]$ the maps H_t^s is order preserving, that is $\phi \le \psi \Rightarrow H_t^s(\phi) \le H_t^s(\psi)$.

ii) For any $t < s \in [0,1]$ it holds

$$H_s^t\big(H_t^s(\phi)\big) = \phi^{c_+^{t,s} c_-^{t,s}} \le \phi,$$
$$H_t^s\big(H_s^t(\phi)\big) = \phi^{c_-^{t,s} c_+^{t,s}} \ge \phi,$$

iii) For any $t, s \in [0,1]$ it holds

$$H_t^s \circ H_s^t \circ H_t^s = H_t^s.$$

Proof The order preserving property is a straightforward consequence of the definition. To prove property (ii) observe that

$$H_s^t\big(H_t^s(\phi)\big)(x) = \sup_y \inf_{x'} \Big(\phi(x') + c^{t,s}(x',y) - c^{t,s}(x,y) \Big),$$

which gives the equality $H_s^t\big(H_t^s(\phi)\big) = \phi^{c_+^{t,s} c_-^{t,s}}$: in particular, choosing $x' = x$ we get claim (the proof of the other equation is similar). For the last property assume $t < s$ (the other case is similar) and observe that

$$\underbrace{H_t^s \circ H_s^t}_{\ge Id} \circ H_t^s \ge H_t^s$$

and

$$H_t^s \circ \underbrace{H_s^t \circ H_t^s}_{\le Id} \le H_t^s.$$

□

Proposition 1.23 *Let ψ be a $c-$convex potential for (μ_0,μ_1) and define $\psi_s := H_0^s(\psi)$. Then it holds*

$$\psi_s(\gamma(s)) = c^{0,s}(\gamma(0),\gamma(s)) + \psi(\gamma(0)),$$

for every curve γ of the kind $t \mapsto \exp_x(tv)$, where $(x,v) \in \operatorname{supp}(\boldsymbol{\gamma})$ and $\boldsymbol{\gamma} \in \mathit{OptTan}(\mu_0,\mu_1)$.

Proof Let γ as in the hypothesis. By definition of H_0^s we have

$$\psi_s(\gamma(s)) \leq c^{0,s}(\gamma(0),\gamma(s)) + \psi(\gamma(0)).$$

To prove the opposite inequality, observe that from the $c-$convexity hypothesis, we know that

$$\begin{aligned}\psi(x) &\geq -c^{0,1}(x,\gamma(1)) - \psi^{c_+}(\gamma(1)) \\ &= -c^{0,1}(x,\gamma(1)) + c^{0,1}(\gamma(0),\gamma(1)) + \psi(\gamma(0)), \qquad \forall x \in M,\end{aligned}$$

thus it holds

$$\begin{aligned}\psi_s(\gamma(s)) &= \inf_{x\in M}\left\{c^{0,s}(\gamma(s),x) + \psi(x)\right\} \\ &\geq \left(\inf_{x\in M}\left\{c^{0,s}(\gamma(s),x) - c^{0,1}(x,\gamma(1))\right)\right\} + c^{0,1}(\gamma(0),\gamma(1)) + \psi(\gamma(0)) \\ &= -c^{s,1}(\gamma(s),\gamma(1)) + c^{0,1}(\gamma(0),\gamma(1)) + \psi(\gamma(0)) \\ &= c^{0,s}(\gamma(0),\gamma(s)) + \psi(\gamma(0)).\end{aligned}$$

□

Proposition 1.24 (Interpolated potentials) *Let ψ be a $c-$convex potential for $\mu_0,\mu_1 \in \mathscr{P}_2(M)$ and let (μ_t) be a constant speed geodesic connecting μ_0 to μ_1. Define $\psi_s := H_0^s(\psi)$ for any $s \in (0,1]$. Then ψ_s is a $c^{t,s}-$*concave *potential for μ_t,μ_s for any $t < s$.*

*Similarly, if ψ is a $c-$concave potential for μ_1,μ_0 and we define $\psi_t := H_1^t(\psi)$, then ψ_t is a $c^{t,s}-$*convex *potential for μ_t,μ_s for any $t < s$.*

Proof We need to show that

$$\psi_s(\gamma(s)) + \psi_s^{c_+^{t,s}}(\gamma(t)) = c^{t,s}(\gamma(t),\gamma(s)).$$

By definition of $c_+^{t,s}$ transform we have

$$\psi_s^{c_+^{t,s}}(\gamma(t)) = \inf_{y\in M} c^{t,s}(\gamma(t),y) + \psi_s(y) \leq c^{t,s}(\gamma(t),\gamma(s)) + \psi_s(\gamma(s)).$$

To prove the other inequality start observing that

$$\begin{aligned}\psi_s(y) = \inf_{x\in M}\left\{c^{0,s}(x,y) + \psi(x)\right\} &\leq c^{0,s}(\gamma(0),y) + \psi(\gamma(0)) \\ &\leq c^{0,t}(\gamma(0),\gamma(t)) + c^{t,s}(\gamma(t),y) + \psi(\gamma(0)),\end{aligned}$$

and conclude by

$$\begin{aligned}\psi_s^{c_+^{t,s}}(\gamma(t)) &= \inf_{y\in M}\Big\{c^{t,s}(\gamma(t),y)+\psi_s(y)\Big\}\\ &\geq -c^{0,t}(\gamma(0),\gamma(t))-\psi(\gamma(0))\\ &= c^{t,s}(\gamma(t),\gamma(s))-c^{0,s}(\gamma(0),\gamma(s))-\psi(\gamma(0))\\ &= c^{t,s}(\gamma(t),\gamma(s))-\psi_s(\gamma(s)).\end{aligned}$$

□

Remark 1.25 (Concavity and convexity of interpolated potentials) A direct consequence of the definition and of remark 1.14 is that if $\mu_0,\mu_1\in\mathcal{P}_c(M)$, then ψ_t is $-\frac{\mathcal{H}_K}{t}$-concave on K, where $K:=\cup_t A_t$. Similarly, φ_t is $-\frac{\mathcal{H}_K}{1-t}$-convex on K. ■

Proposition 1.26 (Regularity properties of the interpolated potentials) *Let ψ be a $c-$convex potential for (μ_0,μ_1) and let $\varphi:=H_0^1(\psi)$. Define $\psi_t:=H_0^t(\psi)$ and $\varphi_t:=H_1^t(\varphi)$ and choose a geodesic (μ_t) from μ_0 to μ_1. Then for every $t\in(0,1)$ it holds:*

i) $\psi_t\geq\varphi_t$ *and both the functions are real valued,*

ii) $\psi_t(x)=\varphi_t(x)$ *for any* $x\in\operatorname{supp}(\mu_t)$,

iii) ψ_t *and* φ_t *are differentiable in the support of* μ_t *and* $\nabla\psi_t(x)=\nabla\varphi_t(x)$ *for any* $x\in\operatorname{supp}(\mu_t)$.

Proof For (i) we have

$$\varphi_t=H_1^t(\varphi)=(H_1^t\circ H_0^1)(\psi)=(\underbrace{H_1^t\circ H_t^1}_{\leq Id}\circ H_0^t)\psi\leq H_0^t(\psi)=\psi_t.$$

Now observe that by definition, $\psi_t(x)<+\infty$ and $\varphi_t(x)>-\infty$ for every $x\in M$, thus it holds

$$+\infty>\psi(x)\geq\varphi(x)>-\infty,\qquad\forall x\in M.$$

To prove (ii), choose $\boldsymbol{\gamma}\in\mathit{OptTan}(\mu_0,\mu_1)$ which induces (μ_t), choose $(x,v)\in\operatorname{supp}(\boldsymbol{\gamma})$ and define $\gamma_t:=\exp_x(tv)$. Recall that it holds

$$\begin{aligned}\psi_t(\gamma_t)&=c^{0,t}(\gamma_0,\gamma_t)+\psi(\gamma_0),\\ \varphi_t(\gamma_t)&=c^{t,1}(\gamma_t,\gamma_1)+\varphi(\gamma_1).\end{aligned}$$

Thus from $\varphi(\gamma_1)=c^{0,1}(\gamma_0,\gamma_1)+\psi(\gamma_0)$ we get that $\psi_t(\gamma_t)=\varphi_t(\gamma_t)$. Since $\operatorname{supp}(\mu_t)=\{\gamma_t\}$ when γ varies in the set of curves of the kind $\exp_x(tv)$ with $(x,v)\in\operatorname{supp}(\boldsymbol{\gamma})$, (ii) follows.

Now we turn to (iii). With the same choice of $t\mapsto\gamma_t$ as above, recall that it holds

$$\begin{aligned}\psi_t(\gamma_t)&=c^{0,t}(\gamma_0,\gamma_t)+\psi(\gamma_0)\\ \psi_t(x)&\leq c^{0,t}(\gamma_0,x)+\psi(\gamma_0),\qquad\forall x\in M,\end{aligned}$$

and that the function $x \mapsto c^{0,t}(\gamma_0, x) + \psi(\gamma_0)$ is superdifferentiable at $x = \gamma_t$. Thus the function $x \mapsto \psi_t$ is superdifferentiable at $x = \gamma_t$. Similarly, φ_t is subdifferentiable at γ_t. Choose $v_1 \in \partial^+\psi_t(\gamma_t)$, $v_2 \in \partial^-\varphi_t(\gamma_t)$ and observe that

$$\psi_t(\gamma_t) + \left\langle v_1, \exp_{\gamma_t}^{-1}(x)\right\rangle + o(\mathrm{d}(x,\gamma_t)) \geq \psi_t(x) \geq \varphi_t(x) \geq \varphi_t(\gamma_t) + \left\langle v_2, \exp_{\gamma_t}^{-1}(x)\right\rangle + o(\mathrm{d}(x,\gamma_t)),$$

which gives $v_1 = v_2$ and the thesis. □

This proposition has deep consequences in term of structure of geodesics in $(\mathscr{P}_2(M), W_2)$: the first of them are given below:

Proposition 1.27 *Let (μ_t) be a geodesic in $(\mathscr{P}_2(M), W_2)$. With the same notation of the previous proposition, define*

$$v_t(x) := \nabla\varphi_t(x) = \nabla\psi_t(x), \qquad \forall t \in (0,1),\ x \in \mathrm{supp}(\mu_t).$$

Then for every $s \in [0,1]$ the map $\exp\big((s-t)v_t\big)$ is the unique optimal map from μ_t to μ_s.

What is interesting in this proposition is the fact that the transport from intermediate times is always unique and induced by a map. It can be proved that this map is locally Lipschitz, but we postpone the proof of this fact to section 2.3, where we refine a bit this well known result.

Proof Given the structure of geodesics, it is sufficient to prove the claim for $s > t$. The previous proposition ensures that the vectors v_t are well defined. Now fix $t \in (0,1)$, $s \in [0,1]$ and choose a plan $\boldsymbol{\gamma} \in \mathit{Opt}(\mu_t, \mu_s)$. By theorem 1.12 we know that $\boldsymbol{\gamma}$ is concentrated on $\partial^{c_+^{t,s}}\psi_t$, and by proposition 1.15 that $\partial^{c_+^{t,s}}\psi_t \subset \{\exp_x((s-t)\partial^+\psi_t)\}$ (well, actually we stated both theorems when dealing with $c-$concavity/convexity, but it is immediate to reformulate them for $c^{t,s}-$concavity/convexity). Since we know that $\partial^+\psi_t(x)$ is single valued for any $x \in \mathrm{supp}(\mu_t)$, and that its value is $v_t(x)$, we get

$$\boldsymbol{\gamma} = \Big(Id, \exp\big((s-t)v_t\big)\Big)_{\#}\mu_t,$$

which is the claim. □

In particular we proved that $v_t \in L^2_{\mu_t}$ for any $t \in (0,1)$ and that $\|v_t\|_{\mu_t} = W_2(\mu_0, \mu_1)$.

1.4 First order differentiable structure

In this section we recall the main features of the first order differentiable calculus in Wasserstein spaces.

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

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

It turns out (see for instance [2, 1.1.2]) 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

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

In order to describe the differentiable structure of the Wasserstein space we start with purely heuristic considerations, as in [20]: the continuity equation

$$\frac{d}{dt}\mu_t + \nabla \cdot (v_t \mu_t) = 0 \tag{1.15}$$

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 exponential of 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 [2, 8.3.1], provides a complete differential characterization of the class of absolutely continuous curves in the Wasserstein space and makes rigorous this picture (in [2] the theorem is proved when the underlying space is $\mathbb{R}^d$, the generalization to the case of manifolds is a straightforward consequence of Nash's embedding theorem and presents no difficulties, we skip the details).

All the curves we consider are defined on $[0, 1]$, unless otherwise stated. We will write (μ_t) for the map $[0, 1] \ni t \mapsto \mu_t \in \mathscr{P}_2(M)$.

Theorem 1.28 *Let (μ_t) be an absolutely continuous curve in $(\mathscr{P}_2(M), W_2)$. Then there exists a velocity field $v_t \in L^2_{\mu_t}$ with $\|v_t\|_{\mu_t} \in L^1(0, 1)$ such that the continuity equation* (1.15) *holds and*

$$\|v_t\|_{\mu_t} \leq |\mu_t'| \qquad \text{for a.e. } t \in (0, 1). \tag{1.16}$$

Conversely, if (μ_t, v_t) satisfies (1.15) *and $\|v_t\|_{\mu_t} \in L^1(0, 1)$, then (μ_t) is absolutely continuous and*

$$\|v_t\|_{\mu_t} \geq |\mu_t'| \qquad \text{for a.e. } t \in (0, 1). \tag{1.17}$$

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 1.28 satisfies, besides (1.16), also

$$v_t \in \overline{\{\nabla\varphi \, : \, \varphi \in C^\infty_c(M)\}}^{L^2_{\mu_t}} \qquad \text{for a.e. } t \in (0, 1). \tag{1.18}$$

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

Definition 1.29 (Tangent bundle of $\mathscr{P}_2(M)$) *Let $\mu \in \mathscr{P}_2(M)$. We define*

$$\mathrm{Tan}_\mu(\mathscr{P}_2(M)) := \overline{\{\nabla\varphi \,:\, \varphi \in C_c^\infty(M)\}}^{L^2_\mu}.$$

We shall call *tangent velocity* field the vector field v_t provided by Theorem 1.28 and we shall denote by $P_\mu : L^2_\mu \to \mathrm{Tan}_\mu(\mathscr{P}_2(M))$ the orthogonal projection.

The definition of Normal space is then:

$$\begin{aligned}\mathrm{Tan}^\perp_\mu(\mathscr{P}_2(M)) &:= \Big\{ w \in L^2_\mu \;:\; \int \langle w, v\rangle \, d\mu = 0,\ \forall v \in \mathrm{Tan}_\mu(\mathscr{P}_2(M)) \Big\} \\ &= \Big\{ w \in L^2_\mu \;:\; \nabla\cdot(w\mu) = 0 \text{ in the sense of distributions } \Big\}.\end{aligned}$$

It turns out that v_t, besides the metric characterization based on (1.16), has also a differential characterization based on (1.18).

Proposition 1.30 *Let (μ_t, v_t) be such that* (1.15) *holds and $\|v_t\|_{\mu_t} \in L^1(0,T)$. Then v_t satisfies* (1.16) *if and only if $v_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$ for a.e. $t \in (0,1)$.*

Proof Suppose that v_t satisfies (1.16), then (1.17) gives that v_t has for a.e. t minimal norm in $L^2_{\mu_t}$ among the solutions of the continuity equation. Now pick any measurable vector field $t \mapsto w_t \in L^2_{\mu_t}$ such that $w_t \in \mathrm{Tan}^\perp_{\mu_t}(\mathscr{P}_2(M))$ for a.e. $t \in [0,1]$ and observe that since $\nabla\cdot(v_t\mu_t) = \nabla\cdot((v_t + w_t)\mu_t)$, the vector field $t \mapsto v_t + w_t$ is compatible with the curve (μ_t) with respect to the continuity equation. Therefore the minimality of the v_t's ensures that for a.e. $t \in [0,1]$ it holds

$$\|v_t\|^2_{\mu_t} \le \|v_t + w_t\|^2_{\mu_t},$$

which easily implies

$$\int \langle v_t, w_t\rangle \, d\mu_t = 0, \qquad a.e.\ t \in [0,1].$$

By the arbitrariness of $w_t \in \mathrm{Tan}^\perp_{\mu_t}(\mathscr{P}_2(M))$ we get $v_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$ for a.e. $t \in [0,1]$.

Conversely, if $v_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$ for a.e. $t \in [0,1]$ and $\tilde v_t$ is the velocity field satisfying (1.16), then $\nabla\cdot((v_t - \tilde v_t)\mu_t) = 0$ as a space-time distribution. This easily implies that

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

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

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 [2] (see also [5, 25]) 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 [2, 8.4.6] and [2, 8.4.7]). The first result relates the tangent field to the infinitesimal behavior 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 1.31 *Let (μ_t) be an absolutely continuous curve and let (v_t) be its tangent velocity field. Then:*

(i) for a.e. $t \in (0,1)$, for any choice of plans $\gamma_h \in \mathit{OptTan}(\mu_t, \mu_{t+h})$, the rescaled transport plans

$$\tilde{\gamma}_h := \frac{1}{h} \cdot \gamma_h$$

converge in $\mathscr{P}_2(TM)$ to $(Id, v_t)_{\#}\mu_t$ (the definition of rescalation of plans is given in 1.5).

(ii) for all $\sigma \in \mathscr{P}_2(\mathbb{R}^d)$ and a.e. $t \in (0,1)$ we have

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

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

$$\lim_{h \to 0} \frac{T_h}{h} = v_t \qquad \text{in } L^2_{\mu_t}. \tag{1.19}$$

It is interesting to note that when the curve (μ_t) is a geodesic, the velocity vector field is well defined of any $t \in (0,1)$, and not just for a.e. t.

Proposition 1.32 *Let (μ_t) be a geodesic in $\mathscr{P}_2(M)$. Then its velocity vector field (v_t) is the one defined in proposition 1.27. Also, it holds*

$$v_t = \frac{d}{dt}\exp\big((t-s)v_s\big), \qquad \forall t, s \in (0,1). \tag{1.20}$$

Proof Using proposition 1.27 and up to splitting the analysis in the two intervals $[0, \frac{1}{2}]$, $[\frac{1}{2}, 1]$, we can assume that it holds $\mu_t = (\exp(tv))_{\#}\mu_0$ for some $v \in L^2_{\mu_0}$ such that $\exp(v)$ is the unique optimal map from μ_0 to μ_1. Define

$$\tilde{v}_t\Big(\exp\big(tv(x)\big)\Big) := \frac{d}{dt}\exp\big(tv(x)\big), \quad \forall t \in [0,1), \ a.e.\ x \in \operatorname{supp}(\mu_0).$$

Since we know from proposition 1.27 that $\exp(tv)$ is invertible for $t < 1$, the above equation defines a vector field $\tilde{v}_t(x)$ for μ_t-a.e. x. Also, $\tilde{v}_t \in L^2_{\mu_t}$ and $\|\tilde{v}_t\|_{\mu_t} = \|v\|_{\mu}$.

We start by proving that $(\tilde{v}_t)$ is the velocity vector field of (μ_t). Choose $\varphi \in C^\infty_c(M)$ and calculate the derivative of $t \mapsto \int \varphi d\mu_t$ to get:

$$\frac{d}{dt}\int \varphi d\mu_t = \frac{d}{dt}\int \varphi \circ \big(\exp(tv)\big) d\mu = \int \Big\langle \nabla\varphi\Big(\exp\big(tv(x)\big)\Big), \frac{d}{dt}\exp\big(tv(x)\big)\Big\rangle d\mu = \langle \nabla\varphi, \tilde{v}_t \rangle_{\mu_t}.$$

This tells that the choice of the vectors $(\tilde{v}_t)$ is admissible with the continuity equation. To prove that they are of minimal norm, just observe that

$$W_2(\mu_0, \mu_1) = \|v\|_{\mu_0} = \int_0^1 \|\tilde{v}_t\|_{\mu_t} dt.$$

It remains to prove $\tilde{v}_t = v_t$. Consider the map $\exp((s-t)\tilde{v}_t)$: it it clear that $\exp((1-t)\tilde{v}_t)_{\#}\mu_t = \mu_s$. Also, it holds

$$\int \mathrm{d}^2\Big(x, \exp_x\big((s-t)\tilde{v}_t(x)\big)\Big) d\mu_t(x) \leq (s-t)^2\|\tilde{v}_t\|^2_{\mu_t} = (s-t)^2\|v\|^2_{\mu_0} = (s-t)^2 W_2^2(\mu_t, \mu_1).$$

This means that this map is optimal. By the uniqueness part of 1.27 we get the thesis. □

Remark 1.33 (Tangent space at singular measures) In the particular case in which the measure $\mu \in \mathscr{P}_2(M)$ is concentrated on an at most countable set, it holds $\mathrm{Tan}_\mu(\mathscr{P}_2(M)) = L^2_\mu$. Indeed, let $\{x_n\}_{n\in\mathbb{N}}$ be the set where μ is concentrated and choose $v \in L^2_\mu$. By density, it is sufficient to deal with the case of v bounded and compactly supported. Thus with a partition of coordinate argument, it is enough to prove the claim in the case $M = \mathbb{R}^d$ and v bounded. In this situation, let $S := \sup_{n\in\mathbb{N}} |v(x_n)|$ and check that for every $N \in \mathbb{N}$ it is possible to find a function $\varphi^N \in C_c^\infty(\mathbb{R}^d)$ satisfying

$$\nabla\varphi^N(x_n) = v_n(x_n), \qquad \forall n \leq N,$$
$$\sup_{x\in\mathbb{R}^d} |\nabla\varphi^N(x)| \leq S.$$

Letting $N \to +\infty$ and using the dominate convergence theorem we get $\|\nabla\varphi^N - v\|_\mu \to 0$. Since $\nabla\varphi^N \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$, the claim is proved. ■

2 Regular curves

Here we introduce the notions of regular curve, starting from those introduced in [1],[10] and going deeper in the analysis. As said in the introduction, regular curves $t \mapsto \mu_t$ are interesting from two points of view:

- From an *algebraic* point of view, as along a regular curve there is a natural notion of 'translation' of vector fields, i.e. natural isometries $\tau_t^s : L^2_{\mu_t} \to L^2_{\mu_s}$, $t, s \in [0,1]$ (see definition 2.11). As we will see in the next chapter, from these maps it comes out a natural notion of absolute continuity for a vector field defined along a regular curve.

- From a *geometric* point of view, as along a regular curve, the tangent space varies 'smoothly in time' (see theorem 2.13). This property will be the key enabler for the proof of existence of parallel transport that we will give in Chapter 4.

All the curves we consider are defined on the unit interval $[0,1]$, unless otherwise stated.

2.1 Cauchy Lipschitz theory on Riemannian manifolds

Definition 2.1 (The constant $\mathrm{L}(v)$) *Let $\xi \in \mathcal{V}(M)$ and define the constant $\tilde{\mathrm{L}}(\xi)$ as*

$$\tilde{\mathrm{L}}(\xi) := \sup_{x \in M} \|\nabla \xi(x)\|_{\mathrm{op}},$$

which is finite by regularity and compactness of support. Then for $\mu \in \mathscr{P}_2(M)$ and $v \in L^2_\mu$ we define $\mathrm{L}(v)$ as:

$$\mathrm{L}(v) := \inf \varliminf_{n \to +\infty} \tilde{\mathrm{L}}(\xi_n),$$

where the infimum is taken among all the sequences $(\xi_n) \subset \mathcal{V}(M)$ converging to v in L^2_μ.

We will say that a vector field $v \in L^2_\mu$ is Lipschitz if $\mathrm{L}(v) < \infty$, and will call $\mathrm{L}(v)$ the Lipschitz constant of v. A simple application of the Ascoli-Arzelà theorem gives that if $\mathrm{L}(v) < \infty$, then there exists a sequence $(\xi_n) \subset \mathcal{V}(M)$ with equibounded Lipschitz constant, which converges to v in L^2_μ and locally uniformly to some vector field $\tilde{v}$ defined on the whole M (which therefore μ-a.e. coincides with v). In particular, any Lipschitz vector field has a continuous representative well defined on the whole $\mathrm{supp}(\mu)$: when dealing with Lipschitz vector fields we will always implicitly assuming to deal with this continuous representative.

In particular, if μ has compact support and $\mathrm{L}(v) < \infty$, then v is bounded. Since we will need this bound later on on the work, let us underline this point with the following definition:

Definition 2.2 (Supremum of Lipschitz vector fields) *Let $\mu \in \mathcal{P}_c(M)$ and $v \in L^2_\mu$ with $\mathrm{L}(v) < \infty$. We will denote by $\mathrm{S}(v)$ the supremum (actually, maximum) of $|v|$ on $\mathrm{supp}(\mu)$.*

A word on notation: in writing $\mathrm{L}(v)$ we are losing the reference to the base measure μ. In most cases, this is not at an issue, as it will be clear from the context who is the measure. However, the value of $\mathrm{L}(\xi)$ may be less clear when $\xi \in \mathcal{V}(M)$ is a smooth vector field defined on all the manifold. The convention we will use is $\mathrm{L}(\xi) := \sup_{x \in M} \|\nabla \xi(x)\|_{\mathrm{op}}$ for these kind of vectors.

Remark 2.3 (The case $M = \mathbb{R}^d$) In the case $M = \mathbb{R}^d$, and $\xi \in \mathcal{V}(M)$, $\mathrm{L}(\xi) = \sup_{x \in M} \|\nabla \xi(x)\|_{\mathrm{op}}$ is precisely the Lipschitz constant of ξ. ■

Before passing to the study of regular curves, we want to discuss an important example which we will need to keep in mind in our analysis. The setting is the following: suppose we have $v \in L^2_\mu$ which is both tangent and Lipschitz. Then we know that there exists a sequence of functions $(\varphi^n) \subset C^\infty_c(M)$ such that $\|\nabla \varphi^n - v\|_\mu \to 0$ and a sequence of vector fields $(\xi^n) \subset \mathcal{V}(M)$ such that $\|\xi^n - v\|_\mu \to 0$ and $\mathrm{L}(\xi^n) \to \mathrm{L}(v)$. A natural question which arises is: can we 'combine' these convergences? In other words, is that true that we can find a sequence $(\varphi^n) \subset C^\infty_c(M)$ such that $\|\nabla \varphi^n - v\|_\mu \to 0$ and $\mathrm{L}(\nabla \varphi^n) \to \mathrm{L}(v)$? The answer is no: actually it may happen that for any sequence of gradients $(\nabla \varphi^n)$ approximating v in L^2_μ the Lipschitz constant $\mathrm{L}(\nabla \varphi^n)$ diverges as $n \to \infty$.

Example 2.4 Let $M = \mathbb{R}^2$ and define, for every $n \in \mathbb{N}$, the set $A_n \subset \mathbb{R}^2$ as

$$A_n := \Big\{ e^{2\pi i \frac{k}{2^n}} \ : \ k = 1, \ldots, 2^n \Big\},$$

where we identified $\mathbb{R}^2$ with the set of complex numbers (in the above formula $i = \sqrt{-1}$). Let $A := \cup_n A_n$ and μ be a measure concentrated on A which gives positive mass to each point of A. Define

$$v(x, y) := (-y, x).$$

Clearly v belongs to L^2_μ and $\mathrm{L}(v) \leq 1$. Also, since μ has countable support, any vector field in L^2_μ is tangent (see remark 1.33), so is v.

Now let $(\varphi_n) \subset C^\infty_c(\mathbb{R}^2)$ be such that $\|v - \nabla\varphi_n\|_\mu \to 0$. In particular, given the atomic structure of μ, $\nabla\varphi_n(x) \to v(x)$ for any $x \in A$. Thus for every $n \in \mathbb{N}$, there exists N_n such that $|\nabla\varphi_{N_n}(x) - v(x)| \leq 1/2$ for any $x \in A_n$. Define the curve $[0, 1] \ni t \mapsto \gamma(t) := e^{2\pi i t}$ and observe that since $\nabla\varphi_{N_n}$ is a gradient, it holds

$$\int_0^1 \big\langle v\big(\gamma(t)\big), \nabla\varphi_{N_n}\big(\gamma(t)\big)\big\rangle \, dt = \int_0^1 \big\langle \gamma'(t), \nabla\varphi_{N_n}\big(\gamma(t)\big)\big\rangle \, dt = 0, \qquad \forall n \in \mathbb{N}.$$

Therefore there must exists some t_0 such that $\big\langle v\big(\gamma(t_0)\big), \nabla\varphi_{N_n}\big(\gamma(t_0)\big)\big\rangle \leq 0$, which, together with $|v(\gamma(t_0))| = 1$, implies that $\big|v\big(\gamma(t_0)\big) - \nabla\varphi_{N_n}\big(\gamma(t_0)\big)\big| \geq 1$. Let x_0 be the nearest point to t_0 among those in A_{N_n} and observe that it holds $|t_0 - x_0| \leq \pi 2^{-N_n}$. Thus

$$\begin{aligned}
\mathrm{Lip}(\nabla\varphi_{N_n}) &\geq \frac{|\nabla\varphi_{N_n}(t_0) - \nabla\varphi_{N_n}(x_0)|}{|t_0 - x_0|} \\
&\geq \frac{|\nabla\varphi_{N_n}(t_0) - v(t_0)| - |v(t_0) - v(x_0)| - |v(x_0) - \nabla\varphi_{N_n}(x_0)|}{|t_0 - x_0|} \\
&\geq \frac{1 - \frac{\pi}{2^{N_n}} - \frac{1}{2}}{\frac{\pi}{2^{N_n}}} = \frac{2^{N_n}}{2\pi} - 1,
\end{aligned}$$

therefore $\mathrm{L}(\nabla\varphi_n) = \mathrm{Lip}(\nabla\varphi_n) \to +\infty$ as $n \to +\infty$, which is our claim.

Observe that vector fields like the one just described, may appear as velocity vector field of an absolutely continuous curve. Indeed, consider the map $Rot_t : \mathbb{R}^2 \to \mathbb{R}^2$ given by the counter-clockwise rotation of an angle t around the origin and let $\mu_t := (Rot_t)_\#\mu$, with μ as above. It is immediate to check that this curve is absolutely continuous, with length 1, and that for every time its velocity vector field is given by the rotation of the vector v described above. ■

We conclude recalling some well known facts about the Cauchy Lipschitz theory on manifolds.

Definition 2.5 (Transport couples and their convergence) *We will call* transport couple *a map $[0, 1] \ni t \mapsto (\mu_t, v_t)$ with $\mu_t \in \mathscr{P}_2(M)$ and $v_t \in \mathrm{L}^2_{\mu_t}$ such that $t \mapsto \|v_t\|_{\mu_t}$ is integrable and the continuity equation*

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

holds in the sense of distribution. Thus a transport couple is just an absolutely continuous curve for which we specified a particular vector field, not necessarily tangent. We say that a sequence of transport couples $n \mapsto (\mu_t^n, v_t^n)$ converges to (μ_t, v_t) if the following three things happen:

- $W(\mu_t^n, \mu_t) \to 0$ *as* $n \to \infty$ *uniformly on* $t \in [0,1]$,
- v_t^n *converges strongly to* v_t *in the sense of definition 1.8 as* $n \to \infty$ *for a.e.* $t \in [0,1]$
- $\int_0^1 \|v_t^n\|_{\mu_t^n} dt \to \int_0^1 \|v_t\|_{\mu_t} dt$ *as* $n \to \infty$.

We recall the following theorem.

Theorem 2.6 (Cauchy-Lipschitz on manifolds) *Let $[0,1] \ni t \mapsto (\mu_t, v_t)$ be a transport couple. Assume that*

$$\int_0^1 \mathrm{L}(v_t) < \infty.$$

Then:

i) There exists a unique family of maps $\mathbf{T}(t,s,\cdot) : \mathrm{supp}(\mu_t) \to \mathrm{supp}(\mu_s)$, $t,s \in [0,1]$, such that the curve $s \mapsto \mathbf{T}(t,s,x)$ is absolutely continuous for every $t \in [0,1]$, $x \in \mathrm{supp}(\mu_t)$ and satisfying:

$$\begin{cases} \mathbf{T}(t,t,x) = x, & \forall t \in [0,1],\ x \in \mathrm{supp}(\mu_t) \\ \dfrac{d}{dr}\mathbf{T}(t,r,x)_{|r=s} = v_s\big(\mathbf{T}(t,s,x)\big), & \forall t \in [0,1],\ x \in \mathrm{supp}(\mu_t),\ a.e.\ s \in [0,1] \end{cases} \tag{2.2}$$

Also, these maps satisfy:

$$\mathbf{T}\big(r,s,\mathbf{T}(t,r,x)\big) = \mathbf{T}(t,s,x), \qquad \mathbf{T}(t,s,\cdot)_{\#}\mu_t = \mu_s$$

for every $t,s,r \in [0,1]$ and $x \in \mathrm{supp}(\mu_t)$.

ii) If the vectors v_t are defined on the whole M and for every t, and $(v_t) \in \mathcal{V}(M \times [0,1])$, the maps $\mathbf{T}(t,s,\cdot)$ are defined on the whole M, $(t,s,x) \mapsto \mathbf{T}(t,s,x)$ is C^∞, and equations (2.2) hold for every choice of $t,s \in [0,1]$ and $x \in M$. Furthermore, there exists a compact set $K \subset M$ such that $\mathbf{T}(t,s,x) = x$ for any $x \notin K$ and any $t,s \in [0,1]$.

iii) (μ_t) may be approximated in the sense of 2.5 by a sequence of transport couples $n \mapsto (\mu_t^n, v_t^n)$, $n \in \mathbb{N}$, such that the vector fields (v_t^n) satisfy the conditions in (ii) and

$$\int_0^1 \Big|\mathrm{L}(v_t^n) - \mathrm{L}(v_t)\Big| dt \to 0,$$
$$\int_0^1 \Big|\|v_t^n\|_{\mu_t^n} - \|v_t\|_{\mu_t}\Big| dt \to 0,$$
$$\int_0^1 \|v_t^n - v_t\|_{\mu_t} dt \to 0,$$

as $n \to \infty$. Also, such transport couples may be chosen to satisfy $\mu_0^n = \mu_0$ for every $n \in \mathbb{N}$.

It is worth underlying that it is *not* part of the approximation result of (*iii*) the fact that the approximating vector fields v_t^n are tangent, not even in the case in which the v_t's are. Approximating with smooth and tangent vector fields is actually possible (we will do this in the appendix 9.1), but requires a quite heavy theoretical machinery - at least in our approach.

2.2 Definition and first properties of regular curves

Just a word on notation: when considering curves of measures, we will write $t \mapsto \mu_t$ or (μ_t) for the curve and μ_t for the single measure which is the value at t of the curve. Given the curve (μ_t), we will always denote by $t \mapsto v_t$ or (v_t) its velocity vector field, and by v_t the value of this vector field at the point t, unless otherwise stated. Recall that (v_t) is identified, up to equality $t-$a.e., among the other solutions of the continuity equation by $v_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$ for a.e. t.

Definition 2.7 (Regular curves) *Let (μ_t) be an absolutely continuous curve on $[0,1]$. We say that (μ_t) is regular if its velocity vector field (v_t) satisfies*

$$\int_0^1 \|v_t\|_{\mu_t}^2 dt < \infty,$$

and

$$\int_0^1 \mathrm{L}(v_t) dt < \infty.$$

It is key in this definition the fact that the vector fields considered are the tangent ones: we will understand why in remark 5.4.

The important part of the definition is the integrability of the Lipschitz constant. The requirement on the square-integrability of the norms of the v_t's (as opposed to plain integrability, which is ensured by the absolute continuity of (μ_t)) is just a 'light' technical assumption: observe that the quantity $\int_0^1 \mathrm{L}(v_t)dt$ is independent on the parametrization of the curve, so if it is finite, we can always reparametrize the curve in order to have $\int_0^1 \|v_t\|_{\mu_t}^2 dt < \infty$.

Also, observe that the regularity of the curve (μ_t) has nothing to do with properties of the underlying measures μ_t's (like absolute continuity).

Whenever (μ_t) is a regular curve and (v_t) its tangent velocity vector field, we will call the associated maps $\mathbf{T}(t,s,\cdot)$ given by theorem 2.6 the *flow maps of* μ_t. Observe that it holds

$$\begin{aligned}\int \mathrm{d}\big(\mathbf{T}(t,s,x),\mathbf{T}(t,s',x)\big)^2 d\mu_t(x) &\leq \int\left(\int_s^{s'}\big|v_r\big(\mathbf{T}(t,r,x)\big)\big|\,dr\right)^2 d\mu_t(x)\\ &\leq (s'-s)\int_s^{s'}\int\big|v_r\big(\mathbf{T}(t,r,x)\big)\big|^2 d\mu_t\,dr\\ &= (s'-s)\int_s^{s'}\|v_r\|^2_{\mu_r}dr, \qquad \forall 0\leq s\leq s'\leq 1.\end{aligned}\tag{2.3}$$

Definition 2.8 (The maps $(\tau_x)_s^t$) *Let (μ_t) be a regular curve and $\mathbf{T}(t,s,\cdot)$ its flow maps. Given $t,s\in[0,1]$ and $x\in\mathrm{supp}(\mu_t)$, we let $(\tau_x)_s^t: T_{\mathbf{T}(t,s,x)}M\to T_xM$ be the map which associate to $v\in T_{\mathbf{T}(t,s,x)}M$ its parallel transport along the absolutely continuous curve $r\mapsto\mathbf{T}(t,r,x)$ from $r=s$ to $r=t$.*

Definition 2.9 (The constant $\mathrm{L}(\mathbf{T}(t,s,\cdot))$) *Let (μ_t) be a regular curve, (v_t) its velocity vector field and $\mathbf{T}(t,s,\cdot)$ its flow maps. Given $t,s\in[0,1]$ we define the constant $\mathrm{L}(\mathbf{T}(t,s,\cdot))\in[0,+\infty)$ by*

$$\mathrm{L}(\mathbf{T}(t,s,\cdot)) := \inf\varliminf_{n\to\infty}\sup_{x\in\mathrm{supp}(\mu_t^n)}\|\nabla(\mathbf{T}^n(t,s,\cdot))(x)-(\tau_x^n)_t^s\|_{\mathrm{op}},$$

where $\|\cdot\|_{\mathrm{op}}$ is the operator norm, the infimum is taken among all the transport couples $n\mapsto(\mu_t^n,v_t^n)$ converging to (μ_t,v_t) as defined in 2.5 and satisfying $(v_t^n)\in\mathcal{V}(M\times[0,1])$, and $\mathbf{T}^n(t,s,\cdot)$ are the maps associated to such curves.

It is easy to check that if the vector field (v_t) of (μ_t) is made of smooth vectors and $\mathrm{supp}(\mu_t)=M$, then $\mathrm{L}(\mathbf{T}(t,s,\cdot))$ is exactly the supremum of $\|\nabla(\mathbf{T}(t,s,\cdot))(x)-(\tau_x)_t^s\|_{\mathrm{op}}$ among all $x\in M$. Furthermore, in the case $M=\mathbb{R}^d$ and $\mathrm{supp}(\mu_t)=\mathbb{R}^d$, it holds $\mathrm{L}(\mathbf{T}(t,s,\cdot))=\mathrm{Lip}\big(\mathbf{T}(t,s,\cdot)-Id\big)$.

The following bounds are just consequences of the definition:

Proposition 2.10 *Let (μ_t) be a regular curve and $\mathbf{T}(t,s,\cdot)$ its flow maps. Then for any $t,s\in[0,1]$ it holds:*

$$\mathrm{Lip}(\mathbf{T}(t,s,\cdot))\leq e^{\left|\int_t^s\mathrm{L}(v_r)dr\right|}\tag{2.4a}$$

$$\mathrm{L}(\mathbf{T}(t,s,\cdot))\leq e^{\left|\int_t^s\mathrm{L}(v_r)dr\right|}-1,\tag{2.4b}$$

where $\mathrm{Lip}(\mathbf{T}(t,s,\cdot))$ *is the Lipschitz constant of* $\mathbf{T}(t,s,\cdot):(\mathrm{supp}(\mu_t),\mathrm{d})\to(\mathrm{supp}(\mu_s),\mathrm{d})$.

Proof It is enough to prove the statement for the case $t\leq s$. Let us assume at first that the curve (μ_t) satisfies the regularity assumption of *(ii)* of theorem 2.6, so that the map $\mathbf{T}(t,s,\cdot)$ is C^∞. In this case it clearly holds $\mathrm{Lip}(\mathbf{T}(t,s,\cdot))=\sup_{x\in\mathrm{supp}(\mu_t)}\|\nabla(\mathbf{T}(t,s,\cdot))(x)\|_{\mathrm{op}}$. Taking the gradient in $\frac{d}{ds}\mathbf{T}(t,s,\cdot)=v_s(\mathbf{T}(t,s,\cdot))$ we get:

$$\frac{d}{ds}\nabla\mathbf{T}(t,s,\cdot)=\nabla\Big(v_s\big(\mathbf{T}(t,s,\cdot)\big)\Big)=\Big((\nabla v_s)\circ\mathbf{T}(t,s,\cdot)\Big)\cdot\big(\nabla\mathbf{T}(t,s,\cdot)\big).$$

Taking the norms this leads to

$$\frac{d}{ds}\|\nabla\mathbf{T}(t,s,\cdot)\|_{\rm op} \leq \|\nabla v_s\|_{\rm op}\|\nabla\mathbf{T}(t,s,\cdot)\|_{\rm op},$$

from which it follows, by the Gronwall lemma and the fact that $\mathrm{Lip}(\mathbf{T}(t,t,\cdot)) = 1$, equation (2.4a).

To prove (2.4b), fix $t \in I$, $x \in \mathrm{supp}(\mu_t)$, $v \in T_xM$ such that $|v|_x = 1$ and consider the function $s \mapsto |((\nabla\mathbf{T}(t,s,\cdot))(x) - (\tau_x)_t^s)(v)|^2$. Its derivative is given by:

$$\frac{d}{ds}\left|\Big(\big(\nabla\mathbf{T}(t,s,\cdot)\big)(x) - (\tau_x)_t^s\Big)(v)\right|^2 = 2\left\langle\Big(\nabla\mathbf{T}(t,s,\cdot)(x) - (\tau_x)_t^s\Big)(v), \Big(\nabla v_s(\mathbf{T}(t,s,\cdot))(x)\Big)(v)\right\rangle.$$

Taking the norms, the supremum over v and x and using (2.4a) this gives

$$\frac{d}{ds}\mathrm{L}(\mathbf{T}(t,s,\cdot)) \leq \mathrm{L}(v_s)e^{\int_t^s \mathrm{L}(v_r)dr}$$

from which it follows (2.4b) by integration.

The case for general maps $\mathbf{T}$ follows from part *(iii)* of theorem 2.6 and from the definition of $\mathrm{L}(\mathbf{T}(t,s,\cdot))$. □

As said, the importance of regular curves comes from the algebraic point that there are well defined translation maps between the L^2 spaces at different measures and on the geometric fact that the *angle between tangent spaces* varies smoothly along them.

We already saw a bit of the algebraic point of view in the definition of the maps $(\tau_x)_s^t$: now we define to the 'global' version of these maps:

Definition 2.11 (Translation of vectors along a regular curve) *Let (μ_t) be a regular curve, $\mathbf{T}(t,s,\cdot)$ its flow maps and $u \in \mathrm{L}^2_{\mu_s}$. We define the translation $\tau_s^t(u) \in \mathrm{L}^2_{\mu_t}$ of u along the curve (μ_t) up to the measure μ_t as:*

$$\tau_s^t(u)(x) := \begin{cases} \textit{the parallel transport of } u(\mathbf{T}(t,s,x)) \textit{ along the curve} \\ r \mapsto \mathbf{T}(t,r,x) \textit{ from } r = s \textit{ to } r = t. \end{cases}$$

In formula, the above definition may be written as:

$$\tau_s^t(u)(x) = (\tau_x)_s^t(u(\mathbf{T}(t,s,x)))$$

It is an immediate consequence of the fact that the parallel transport is norm preserving and of the identity $\mathbf{T}(s,t,\cdot)_\#\mu_s = \mu_t$, the fact that the map τ_s^t in an isometry from $\mathrm{L}^2_{\mu_s}$ to $\mathrm{L}^2_{\mu_t}$ for any $t,s \in [0,1]$. Furthermore, from the group property of the flow maps, it follows

$$\tau_s^t \circ \tau_r^s = \tau_r^t \tag{2.5}$$

for every $t,r,s \in [0,1]$.

It is worth underlying that in general nothing ensures that $\tau_s^t(u)$ is tangent if u is.

Remark 2.12 (The case $M = \mathbb{R}^d$) In the case $M = \mathbb{R}^d$ the translation maps are nothing but the composition with the flow maps. This may help understanding why $\tau_s^t(u)$ may be not tangent if u is. Indeed, let $u = \nabla\varphi$ for some $\varphi \in C_c^\infty(\mathbb{R}^d)$: we have

$$\tau_t^s(\nabla\varphi) = (\nabla\varphi) \circ \mathbf{T}(t, s, \cdot),$$

and clearly there is no reason for the right hand side to be a gradient. ■

Before passing to the analysis of the geometric point of view, let us describe the heuristic idea. We just said that to a given tangent vector $u \in \mathrm{Tan}_{\mu_s}(\mathscr{P}_2(M))$, we can naturally associate the vector $\tau_s^t(u)$ in $\mathrm{L}^2_{\mu_t}$. Such a vector is certainly a well defined vector in $\mathrm{L}^2_{\mu_t}$, but, as said, in general nothing ensures that it is tangent. However, if 'the angle varies smoothly', we can hope that the distance from this vector to the tangent space $\mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$ is controlled uniformly in $\|u\|_{\mu_s}$ by a function of s, t which goes to 0 as t tends to s. This is actually the case, as we are going to show now: this result is the main justification for the introduction of regular curves.

Theorem 2.13 *Let (μ_t) be a regular curve, $\mathbf{T}(t, s, \cdot)$ its flow maps and $u_s \in \mathrm{Tan}_{\mu_s}(\mathscr{P}_2(M))$. Then it holds:*

$$\|\mathrm{P}^\perp_{\mu_t}(\tau_s^t(u_s))\|_{\mu_t} \le \mathrm{L}(\mathbf{T}(t, s, \cdot))\|u_s\|_{\mu_s}. \tag{2.6}$$

Proof The statement is equivalent to

$$\|\tau_s^t(\nabla\varphi) - \mathrm{P}_{\mu_t}(\tau_s^t(\nabla\varphi))\|_{\mu_t} \le \left(e^{\left|\int_t^s \mathrm{L}(v_r)dr\right|} - 1\right)\|\nabla\varphi\|_{\mu_s}, \qquad \forall \varphi \in C_c^\infty(M). \tag{2.7}$$

Thus to prove the theorem we need to find a smooth function $\psi \in C_c^\infty(M)$ such that $\|\tau_s^t(\nabla\varphi) - \nabla\psi\|_{\mu_t}$ is small. Let us assume for a moment that the vectors v_t's satisfy the regularity assumption in *(ii)* of theorem 2.6. In this case our candidate is $\psi := \varphi \circ \mathbf{T}(t, s, \cdot)$, which belongs to $C_c^\infty(M)$, since the map $\mathbf{T}(t, s, \cdot)$ is different from the identity only in a compact set. The gradient of ψ is given by $\nabla\psi = \big(\nabla\mathbf{T}(t, s, \cdot)\big)^t \cdot \nabla\varphi \circ \mathbf{T}(t, s, \cdot)$, thus we have

$$\begin{aligned}
\|\tau_s^t(\nabla\varphi) - \mathrm{P}_{\mu_t}(\tau_s^t(\nabla\varphi))\|_{\mu_t} &\le \|\tau_s^t(\nabla\varphi) - \nabla\psi\|_{\mu_t} \\
&= \sqrt{\int \left|(\tau_x)_s^t\Big(\nabla\varphi\big(\mathbf{T}(t, s, x)\big)\Big) - \big(\nabla\mathbf{T}(t, s, x)\big)^{\mathrm{t}} \cdot \nabla\varphi(\mathbf{T}(t, s, x))\right|^2 d\mu_t(x)} \\
&= \sqrt{\int \left|\Big((\tau_x)_s^t - \big(\nabla\mathbf{T}(t, s, x)\big)^{\mathrm{t}}\Big)\big(\nabla\varphi(\mathbf{T}(t, s, x))\big)\right|^2 d\mu_t(x)} \\
&\le \sqrt{\int \left\|(\tau_x)_s^t - \big(\nabla\mathbf{T}(t, s, x)\big)^{\mathrm{t}}\right\|^2_{\mathrm{op}} \left|\nabla\varphi\big(\mathbf{T}(t, s, x)\big)\right|^2 d\mu_t(x)} \\
&= \sqrt{\int \left\|(\tau_x)_t^s - \nabla\mathbf{T}(t, s, x)\right\|^2_{\mathrm{op}} \left|\nabla\varphi\big(\mathbf{T}(t, s, x)\big)\right|^2 d\mu_t(x)} \\
&\le \mathrm{L}(\mathbf{T}(t, s, \cdot))\sqrt{\int \left|\nabla\varphi\big(\mathbf{T}(t, s, x)\big)\right|^2 d\mu_t(x)} \\
&= \mathrm{L}(\mathbf{T}(t, s, \cdot))\|\nabla\varphi\|_{\mu_s},
\end{aligned} \tag{2.8}$$

which implies, thanks to (2.4b), equation (2.7).

To remove the smoothness assumption on the velocity vector fields, it is enough to use the approximation result in (iii) of theorem 2.6. Indeed, fix $t, s \in [0,1]$ and let (μ_t^n, v_t^n) be a sequence of transport couples satisfying the assumptions in (ii) of theorem 2.6, converging to (μ_t, v_t) in the sense of 2.5 and such that

$$\mathrm{L}(\mathbf{T}(t,s,\cdot)) = \lim_{n\to\infty} \sup_{x\in \mathrm{supp}(\mu_t^n)} \|\nabla(\mathbf{T}(t,s,\cdot))(x) - (\tau_x^n)_t^s\|_{\mathrm{op}}, \tag{2.9}$$

where $\mathbf{T}^n(t,s,\cdot)$ are the maps associated to (μ_t^n, v_t^n). Under this assumptions, it is not hard to check that

$$\begin{aligned} \lim_{n\to\infty} (\tau^n)_s^t(\nabla\varphi) &= \tau_s^t(\nabla\varphi), \qquad \text{in } L^2_{\mu_t}, \\ \lim_{n\to\infty} \mathbf{T}^n(t,s,\cdot)_\# \mu_t &= \mu_s \qquad \text{in } (\mathscr{P}_2(M), W_2). \end{aligned} \tag{2.10}$$

Now let $\psi^n := \varphi \circ \mathbf{T}^n(t,s,\cdot)$, observe that $\psi^n \in C_c^\infty(M)$ and therefore

$$\begin{aligned} \|\tau_s^t(\nabla\varphi) - \mathrm{P}_{\mu_t}(\tau_s^t(\nabla\varphi))\|_{\mu_t} &\le \|\tau_s^t(\nabla\varphi) - \nabla\psi^n\|_{\mu_t} \\ &\le \|\tau_s^t(\nabla\varphi) - (\tau^n)_s^t(\nabla\varphi)\|_{\mu_t} + \|(\tau^n)_s^t(\nabla\varphi) - \nabla\psi^n\|_{\mu_t}. \end{aligned}$$

Arguing as in (2.8) we obtain

$$\|(\tau^n)_s^t(\nabla\varphi) - \nabla\psi^n\|_{\mu_t} \le \mathrm{L}(\mathbf{T}^n(t,s,\cdot))\|\nabla\varphi\|_{\mathbf{T}^n(t,s,\cdot)_\#\mu_t}.$$

Therefore using (2.9) and (2.10) we obtain

$$\begin{aligned} \|\tau_s^t(\nabla\varphi) - \mathrm{P}_{\mu_t}(\tau_s^t(\nabla\varphi))\|_{\mu_t} &\le \varlimsup_{n\to\infty} \Big(\|\tau_s^t(\nabla\varphi) - (\tau^n)_s^t(\nabla\varphi)\|_{\mu_t} + \mathrm{L}(\mathbf{T}^n(t,s,\cdot))\|\nabla\varphi\|_{\mathbf{T}^n(t,s,\cdot)_\#\mu_t} \Big) \\ &\le \mathrm{L}(\mathbf{T}(t,s,\cdot))\|\nabla\varphi\|_{\mu_s}, \end{aligned}$$

which is the thesis. □

Remark 2.14 (Topological restriction) If (μ_t) is a regular curve, then all the supports $\mathrm{supp}(\mu_t)$ are homeomorphic. Indeed, the flow maps $\mathbf{T}(t,s,\cdot)$ are Lipschitz, and satisfy

$$\mathbf{T}\big(t,s,\mathrm{supp}(\mu_t)\big) = \mathrm{supp}(\mu_s),$$

thus, since their inverse is Lipschitz as well, they provide a bi-Lipschitz homeomorphism of the supports. ■

Having introduced the notion of regular curve, it is natural to ask whether these curves are dense or not, and whether geodesics are regular or not. The answer to the first question is affirmative, and actually it holds a more general result, according to which any transport couple may be approximated by regular curves made of 'smooth measures with smooth velocity'. We postpone the proof of this fact to the appendix, as we won't need this result in the rest of the work and the proof is a bit technical.

In the next section we discuss the problem of regularity of geodesics.

2.3 On the regularity of geodesics

It is worth underlying from the beginning that from remark 2.14 it follows that in general geodesics are not regular: it is sufficient to choose μ_0, μ_1 such that their supports are not homeomorphic to have that no geodesic between them is regular. Thus the best we can hope is to replicate the result valid in the case $M = \mathbb{R}^d$, where the restriction of a geodesic on $[0,1]$ to an interval of the kind $[\varepsilon, 1-\varepsilon]$ is always regular (see [1], [10]). This is what we will prove here for the case of a geodesic connecting two measures with compact support. The general case seems to be more tricky, see remark 2.24. Still, observe that as soon as we prove that restriction of geodesics between measures with compact support are regular, we easily derive that the class of regular geodesic is dense in the class of all geodesics in $\mathscr{P}_2(M)$.

The proof of regularity of restriction of geodesics is a consequence of the following proposition, due to Fathi (see appendix of [8]), which we restate in our terminology:

Proposition 2.15 *For every compact set $K \subset M$ there exists a constant C, such that the following is true. Assume that $\mu \in \mathcal{P}_c(M)$ is concentrated on K and $\varphi, \psi : M \to \mathbb{R}$ satisfy: φ is -1-concave, ψ is -1-convex, $\varphi \geq \psi$ on M and $\varphi = \psi$ on $\operatorname{supp}(\mu)$. Then the vector field defined by $v(x) = \nabla\varphi(x) = \nabla\psi(x)$ for every $x \in \operatorname{supp}\mu$ belongs to L^2_μ (the fact that v is well defined comes from an argument similar to that of proposition 1.26) and $\mathrm{L}(v) \leq C$.*

Idea of the proof. Use a partition of the unit subordinate to a cover by charts to reduce the problem to a problem in $\mathbb{R}^d$. The fact that K is compact, and thus only a finite number of coordinate charts are needed to cover it, ensures that the functions φ, ψ read in each of the charts are $-D$-concave and $-D$-convex respectively, for some $D \in \mathbb{R}$.

Now look at the problem in $\mathbb{R}^d$. Let $A := \{\varphi = \psi\}$, so that v is well defined on A. From the hypothesis we have that for any $x \in A$, $y \in \mathbb{R}^d$ it holds:

$$\langle v(x), y - x\rangle - \frac{D}{2}|x-y|^2 \leq \psi(y) - \psi(x) \leq \varphi(y) - \varphi(x) \leq \langle v(x), y-x\rangle - \frac{D}{2}|x-y|^2.$$

Therefore

$$|\varphi(y) - \varphi(x) - \langle v(x), y-x\rangle\,| \leq \frac{D}{2}|y-x|^2, \qquad \forall x \in A,\ y \in \mathbb{R}^d.$$

In particular, for any couple $x_1, x_2 \in A$ and any $y \in \mathbb{R}^d$ we have

$$\begin{aligned} \varphi(x_2) - \varphi(x_1) - \langle v(x_1), x_2 - x_1\rangle &\leq \frac{D}{2}|x_2 - x_1|^2, \\ \varphi(x_2 + y) - \varphi(x_2) - \langle v(x_2), y\rangle &\leq \frac{D}{2}|y|^2, \\ -\varphi(x_2+y) + \varphi(x_1) + \langle v(x_1), x_2 + y - x_1\rangle &\leq \frac{D}{2}|x_2 + y - x_1|^2. \end{aligned}$$

Adding up these inequalities we obtain

$$\begin{aligned} \langle v(x_1), y\rangle - \langle v(x_2), y\rangle &\leq \frac{D}{2}\Big(|x_2 - x_1|^2 + |y|^2 + |x_2 + y - x_1|^2\Big), \\ &\leq \frac{3D}{2}\Big(|x_2 - x_1|^2 + |y|^2\big). \end{aligned}$$

Choosing $y = \frac{v(x_1)-v(x_2)}{3D}$ we obtain

$$|v(x_1) - v(x_2)|^2 \leq 9D^2|x_1 - x_2|^2.$$

Reading back this result in the manifold we obtain the thesis. $\square$

Notice that the value of the constant C provided by Fathi's argument is, in some sense, not intrinsic as it depends on the partition of the unit chosen to put the compact K into charts.

An immediate consequence of Fathi's proposition is the regularity of the restriction of geodesics:

Corollary 2.16 (Regularity of restriction of geodesics) *Let (μ_t) be a geodesic in $\mathcal{P}_c(M)$. Then its restriction to any interval of the kind $[\varepsilon, 1-\varepsilon]$, $0 < \varepsilon < \frac{1}{2}$ is regular.*

Proof Let $K \subset M$ be a compact set whose interior contains the support of all the μ_t's and C the constant associated to it via Fathi's result. We know from propositions 1.27 and 1.32 that the velocity vector field (v_t) of (μ_t) is given by the formula

$$v_t(x) = \nabla\varphi_t(x) = \nabla\psi_t(x), \qquad \forall t \in (0,1)\ x \in \mathrm{supp}(\mu_t),$$

where ψ_t is $\frac{\mathcal{H}(K)}{t}$-concave and φ_t is $\frac{\mathcal{H}(K)}{1-t}$-convex on K, $\psi_t \geq \varphi_t$ and $\psi_t = \varphi_t$ in $\mathrm{supp}(\mu_t)$. Then by the result of Fathi we deduce:

$$\mathrm{L}(v_t) \leq C\frac{\mathcal{H}(K)}{\min\{t, 1-t\}},$$

whence the thesis follows. $\square$

Now that regularity of restriction of geodesic is proven, we pass to a more detailed analysis of the Lipschitz constant of the velocity vector fields. The result we want to prove is the following: for any $t \in (0,1)$ there exists a sequence $n \mapsto \varphi_t^n \in C_c^\infty(M)$ of functions such that $\|\nabla\varphi_t^n - v_t\|_{\mu_t} \to 0$ as $n \to \infty$ and $\sup_n \mathrm{L}(\nabla\varphi_t^n) < \infty$. Because of example 2.4, this fact is not a consequence of $\mathrm{L}(v_t) < \infty$ and $v_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$. Our argument does *not* improve the value of the constant C provided by Fathi's result, as we will do a similar reduction to coordinate charts.

We advice the reader that we won't need the result proven below until the end of the work, where we study Jacobi fields, thus those more interested in the calculus with vector fields on $(\mathscr{P}_2(M), W_2)$ may skip this part at a first reading.

Our strategy consists in reducing the problem to a problem in $\mathbb{R}^d$ and then in showing that whenever we have a -1-concave function ψ and a -1-convex function φ satisfying $\psi \geq \varphi$ on the whole $\mathbb{R}^d$, then there exists a third function ϕ which is both -1-convex and -1-concave and satisfies:

$$\psi \geq \phi \geq \varphi,$$

on the whole $\mathbb{R}^d$. Thus the vector field v defined as $\nabla\psi = \nabla\varphi$ on the contact set, may be actually seen as the gradient of ϕ on the whole $\mathbb{R}^d$ (as since ϕ is both -1-concave and -1-convex, it is differentiable everywhere with Lipschitz gradient). Thus to approximate v with gradients of smooth functions we can simply consider the gradients of some functions ϕ^n converging to ϕ in a sufficiently smooth way.

Our proof is based on a purely geometrical result valid for functions on $\mathbb{R}^d$. To state it, we need the following definition.

Definition 2.17 (Upper and lower envelope) *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{\pm\infty\}$. The upper and lower envelopes* $\mathrm{Upp}(f)$*,* $\mathrm{Low}(f)$ *of f are defined as:*

$$\begin{aligned}
\mathrm{Upp}(f)(x) &:= \inf_y \sup_{x'} |x-y|^2 - |x'-y|^2 + f(x'), \\
\mathrm{Low}(f)(x) &:= \sup_y \inf_{x'} -|x-y|^2 + |x'-y|^2 + f(x').
\end{aligned}$$

It is immediate to verify that

$$\mathrm{Low}(f) \le f \le \mathrm{Upp}(f)$$

and that $\mathrm{Upp}(f)$ is the infimum on y, a of all the functions of the kind $x \mapsto |x-y|^2 + a$ which are bigger or equal than f on the whole $\mathbb{R}^d$; similarly, $\mathrm{Low}(f)$ is the supremum on y, a of the functions of the kind $-|x-y|^2 + a$ which are lesser or equal than f on the whole $\mathbb{R}^d$. For these kind of functions we deserve a name:

Definition 2.18 (Parabolas) *Let $y \in \mathbb{R}^d$, $a \in \mathbb{R}$. The* parabola $P_+(y,a)$ *is the function from $\mathbb{R}^d$ to $\mathbb{R}$ defined by $P_+(y,a)(x) := |x-y|^2 + a$. Similarly, the parabola $P_-(y,a) : \mathbb{R}^d \to \mathbb{R}$ is defined by $P_-(y,a)(x) = -|x-y|^2 + a$.*

In the language of c-transforms, for $c(x,y) = |x-y|^2$, the upper and lower envelope may be written as

$$\begin{aligned}
\mathrm{Upp}(f) &= f^{c_+ c_+}, \\
\mathrm{Low}(f) &= f^{c_- c_-}.
\end{aligned}$$

Lemma 2.19 *Let $P_+(y,a)$ and $P_-(\overline{y},\overline{a})$ be given. Then $P_+(y,a) \ge P_-(\overline{y},\overline{a})$ on the whole $\mathbb{R}^d$ if and only if*

$$\frac{|y-\overline{y}|^2}{2} \ge \overline{a} - a.$$

Similarly, it holds $P_+(y,a)(x) \le P_-(\overline{y},\overline{a})(x)$ for some $x \in \mathbb{R}^d$ if and only if

$$\frac{|y-\overline{y}|^2}{2} \le \overline{a} - a.$$

Thus in particular it holds $P_+(y,a)(x) \ge P_-(\overline{y},\overline{a})(x)$ on the whole $\mathbb{R}^d$ with equality for some point $x \in \mathbb{R}^d$ if and only if

$$\frac{|y-\overline{y}|^2}{2} = \overline{a} - a.$$

Proof It holds $|x-y|^2 + a \geq -|x-\overline{y}|^2 + \overline{a}$ for every $x \in \mathbb{R}^d$, if and only if

$$\inf_{x \in \mathbb{R}^d} |x-y|^2 + |x-\overline{y}|^2 \geq \overline{a} - a.$$

The infimum on the right hand side is attained at $x = \frac{y+\overline{y}}{2}$, and the value is $\frac{|y-\overline{y}|^2}{2}$.

The claims follow. □

The following is the crucial and purely geometrical result on which is based our argument.

Lemma 2.20 *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{\pm\infty\}$ be a function with compact support. Then it holds*

$$\mathrm{Upp}(\mathrm{Low}(\mathrm{Upp}(f))) = \mathrm{Low}(\mathrm{Upp}(f)). \tag{2.11}$$

Proof If $\mathrm{Upp}(f)(x_0) = +\infty$ for some $x_0 \in \mathbb{R}^d$, then $\mathrm{Upp}(f) \equiv +\infty$, and in this case the thesis is obvious. Also, if $\mathrm{Upp}(f)(x_0) = -\infty$, then $\mathrm{Low}(\mathrm{Upp}(f)) \equiv -\infty$, and this case the thesis is obvious as well. Similarly, if $\mathrm{Low}(\mathrm{Upp}(f))(x_0) = -\infty$, then $\mathrm{Low}(\mathrm{Upp}(f)) \equiv -\infty$ and there is nothing more to prove; while if $\mathrm{Low}(\mathrm{Upp}(f))(x_0) = +\infty$, then $\mathrm{Upp}(x_0) \geq \mathrm{Low}(\mathrm{Upp}(f))(x_0) = +\infty$ and, as before, we conclude $\mathrm{Upp}(f) = \mathrm{Low}(\mathrm{Upp}(f)) \equiv +\infty$. Thus we may assume that both $\mathrm{Upp}(f)$ and $\mathrm{Low}(\mathrm{Upp}(f))$ are real valued: in particular they are continuous.

For better clarity, we define:

$$\begin{aligned} g &:= \mathrm{Upp}(f), \\ h &:= \mathrm{Low}(g) = \mathrm{Low}(\mathrm{Upp}(f)), \end{aligned}$$

so that the problem consists in proving $\mathrm{Upp}(h) \leq h$.

Step 1: existence of 'optimal parabola'. Our first claim is that for every $x_0 \in \mathbb{R}^d$ there exists $y_0 \in \mathbb{R}^d$ and $a \in \mathbb{R}$ such that

$$\begin{aligned} g(x_0) &= |x_0 - y_0|^2 + a, \\ g(x) &\leq |x - y_0|^2 + a, \qquad \forall x \in \mathbb{R}^d. \end{aligned}$$

Indeed, pick a minimizing sequence y^n in the definition of $g(x_0)$: that is, a sequence satisfying

$$\begin{aligned} g(x_0) &= -\varepsilon^n + |x_0 - y^n|^2 + \sup_{x'} -|x' - y^n|^2 + f(x'), \\ g(x) &\leq |x - y^n|^2 + \sup_{x'} -|x' - y^n|^2 + f(x'), \qquad \forall x \in \mathbb{R}^d, \end{aligned}$$

for some sequence $\varepsilon^n \downarrow 0$. To prove the claim it is sufficient to show that $|y^n|$ is bounded. Argue by contradiction and assume that $\lim_n |y^n| = +\infty$; let $x^n := x_0 + \frac{y^n - x_0}{|y^n - x_0|}$. Observe that $|x_0 - x^n| = 1$ and

$$\begin{aligned} g(x^n) - g(x_0) &\leq |x^n - y^n|^2 - |x_0 - y^n|^2 + \varepsilon^n \\ &= |x^n - x_0|^2 + 2\langle x^n - x_0, x_0 - y^n \rangle + \varepsilon^n \\ &= 1 - 2|x_0 - y^n| + \varepsilon^n \to -\infty, \end{aligned}$$

which contradicts the continuity of g. Thus our claim is proved; similarly, it can be proved that for every $x_0 \in \mathbb{R}^d$ there exists $y_0 \in \mathbb{R}^d$ and $a_0 \in \mathbb{R}$ such that

$$\begin{aligned} h(x_0) &= -|x_0 - y_0|^2 + a_0, \\ h(x) &\geq -|x - y_0|^2 + a_0, \qquad \forall x \in \mathbb{R}^d. \end{aligned} \tag{2.12}$$

Step 2: the contact set. For any $x_0 \in \mathbb{R}^d$ and any y_0 satisfying (2.12) for some a_0, we define the contact set $C(x_0, y_0) \subset \mathbb{R}^d$ of those points where the graph of the parabola $P_-(y_0, a_0)$ touches the graph of g, i.e.:

$$C(x_0, y_0) := \Big\{ x \in \mathbb{R}^d \ : \ g(x) = P_-(y_0, a_0)(x) \Big\}.$$

Clearly $C(x_0, y_0)$ is a closed sets.

An important property of the contact set is the following: if $x \in C(x_0, y_0)$ then

$$\begin{aligned} g(x) &= P_+\Big(2x - y_0,\ -2|x - y_0|^2 + a_0\Big)(x), \\ g(x') &\leq P_+\Big(2x - y_0,\ -2|x - y_0|^2 + a_0\Big)(x'), \qquad \forall x' \in \mathbb{R}^d. \end{aligned} \tag{2.13}$$

Indeed, the parabola $P_+(2x - y_0, -2|x - y_0|^2 + a_0)$ is the *only* parabola of the kind $P_+(y, a)$ which stays above the parabola $P_-(y_0, a_0)$ on the whole $\mathbb{R}^d$ and for which it holds

$$P_+(y, a)(x) = P_-(y_0, x_0)(x).$$

Therefore since the value of $g(x)$ must be attained for some convex parabola, it is attained for this one, and the claim follows.

Step 3: interpolation of parabolas. Given two parabolas $P_+(y_0, a_0)$ and $P_+(y_1, a_1)$ we define their interpolation as the family of parabolas depending on a parameter $t \in [0, 1]$, defined as $P_+(y_t, a_t)$, where y_t and a_t are given by:

$$\begin{aligned} y_t &:= (1 - t)y_0 + t y_1, \\ a_t &:= (1 - t)a_0 + t a_1 + t(1 - t)\frac{|y_0 - y_1|^2}{2}. \end{aligned}$$

The interpolation of two parabolas has two important properties: the first is that if for some $y \in \mathbb{R}^d$ and $a \in \mathbb{R}$ it holds $P_-(y, a) \leq P_+(y_0, a_0)$ and $P_-(y, a) \leq P_+(y_1, a_1)$ on the whole $\mathbb{R}^d$, then it holds $P_-(y, a) \leq P_+(y_t, a_t)$ on the whole $\mathbb{R}^d$ for any $t \in [0, 1]$. The second property is that if for some $y \in \mathbb{R}^d$ and $a \in \mathbb{R}$ the parabola $P_-(y, a)$ is tangent to both $P_+(y_0, a_0)$ and $P_+(y_1, a_1)$, then it is tangent to all the $P_+(y_t, a_t)$, for any $t \in [0, 1]$.

Let us start proving the first property. From lemma 2.19 we know that the hypothesis $P_-(y, a) \leq P_+(y_0, a_0)$ and $P_-(y, a) \leq P_+(y_1, a_1)$, is equivalent to

$$\begin{aligned} \frac{|y - y_0|^2}{2} &\geq a - a_0, \\ \frac{|y - y_1|^2}{2} &\geq a - a_1. \end{aligned}$$

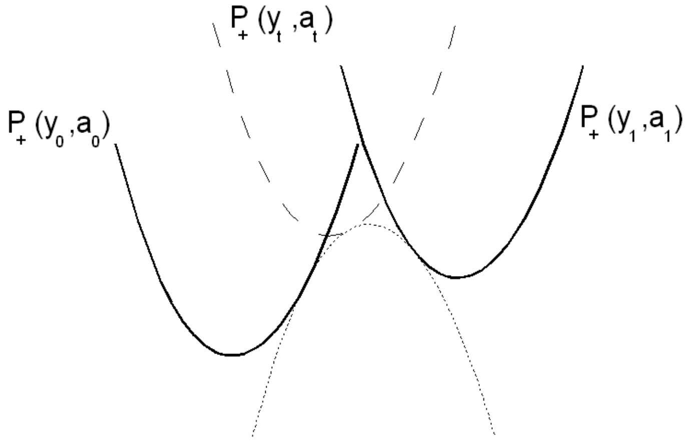


Figure 1: The envelope of the interpolated parabolas is a concave parabola

Then it holds

$$\frac{|y-y_t|^2}{2} = (1-t)\frac{|y-y_0|^2}{2} + t\frac{|y-y_1|^2}{2} - t(1-t)\frac{|y_0-y_1|^2}{2} \geq a - a_t, \quad \forall t \in [0,1],$$

and this equation is, again by lemma 2.19, equivalent to $P_-(y,a) \leq P_+(y_t, a_t)$ on the whole $\mathbb{R}^d$ for any $t \in [0,1]$.

The second property is proven in a similar way: it is enough to substitute the $\geq$ signs with equalities.

Step 4: convex hull of the contact set. Fix $x_0 \in \mathbb{R}^d$ and pick y_0 satisfying (2.12) for some $a_0 \in \mathbb{R}$. Let $\overline{C}(x_0, y_0)$ be the convex hull of $C(x_0, y_0)$. To each point $\overline{x} \in \overline{C}(x_0, y_0)$ we associate the parabola $P_+(\overline{y}, \overline{a})$ given by:

$$\begin{aligned} \overline{y} &:= 2\overline{x} - y_0, \\ \overline{a} &:= -2|\overline{x} - y_0|^2 + a_0. \end{aligned} \tag{2.14}$$

It is just a matter of algebraic manipulations to see that if $x^0, x^1 \in \overline{C}(x_0, y_0)$ are two given points, then the parabola associated to $(1-t)x^0 + tx^1$ is precisely the interpolation of the parabolas associated to x^0 and x^1 in the sense described above.

Here it comes the main idea of the proof: we claim that $x_0 \in \overline{C}(x_0, y_0)$. To prove this, we argue by contradiction. Assume that $x_0 \notin \overline{C}(x_0, y_0)$ and let $\overline{x} \in \overline{C}(x_0, y_0)$ be the point in $\overline{C}(x_0, y_0)$ which realizes the distance between x_0 and $\overline{C}(x_0, y_0)$ (which exists and is unique because $\overline{C}(x_0, y_0)$ is closed and convex). Let $P_+(\overline{y}, \overline{a})$ be the parabola associated to $\overline{x}$ via equations (2.14). Define

$$\begin{aligned} y_t &:= y_0 + t(2x_0 - \overline{y} - y_0), \\ a_t &:= \frac{|y_t - \overline{y}|^2}{2} - \frac{|y_0 - \overline{y}|^2}{2} + a_0. \end{aligned}$$

We claim that for $t > 0$ sufficiently small, the parabola $P_-(y_t, a_t)$ is less or equal than g on the whole $\mathbb{R}^d$. To prove this we argue by contradiction: thus we assume that there exists a sequence $t_n \downarrow 0$ and a sequence $(x_n) \subset \mathbb{R}^d$ such that

$$P_-(y_{t_n}, a_{t_n})(x_n) > g(x_n). \tag{2.15}$$

The fact that f has compact support and that $P_-(y_t, a_t)(x) \to -\infty$ as $|x| \to \infty$ gives that the sequence (x_n) is bounded. Thus, extracting if necessary a subsequence, we may assume that (x_n) converges to some $\tilde{x}$. Clearly $P_-(y_0, a_0)(\tilde{x}) = g(\tilde{x})$, that is $\tilde{x} \in C(x_0, y_0)$. Now observe that

$$\begin{aligned} P_-(y_t, a_t)(x) - P_-(y_0, a_0)(x) &= |x - y_t|^2 - |x - y_0|^2 + a_t - a_0 \\ &= t\,\langle 2x - \overline{y} - y_0, 2x_0 - \overline{y} - y_0\rangle + \frac{3t^2}{2}|2x_0 - \overline{y} - y_0|^2, \end{aligned}$$

thus recalling that $\overline{x} = \frac{y_0 + \overline{y}}{2}$ and that $\overline{x} \in \overline{C}(x_0, y_0)$ realizes the distance from x_0 to $\overline{C}(x_0, y_0)$, we get that $\langle \overline{x} - x_0, \overline{x} - \tilde{x}\rangle < 0$ and therefore

$$P_-(y_t, a_t)(\tilde{x}) - P_-(y_0, a_0)(\tilde{x}) = 4t\,\langle \overline{x} - x_0, \overline{x} - \tilde{x}\rangle + 6t^2|x_0 - \overline{x}|^2 < 0$$

for sufficiently small t. We deduce that for sufficiently small t the above strict inequality holds in a neighborhood of $\tilde{x}$, which gives $P_-(y_t, a_t)(x) < P_-(y_0, a_0)(x) \le g(x)$ for x near to $\tilde{x}$. This contradicts inequality (2.15) and gives the desired absurdum.

What we have proved is that if $x_0 \notin \overline{C}(x_0, y_0)$, then $P_-(y_t, a_t)$ stays below g on the whole $\mathbb{R}^d$ for sufficiently small t. To conclude that this is absurdum, and thus that $x_0 \in \overline{C}(x_0, y_0)$, observe that

$$\frac{d}{dt} P_-(y_t, a_t)(x_0)_{|t=0} = 4|\overline{x} - x_0|^2 > 0,$$

where the inequality is strict because we assumed $x_0 \notin \overline{C}(x_0, y_0) \ni \overline{x}$. This contradicts our choice of y_0: indeed since for small t we have $P_-(y_t, a_t) \le g$ on the whole $\mathbb{R}^d$ and $P_-(y_t, a_t)(x_0) > P_-(y_0, a_0)(x_0)$, the value of $h(x_0)$ is strictly greater than $P_-(y_0, a_0)(x_0)$, which is absurdum.

Thus $x_0 \in \overline{C}(x_0, y_0)$.

Step 5: conclusion. Now we summarize all what we proved up to now to conclude our proof. Choose $x_0 \in \mathbb{R}^d$: we want to prove that $\mathrm{Upp}(h)(x_0) \le h(x_0)$. Choose y_0, a_0 satisfying (2.12): since

$$P_-(y_0, a_0)(x_0) = P_+\Big(2x_0 - y_0, -2|x_0 - y_0|^2 + a_0\Big)(x_0),$$

to achieve the thesis it is sufficient to show that $P_+(2x_0 - y_0, 2h(x_0) - |x_0 - y_0|^2)$ is greater or equal than h on the whole $\mathbb{R}^d$. For any point $x \in \overline{C}(x_0, y_0)$ define $Y(x) \in \mathbb{R}^d$ and $A(x) \in \mathbb{R}$ as:

$$\begin{aligned} Y(x) &:= 2x - y_0, \\ A(x) &:= -2|x - y_0|^2 + a_0. \end{aligned}$$

Also, define the set $C^1(x_0, y_0)$ as the union of all the segments joining elements of $C(x_0, y_0)$, $C^2(x_0, y_0)$ as the union of all the segments joining elements of $C^1(x_0, y_0)$ and so on, so that, in particular, it holds $C^d(x_0, y_0) = \overline{C}(x_0, y_0)$.

Choose any $x^0 \in C(x_0, y_0)$ and use equations (2.13) to gain

$$\begin{aligned} P_+(Y(x^0), A(x^0))(x^0) &= g(x^0) = h(x^0) = P_-(y_0, a_0)(x^0), \\ P_+(Y(x^0), A(x^0))(x) &\geq g(x) \geq h(x) \geq P_-(y_0, a_0)(x), \qquad \forall x \in \mathbb{R}^d. \end{aligned} \tag{2.16}$$

Since this is valid for any $x^0 \in C(x_0, y_0)$, by the results on the interpolation of parabolas and the definitions of $Y(x)$, $A(x)$, we get that the equations (2.16) are valid also for any $x^1 \in C^1(x_0, y_0)$. Again by interpolation, equations (2.16) must hold for points in $C^2(x_0, y_0)$. Continue this way up to get that they are valid for any element of $\overline{C}(x_0, y_0)$. Thus in particular they are valid for x_0 and we obtain

$$\begin{aligned} P_+\Big(2x_0 - y_0, -2|x_0 - y_0|^2 - a_0\Big)(x_0) &= g(x_0) = h(x_0) = P_-(y_0, a_0)(x_0), \\ P_+\Big(2x_0 - y_0, -2|x_0 - y_0|^2 - a_0\Big)(x) &\geq g(x) \geq h(x) \geq P_-(y_0, a_0)(x), \qquad \forall x \in \mathbb{R}^d. \end{aligned} \tag{2.17}$$

□

Now we come back to our discussion about velocity vector field of a regular geodesic. Fix a geodesic $(\mu_t) \subset \mathcal{P}_c(M)$, let (v_t) be its velocity vector field and let $K \subset M$ be a compact set whose interior contains the support of all the μ_t's. Also, let ψ_t and φ_t as in proposition 1.26.

Finally, consider a partition of the unit of M subordinate to a cover by charts, and let $\{\theta_i\}_{i=1\dots,N}$ be those functions θ in the partition such that $\mathrm{supp}(\theta) \cap K \neq \emptyset$. Let $\Omega_i \subset M$, $i = 1, \dots, N$, be the open set containing $\mathrm{supp}(\theta_i)$ which is diffeomorphic to $\mathbb{R}^{\dim M}$ and let $\iota_i : \Omega_i \to \mathbb{R}^{\dim M}$ be the diffeomorphism.

Theorem 2.21 *With the notation just described, there exist two constants $A, B \in \mathbb{R}$ depending on $\{\theta_i\}, \{\iota_i\}, K$ such that for every $t \in (0,1)$, there exists a function ϕ_t which is both $-\big(\frac{A}{\min\{t,1-t\}} + B\big)$ concave and $-\big(\frac{A}{\min\{t,1-t\}} + B\big)$ convex satisfying*

$$\varphi_t \leq \phi_t \leq \psi_t. \tag{2.18}$$

Proof Fix $t \in (0,1)$ and observe that for any $y \in K$, the gradient of $x \mapsto \mathrm{d}^2(x, y)$ is bounded by $\mathrm{diam}(K)$ in K (that is: for any $x \in K$ and any v in the superdifferential of $\mathrm{d}^2(\cdot, y)$ at x it holds $|v| \leq \mathrm{diam}(K)$). Therefore any element of the superdifferential of $\psi_t(x)$ for $x \in K$ has norm bounded by $\frac{\mathrm{diam}(K)}{t}$. Similarly, any element of the subdifferential of $\varphi_t(x)$ for $x \in K$ has norm bounded by $\frac{\mathrm{diam}(K)}{1-t}$.

Up to adding to both ψ_t and φ_t the same constant, we may assume that $\psi_t(x_0) = \varphi_t(x_0) = 0$ for some $x_0 \in K$ (this doesn't affect our problem), thus from the above bound on the norm of elements in sub/super-differential it follows

$$\max\Big\{\sup_{x \in K} |\psi_t(x)|, \sup_{x \in K} |\varphi_t(x)|\Big\} \leq \frac{\big(\mathrm{diam}(K)\big)^2}{\min\{t, 1-t\}}.$$

Now fix $i \in \{1, \dots, N\}$ and consider the functions $\psi_t\theta_i$, $\varphi_t\theta_i$. The differential identities

$$\begin{aligned} \nabla(fg) &= (\nabla f)g + f(\nabla g), \\ \nabla^2(fg) &= (\nabla^2 f)g + f(\nabla^2 g) + \nabla f \otimes \nabla g + \nabla g \otimes \nabla f, \end{aligned}$$

valid for any couple of smooth functions $f, g : \mathbb{R}^{\dim M} \to \mathbb{R}$, gives that the norm of the gradient of $\psi_t\theta_i$ and $\varphi_t\theta_i$ is bounded by $\frac{A_1}{t} + B_1$ and that $\psi_t\theta_i$ is $-(\frac{A_1}{t} + B_1)$ concave on K and $\varphi_t\theta_i$ is $-(\frac{A_1}{1-t} + B_1)$ convex on K, where A_1, B_1 depend only on K and $\{\theta_i\}$.

Now consider the function $(\psi_t\theta_i) \circ \iota_i^{-1} : \mathbb{R}^{\dim M} \to \mathbb{R}$. The differential identity

$$\nabla^2(f \circ g) = (\nabla^2 f) \circ g \cdot (\nabla g)^2 + (\nabla f) \circ g \cdot \nabla^2 g,$$

valid for any couple of smooth functions $f : \mathbb{R}^{\dim M} \to \mathbb{R}$, $g : \mathbb{R}^{\dim M} \to \mathbb{R}^{\dim M}$ gives that the distributional Hessian of $(\psi_t\theta_i) \circ \iota_i^{-1}$ is bounded above by $(\frac{A_2}{t} + B_2)$, where A_2, B_2 depend only on A_1, B_1 and ι_i. Similarly, the distributional Hessian of $(\varphi_t\theta_i) \circ \iota_i^{-1} : \mathbb{R}^{\dim M} \to \mathbb{R}$ is bounded below by $-(\frac{A_2}{1-t} + B_2)$. In particular, the functions

$$\eta_{1,i,t} := (\psi_t\theta_i) \circ \iota_i^{-1} \left(\frac{A_2}{\min\{t, 1-t\}} + B_2 \right)^{-1}$$
$$\eta_{2,i,t} := (\varphi_t\theta_i) \circ \iota_i^{-1} \left(\frac{A_2}{\min\{t, 1-t\}} + B_2 \right)^{-1}$$

satisfy: $\eta_{1,i,t}$ is -1 concave and $\eta_{2,i,t}$ is -1 convex.

The fact that $\eta_{1,i,t}$ is -1 concave implies that $\mathrm{Upp}(\eta_{1,i,t}) = \eta_{1,i,t}$. Define $\eta_{i,t} := \mathrm{Low}(\eta_{1,i,t})$. By definition, $\eta_{i,t}$ is -1 convex; also, it holds $\eta_{i,t} \geq \eta_{2,i,t}$, since $\eta_{2,i,t}$ is a -1 convex function everywhere less or equal than $\eta_{1,i,t}$ and $\eta_{i,t}$ is the supremum of all such functions. Here we apply lemma 2.20: since

$$\begin{aligned} \mathrm{Upp}(\eta_{i,t}) &= \mathrm{Upp}(\mathrm{Low}(\eta_{1,i,t})) = \mathrm{Upp}(\mathrm{Low}(\mathrm{Upp}(\eta_{1,i,t}))) \\ &= \mathrm{Low}(\mathrm{Upp}(\eta_{1,i,t})) = \mathrm{Low}(\eta_{1,i,t}) = \eta_{i,t}, \end{aligned}$$

we obtain that $\eta_{i,t}$ is also -1 concave.

Therefore the function

$$\phi_{i,t} := \left(\frac{A_2}{\min\{t, 1-t\}} + B_2 \right)^{-1} \eta_{i,t}$$

is both $-(\frac{A_2}{\min\{t,1-t\}} + B_2)^{-1}$ convex and $-(\frac{A_2}{\min\{t,1-t\}} + B_2)^{-1}$ concave and satisfies

$$(\psi_t\theta_i) \circ \iota_i^{-1} \geq \phi_{i,t} \geq (\varphi_t\theta_i) \circ \iota_i^{-1},$$

on the whole $\mathbb{R}^{\dim M}$. Also, since both $(\psi_t\theta_i) \circ \iota_i^{-1}$ and $(\varphi_t\theta_i) \circ \iota_i^{-1}$ have compact support, the same is true for $\phi_{i,t}$. Observe that the 'convexity-concavity' property of $\phi_{i,t}$ gives that its distributional Hessian is absolutely continuous.

Now it is just a matter of 'coming back' to the manifold. The function $\phi_{i,t} \circ \iota_i$ satisfies

$$|\nabla(\phi_{i,t} \circ \iota_i)(x)| \leq \frac{A_3}{\min\{t, 1-t\}} + B_3, \qquad \forall x \in \Omega_i$$
$$\|\nabla^2(\phi_{i,t} \circ \iota_i)(x)\|_{\mathrm{op}} \leq \frac{A_3}{\min\{t, 1-t\}} + B_3, \qquad a.e.\ x \in \Omega_i,$$

where A_3, B_3 depend only on A_2, B_2 and $\{\iota_i\}$.

Finally, the function $\phi_t := \sum_i \phi_{i,t} \circ \iota_i \theta_i$ satisfies

$$\|\nabla^2 \phi_t(x)\|_{\text{op}} \leq \frac{A_4}{\min\{t, 1-t\}} + B_4, \qquad a.e.\ x \in K,$$

for some A_4, B_4 depending only on A_3, B_3 and $\{\theta_i\}$. By construction, it also satisfies equation (2.18); thus the proof is achieved. □

The fact that the vectors v_t may be approximated by gradients of smooth functions with bounded Lipschitz constant now follows:

Corollary 2.22 *With the same notation as above, for every $t \in (0,1)$ there exists a sequence $\phi_t^n \in C_c^\infty(M)$ such that*

$$\lim_{n \to \infty} \|v_t - \nabla \phi_t^n\|_{\mu_t} = 0,$$
$$\sup_{n \in \mathbb{N}} \mathrm{L}(\nabla \phi_t^n) < \infty.$$

Proof In the case $M = \mathbb{R}^d$, just find ϕ_t as in the statement of theorem 2.21 and define ϕ^n as approximation by convolution of ϕ: in this case $\nabla \phi_t^n$ converges to $\nabla \phi_t$ uniformly, and $\mathrm{L}(\nabla \phi_t^n) \leq \mathrm{L}(\nabla \phi_t)$ for any $n \in \mathbb{N}$.

For the general case, just pass to a finite set of coordinate system, approximate as above, and come back to the manifold. The uniform convergence of $\nabla \phi_t^n$ to $\nabla \phi_t$ is still true, while for the Lipschitz constants it holds

$$\mathrm{L}(\nabla \phi_t^n) \leq C \mathrm{L}(\nabla \phi_t), \qquad \forall n \in \mathbb{N},$$

where $C \in \mathbb{R}$ depends on the partition of the unit and on the charts chosen. □

Remark 2.23 (The case $M = \mathbb{R}^d$) In the Euclidean case, a direct application of lemma 2.20 shows that if $(\mu_t) \subset \mathscr{P}_2(\mathbb{R}^d)$ is a geodesic and (v_t) its velocity vector field, then there are functions ϕ_t which are both $-\frac{1}{\min\{t,1-t\}}$-convex and $-\frac{1}{\min\{t,1-t\}}$-concave, such that μ-a.e. it holds $v_t = \nabla \phi_t$. thus the result of theorem 2.21 is true for the Euclidean case regardless of the compactness assumption. ■

Remark 2.24 (The non compact case) The assumption of compactness of the supports of the μ_t's was used in two points: in order to apply remark 1.14 - to obtain the semiconvexity/semiconcavity of the interpolated potentials -, and in order to use only a finite number of charts. If the manifold M has non-negative sectional curvature, then remark 1.13 ensures that the potentials φ_t and ψ_t are $-\frac{1}{1-t}$-convex and $-\frac{1}{t}$-concave respectively. Therefore it seems there is some hope to generalize theorem 2.21 to such manifolds without imposing the compactness assumption.

We don't have a guess on the validity of the same theorem in the general case. ■

3 Absolutely continuous vector fields

3.1 Definition and first properties

In this section we introduce the notion of absolutely continuous vector fields defined along a regular curve and analyze the first properties of these fields. It is worth underlying that the vector fields we consider here may be not tangent, and their derivative in the sense of definition 3.6 below, is not the covariant derivative. A good analogy to think at is the following. Imagine a Riemannian manifold embedded on $\mathbb{R}^D$ and think at a smooth non tangent vector field defined along a smooth curve on the manifold: in this analogy, the Riemannian manifold is $\mathscr{P}_2(M)$, the smooth curve is a regular curve, and the smooth vector field is the absolutely continuous vector fields we analyze here. The derivative we are going to define corresponds, in this analogy, to the time derivative of the vector field along the curve, which we think as a vector in $\mathbb{R}^D$ which varies in time. The result of this derivation process does not produce, in general, a tangent vector field, not even if we assume that the starting vector field is tangent: we are going to show later on Chapter 5 how from these concepts it arises the natural Levi-Civita connection on $\mathscr{P}_2(M)$.

Definition 3.1 (Vector fields along a curve) *Let (μ_t) be a curve in $\mathscr{P}_2(M)$. A* vector field *along (μ_t) is a measurable map $t \mapsto u_t$ from $[0,1]$ to the set of measurable vector fields on M such that $u_t \in L^2_{\mu_t}$ for any t. We will denote it by (u_t) or by $t \mapsto u_t \in L^2_{\mu_t}$.*

There is some ambiguity with this terminology, as we say 'vector field' both for the map $t \mapsto u_t$ and for a single vector field u belonging to some L^2_μ. Hopefully, the context should always clarify in which sense we intend 'vector field'. Also, in the following we will often deal with vector fields defined only for a.e. t: the typical example being the velocity vector field (v_t) of a given absolutely continuous curve.

We will say that a vector field (u_t) is L^1 if $\int_0^1 \|u_t\|_{\mu_t} dt < \infty$, and that it is continuous if it is continuous w.r.t. the strong convergence of maps (definition 1.8).

For regularity higher than mere continuity, plain convergence is not sufficient to give a good definition, as it doesn't quantify the variation in time of the vector field.

The way to control this variation we propose is the following. Suppose our curve (μ_t) is regular. Then we can read the regularity of a vector field along it, by 'translating' the vector field (a priori defined in different L^2 spaces for different times) onto the same reference space using the translation maps τ_t^s.

Definition 3.2 (Regularity of vector fields) *Let (μ_t) be a regular curve and (u_t) a vector field along it. We say that (u_t) is absolutely continuous (or C^n, or $C^{n,1}$ or C^∞) if the map $t \mapsto \tau_t^s(u_t) \in L^2_{\mu_s}$ is absolutely continuous (or C^n, or $C^{n,1}$ or C^∞) for any $s \in [0,1]$.*

As said, the definition does *not* require the vector field to be tangent.

Remark 3.3 Given that τ_s^t is an isometry from $\mathrm{L}^2_{\mu_s}$ onto $\mathrm{L}^2_{\mu_t}$, to check the desired regularity of a vector field it is sufficient to check the regularity of the map $t \mapsto \tau_t^s(u_t)$ for *some* $s \in [0,1]$. ∎

Remark 3.4 It is easy to check that a vector field (u_t) is L^1 if and only if the map $t \mapsto \tau_t^s(u_t) \in L^2_{\mu_s}$ is L^1 for every $s \in [0,1]$, and that it is continuous if and only if the map $t \mapsto \tau_t^s(u_t) \in L^2_{\mu_s}$ is continuous. Thus the above definition works also for lower kind of regularity. ■

Remark 3.5 (Lebesgue points) For an L^1 vector field (u_t) they are well defined its *Lebesgue points*: we say that t_0 is a Lebesgue point of (u_t) if and only if t_0 is a Lebesgue point of $t \mapsto \tau_t^0(u_t) \in L^2_{\mu_0}$. ■

The definition of derivative of an absolutely continuous vector field, now comes quite natural:

Definition 3.6 (Total derivative of absolutely continuous vector fields) *Let (μ_t) be a regular curve and (u_t) an absolutely continuous vector field along it. We denote by $t \mapsto \frac{\boldsymbol{d}}{dt}u_t \in L^2_{\mu_t}$ its total derivative, defined by*

$$\frac{\boldsymbol{d}}{dt}u_t := \lim_{s\to t}\frac{\tau_s^t(u_s)-u_t}{s-t},$$

where the limit is intended to be in $\mathrm{L}^2_{\mu_t}$.

The reader should not be offended if we underline once again that this is *not* the definition of covariant derivative: nothing ensures that $\frac{\boldsymbol{d}}{dt}u_t$ is tangent, provided u_t is.

Proposition 3.7 (First properties of the derivation of vector fields) *Let (u_t) be an absolutely continuous vector field along the regular curve (μ_t). Then it holds:*

$$\frac{\boldsymbol{d}}{dt}u_t = \tau_s^t\left(\frac{d}{dt}\left(\tau_t^s(u_t)\right)\right) \qquad a.e.\ t,\ \forall s\in[0,1]. \tag{3.1}$$

The derivation of absolutely continuous vector fields is a linear operator and satisfies the Leibniz rule:

$$\frac{d}{dt}\left\langle u_t^1,u_t^2\right\rangle_{\mu_t} = \left\langle \frac{\boldsymbol{d}}{dt}u_t^1,u_t^2\right\rangle_{\mu_t} + \left\langle u_t^1,\frac{\boldsymbol{d}}{dt}u_t^2\right\rangle_{\mu_t}, \tag{3.2}$$

for any couple of absolutely continuous vector fields $(u_t^1),(u_t^2)$.

Proof The first equation follows directly from (2.5) and the group property of the maps τ_t^s. To prove the Leibniz rule observe that

$$\begin{aligned}\frac{d}{dt}\left\langle u_t^1,u_t^2\right\rangle_{\mu_t} &= \frac{d}{dt}\left\langle \tau_t^0(u_t^1),\tau_t^0(u_t^2)\right\rangle_{\mu_0}\\ &= \left\langle \frac{d}{dt}\tau_t^0(u_t^1),\tau_t^0(u_t^2)\right\rangle_{\mu_0} + \left\langle \tau_t^0(u_t^1),\frac{d}{dt}\tau_t^0(u_t^2)\right\rangle_{\mu_0}\\ &= \left\langle \frac{\boldsymbol{d}}{dt}u_t^1,u_t^2\right\rangle_{\mu_t} + \left\langle u_t^1,\frac{\boldsymbol{d}}{dt}u_t^2\right\rangle_{\mu_t}\end{aligned}$$

□

The first part of the statement above ensures that the total derivative of an absolutely continuous vector field is defined for a.e. t and is an L^1 vector field, as expected. It is interesting to observe that it is possible to 'integrate' vector fields as well.

Proposition 3.8 (Integral of vector fields) *Let (μ_t) be a regular curve, (u_t) be an L^1 vector field defined along it, and $U^0 \in L^2_{\mu_0}$. Then there exists a unique absolutely continuous vector field $t \mapsto U_t \in L^2_{\mu_t}$ satisfying*

$$\begin{cases} U_0 = U^0, \\ \dfrac{\boldsymbol{d}}{dt} U_t = u_t, \quad \text{a.e. } t. \end{cases}$$

Proof (U_t) solves the problem if and only if the vector field (V_t) defined by $V_t := \tau_t^0(U_t) \in L^2_{\mu_0}$ solves

$$\begin{cases} V_0 = U^0, \\ \dfrac{d}{dt} V_t = \tau_t^0(u_t). \end{cases}$$

The conclusion follows. □

Another straightforward consequence of equation (3.1) is that if a vector field is C^1, then its derivative is defined for every $t \in [0,1]$ (and not just for almost every) and is a continuous vector field. Similarly, if a vector field (u_t) is C^n, then its i-th derivative $\frac{\boldsymbol{d}^i}{dt^i} u_t$, $i \le n$, can be computed by

$$\frac{\boldsymbol{d}^i}{dt^i} u_t = \tau_s^t \left(\frac{d^i}{dt^i} \tau_t^s(u_t) \right),$$

so that (thankfully!) a C^n vector field has total derivatives up to order n and all of them are continuous.

A couple of interesting examples of vector fields along a regular curve (μ_t) are the following. The first one is given by $t \mapsto \xi \in L^2_{\mu_t}$ for a given $\xi \in \mathcal{V}(M)$ independent on time. Even if it may seem that the vector field does not depend on the time variable, this vector field should *not* be thought as a constant vector field along (μ_t). Observe that from

$$\|\tau_t^s(\xi) - \xi\|^2_{\mu_s} \le \mathrm{L}(\xi)(s-t) \int_t^s \|v_r\|^2_{\mu_r} dr, \qquad \forall t < s \in [0,1],$$

we have that (ξ) is absolutely continuous; now compute its total derivative from the definition to get:

$$\frac{\boldsymbol{d}}{dt}\xi = \lim_{h \to 0} \frac{\tau_{t+h}^t(\xi) - \xi(x)}{h} = \nabla\xi \cdot \frac{d}{ds}\mathbf{T}(t,s,\cdot)_{|s=t} = \nabla\xi \cdot v_t, \quad a.e.\ t \in [0,1].$$

The fact that the derivative is not zero explains why these vector fields are not 'constant'. More generally, any smooth time dependent vector field $(\xi_t) \in \mathcal{V}(M \times [0,1])$, is absolutely continuous and its derivative is given by

$$\frac{\boldsymbol{d}}{dt}\xi_t = \partial_t \xi_t + \nabla\xi \cdot v_t, \qquad a.e.\ t \in [0,1]. \tag{3.3}$$

Vector fields of the kind $(\xi_t) \in \mathcal{V}(M \times [0,1])$, despite their smoothness, in general are no more regular than absolutely continuous: their regularity is strictly linked to the regularity of the velocity vector field (v_t) of the curve, as shown in the following proposition.

Proposition 3.9 *Let (μ_t) be a regular curve and $(\xi_t) \in \mathcal{V}(M \times [0,1])$. Then $t \mapsto \xi_t \in L^2_{\mu_t}$ is C^n, $n \geq 1$, if and only if (v_t) is C^{n-1}.*

Proof Let $n = 1$. We already know that $t \mapsto \xi_t \in L^2_{\mu_t}$ is absolutely continuous; its derivative is given by $\partial_t \xi_t + \nabla \xi \cdot v_t$. Clearly this vector field is continuous if and only if (v_t) is. The rest follows by a recursion argument. For instance: if (v_t) is C^1, the derivative of $(\frac{\boldsymbol{d}}{dt}\xi_t)$ is given by

$$\frac{\boldsymbol{d}}{dt}\Big(\partial_t \xi_t + \nabla \xi_t \cdot v_t\Big) = \partial_{tt}\xi_t + \nabla(\partial_t \xi_t) \cdot v_t + \nabla^2 \xi(v_t, v_t) + \nabla \xi \cdot \frac{\boldsymbol{d}}{dt} v_t,$$

so that $\frac{\boldsymbol{d}^2}{dt^2}\xi$ is continuous; the same formula shows that if $\frac{\boldsymbol{d}^2}{dt^2}\xi$ is continuous, then (v_t) has to be C^1. $\square$

In the following, when speaking about time regularity of vector fields we will always refer to definition 3.2, which, in view of what just proved, is not equivalent to the usual concept of smoothness in time.

The second example of vector field we want to mention now is $\big(\tau_{t_0}^t(u)\big)$ for some $u \in L^2_{\mu_{t_0}}$. These are the vector fields that we can call 'constant': indeed it is an immediate consequence of the definition the fact that $\big(\tau_{t_0}^t(u)\big)$ is absolutely continuous and that its derivative is 0. Thus these vector field are C^∞. An important class of vector fields of this kind is the velocity vector field of a geodesic:

Proposition 3.10 *Let $(\mu_t) \subset \mathcal{P}_c(M)$ be the restriction to $[0,1]$ of a geodesic defined on some larger interval $[-\varepsilon, 1+\varepsilon]$, and (v_t) its velocity vector field. Then it holds*

$$v_s = \tau_t^s(v_t), \qquad \forall t, s \in [0,1]. \tag{3.4}$$

In particular, (v_t) is of class C^∞ and it holds

$$\frac{\boldsymbol{d}}{dt} v_t = 0, \quad \forall t \in (0,1)$$

.

Proof Recall that from corollary 2.16 we know that (μ_t) is regular, so that the statement makes sense. Now observe that a direct consequence of propositions 1.27 and 1.32 is that the flow maps are given by

$$\mathbf{T}(t, s, \cdot) = \exp\big((s-t)v_t\big).$$

Equation (3.4) is then just a restatement of equation (1.20). The rest follows. $\square$

Remark 3.11 We will see later that the equation $\frac{\boldsymbol{d}}{dt}v_t = 0$ implies that the covariant derivative of the vector field (v_t) is 0. It is curious the fact that for the velocity vector field of a geodesic, not only the covariant derivative is 0 - as expected by Riemannian analogy - but also the total one.

Also, observe that from a formal point of view, the equation $\frac{\boldsymbol{d}}{dt}v_t = 0$ may be written as

$$\partial_t v_t + \nabla v_t \cdot v_t = 0,$$

which is nothing but the spatial gradient of the Hamilton-Jacobi equation satisfied by the Kantorovich potentials (ϕ_t) along a geodesic:

$$\partial \phi_t + \frac{|\nabla \phi_t|^2}{2} = 0.$$

■

We conclude this introductory section by giving an equivalent characterization of absolute continuity. Let us start recalling the following result valid on Hilbert spaces.

Lemma 3.12 *Let $\mu \in \mathscr{P}_2(M)$ and $(w_t) \subset L^2_\mu$ be a given time dependent family of vectors. Then (w_t) is absolutely continuous as a curve from $[0,1]$ to L^2_μ if and only if for $\mu-$a.e. x, the curve $t \mapsto w_t(x) \in T_xM$ admits an absolutely continuous representative - not relabeled - and the vector field $(\overline{w}_t)$ defined by*

$$\overline{w}_t(x) := \frac{d}{dt} w_t(x), \qquad \mu \times \mathcal{L}^1 - a.e.\ (x,t),$$

where $\mathcal{L}^1$ is the 1-dimensional Lebesgue measure, is an L^1 family in L^2_μ, i.e. $\int_0^1 \|\overline{w}_t\|_\mu < \infty$ (in this case the family $(\overline{w}_t)$ is also the derivative of (w_t) in the L^2 sense).

From lemma 3.12 we have the following equivalent characterization of absolute continuity of a vector field defined along a regular curve.

Proposition 3.13 (Equivalent formulation of absolute continuity) *Let (μ_t) be a regular curve, $\mathbf{T}(t,s,\cdot)$ its flow maps and (u_t) a vector field defined along it. Then (u_t) is absolutely continuous if and only if:*

- *for μ_0-a.e. x the vector field $t \mapsto u_t(\mathbf{T}(0,t,x))$ defined along the curve $t \mapsto \mathbf{T}(0,t,x)$ admits an absolutely continuous representative (not relabeled),*

- *the vector field (u'_t) defined by $u'_t(\mathbf{T}(0,t,x)) = \frac{d}{dt}u_t(\mathbf{T}(0,t,x))$ (where $\frac{d}{dt}$ is the covariant derivative along $t \mapsto \mathbf{T}(0,t,x)$), belongs to $L^2_{\mu_t}$ for almost every t and the map $t \mapsto \|u'_t\|_{\mu_t}$ is integrable.*

In this case it holds $\frac{\boldsymbol{d}}{dt}u_t = u'_t$ in $L^2_{\mu_t}$ for a.e. t.

Proof It is enough to apply lemma 3.12 to the measure $\mu := \mu_0$ and the family $w_t := \tau_t^0(u_t)$.

□

A simple recursion argument shows that if the vector field (u_t) is $C^{n,1}$, i.e. the $n-$th derivative is absolutely continuous, then for $\mu-$a.e. x the vector field $t \mapsto u_t(x)$ is $C^{n,1}$ as well: for instance, if (u_t) is $C^{1,1}$, then its derivative (u'_t) is absolutely continuous, therefore $t \mapsto u'_t(x)$ is absolutely continuous and thus $t \mapsto u_t(x)$ is $C^{1,1}$.

Observe that in general, to know that (u_t) is C^1 it is *not* sufficient to derive that $t \mapsto u_t(\mathbf{T}(0,t,x))$ is C^1 as well for $\mu-$a.e. x, as the following example shows.

Example 3.14 Let $\mu := \mathcal{L}^1|_{[0,1]} \in \mathscr{P}(\mathbb{R})$ and $u_t \in L^2_\mu$ be defined as

$$u_t(x) := \begin{cases} 0 & \text{if } x \le t, \\ 1 & \text{if } x > t, \end{cases}$$

and $w_t := \int_0^t u_s\, ds$. Since (u_t) is continuous as a time depending function with values in L^2_μ, (w_t) is C^1. However, the map $t \mapsto w_t(x)$ is given by

$$w_t(x) := \begin{cases} t & \text{if } x \ge t, \\ x & \text{if } x < t, \end{cases}$$

and thus is not C^1 for any $x \in (0,1)$. ■

3.2 Approximation of absolutely continuous vector fields

Here we discuss the problem of approximating a given absolutely continuous vector field with regular ones.

The definition of convergence of vector fields is quite natural:

Definition 3.15 (Convergence of absolutely continuous vector fields) *Let (μ_t) be a regular curve and (u_t), (u^n_t), $n \in \mathbb{N}$, be given absolutely continuous vector fields along it. We say that the sequence $n \mapsto (u^n_t)$ converges to (u_t) if:*

i) $u^n_t \to u_t$ in $L^2_{\mu_t}$ as $n \to \infty$ for any $t \in [0,1]$,

ii) $\displaystyle\int_0^1 \left\| \frac{\boldsymbol{d}}{dt} u^n_t \right\|_{\mu_t} dt \to \int_0^1 \left\| \frac{\boldsymbol{d}}{dt} u_t \right\|_{\mu_t} dt$ as $n \to \infty$.

It should be clear that instead of asking pointwise convergence in (i), one could ask for uniform convergence, and that condition (ii) is equivalent to

$$\int_0^1 \left\| \frac{\boldsymbol{d}}{dt} u^n_t - \frac{\boldsymbol{d}}{dt} u_t \right\|_{\mu_t} dt \to 0$$

as $n \to \infty$, that is: 'the sequence of functions $(\|\frac{\boldsymbol{d}}{dt} u^n_t\|_{\mu_t})$ converges to $(\|\frac{\boldsymbol{d}}{dt} u_t\|_{\mu_t})$ in $L^1(0,1)$.

We start with the following simple result.

Proposition 3.16 (Approximation with vector fields regular in time) *Let (μ_t) be a regular curve and (u_t) an absolutely continuous vector field along it. Then there exists a sequence (u^n_t) of C^∞ vector fields along (μ_t) which converges to (u_t).*

Proof Consider the vector field $t \mapsto \overline{u}_t := \tau_t^0(u_t) \in L^2_{\mu_0}$ which is, by definition, absolutely continuous in $L^2_{\mu_0}$. In the Hilbert space $L^2_{\mu_0}$ we can find (e.g. by convolution) a sequence $(\overline{u}_t^n)$ of C^∞ vector fields uniformly converging to $(\overline{u}_t)$ and satisfying

$$\lim_{n\to\infty} \int_t^s \left\| \frac{d}{dt}\overline{u}_r^n \right\|_{\mu_0} dr = \int_t^s \left\| \frac{d}{dt}\overline{u}_r \right\|_{\mu_0} dr,$$

for any $t < s \in [0,1]$. Then it is sufficient to define $u_t^n := \tau_0^t(\overline{u}_t^n)$ for any $n \in \mathbb{N}$, $t \in [0,1]$. □

Thus the approximation of absolutely continuous vector fields with ones which are regular in time, is easy obtainable from the definitions. More delicate (but still possible) is the question of approximating a vector field with others 'regular in space'[1], i.e. with vector fields of the kind $(\xi_t) \in \mathcal{V}(M \times [0,1])$.

The following lemma reduces this problem to the one of approximating vector fields of the kind $(\tau_0^t(u_0))$, for $u_0 \in L^2_{\mu_0}$.

Lemma 3.17 *Let (μ_t) be a regular curve and (u_t) an absolutely continuous vector field along it. Fix $n \in \mathbb{N}$ and define the vector field (u_t^n) by*

$$u^n(t) := (1 - nt + i)\tau^t_{\frac{i}{n}}\left(u_{\frac{i}{n}}\right) + (nt - i)\tau^t_{\frac{i+1}{n}}\left(u_{\frac{i+1}{n}}\right), \qquad \forall t \in \left[\frac{i}{n}, \frac{i+1}{n}\right].$$

Then (u_t^n) is absolutely continuous for every $n \in \mathbb{N}$ and the sequence $n \mapsto (u_t^n)$ converges to (u_t).

Proof Defining $\overline{u}_t := \tau_t^0(u_t)$ and analogously $\overline{u}_t^n$, we see that $(\overline{u}_t^n)$ is nothing but a piecewise affine interpolation of $\overline{u}_t$. Whence the result follows. □

Thus, our problem of approximating general absolutely continuous vector fields is reduced to the problem of approximating vector fields of the kind $\tau_0^t(u_0)$ for $u_0 \in L^2_{\mu_0}$. A way to produce such an approximation is given in the following proposition.

Proposition 3.18 *Let (μ_t) be a regular curve and $u_0 \in L^2_{\mu_0}$. Then there exists a sequence $n \to (\xi_t^n)$ of vector fields of in $\mathcal{V}(M \times [0,1])$ which converges to $(\tau_0^t(u_0))$. Also, if $u_0 \in \mathcal{V}(M)$, such a sequence may be chosen to satisfy $\xi_0^n = u_0$ for every $n \in \mathbb{N}$.*

Proof It is clear that we can assume $u_0 = \xi \in \mathcal{V}(M)$. Use part (iii) of 2.6 to find a sequence of regular curves (μ_t^n) whose velocity vector fields (v_t^n) satisfy $(v_t^n) \in \mathcal{V}(M \times [0,1])$ and the transport couples (μ_t^n, v_t^n) converge to the transport couple (μ_t, v_t). Let $\mathbf{T}^n(t,s,\cdot)$ be the flow maps of (μ_t^n) and define $\xi_t^n := (\tau^n)_0^t(\xi)$. Observe that due to the smoothness of the vectors v_t^n, we have that $(\xi_t^n) \in \mathcal{V}(M \times [0,1])$, thus, in particular, (ξ_t^n) is an absolutely continuous vector field along (μ_t). We claim that the sequence (ξ_t^n) converges to $(\tau_t^0(\xi))$ as $n \to \infty$.

[1]The approximating vector fields we are looking for are C_c^∞ in time and space in the sense that they belong to $\mathcal{V}(M \times [0,1])$, still, as already said, we prefer to say that this is an approximation with vectors regular in space (thus not mentioning the regularity in time) since we prefer to use the terminology of regularity in time only in the sense of definition 3.2.

To prove this, observe that from the convergence of the transport couples and the uniform bound, in n, on the Lipschitz constants of the vector fields v_t^n it follows that ξ_t^n converges in $L^2_{\mu_t}$ to $\tau_t^0(\xi)$ as $n \to \infty$ for any $t \in [0,1]$ (we omit the details). Thus we only have to check the convergence of derivatives. We have:

$$\frac{\boldsymbol{d}}{dt}\xi_t^n = \partial_t \xi_t^n + \nabla \xi_t^n \cdot v_t = \nabla \xi \cdot \frac{d}{dt}\mathbf{T}^n(t,0,\cdot) + \nabla \xi \cdot \nabla \mathbf{T}^n(t,0,\cdot) \cdot v_t. \tag{3.5}$$

Now observe that derivating in time the identity

$$\mathbf{T}^n(t,0,\mathbf{T}^n(0,t,x)) = x, \qquad \forall x \in M,$$

we get

$$\begin{aligned} 0 &= \frac{d}{dt}\Big(\mathbf{T}^n\big(t,0,\mathbf{T}^n(0,t,\cdot)\big)\Big) \\ &= \left(\frac{d}{dt}\mathbf{T}^n(t,0,\cdot)\right) \circ \mathbf{T}^n(0,t,\cdot) + \big(\nabla \mathbf{T}^n(t,0,\cdot)\big) \circ \mathbf{T}^n(0,t,\cdot) \cdot v_t^n \circ \mathbf{T}^n(0,t,\cdot), \end{aligned}$$

which means

$$\frac{d}{dt}\mathbf{T}^n(t,0,\cdot) = -\nabla \mathbf{T}^n(t,0,\cdot) \cdot v_t^n.$$

Substituting in (3.5) we obtain

$$\begin{aligned} \frac{\boldsymbol{d}}{dt}\xi_t^n &= -\nabla \xi \cdot \nabla \mathbf{T}^n(t,0,\cdot) \cdot v_t^n + \nabla \xi \cdot \nabla \mathbf{T}^n(t,0,\cdot) \cdot v_t \\ &= \nabla \xi \cdot \nabla \mathbf{T}^n(t,0,\cdot) \cdot \big(v_t - v_t^n\big), \end{aligned}$$

and therefore

$$\left\| \frac{\boldsymbol{d}}{dt}\xi_t^n \right\| \leq \mathrm{L}(\xi)\mathrm{Lip}\big(\mathbf{T}^n(t,0,\cdot)\big)\|v_t - v_t^n\|_{\mu_t} \leq \mathrm{L}(\xi) e^{\int_0^1 \mathrm{L}(v_r^n)dr}\|v_t - v_t^n\|_{\mu_t}.$$

The conclusion follows from the $L^1(0,1)$ convergences of $(\mathrm{L}(v_t^n))$ to $(\mathrm{L}(v_t))$, of $(\|v_t - v_t^n\|_{\mu_t})$ to 0 and the dominated convergence theorem. □

Corollary 3.19 *Let (μ_t) be a regular curve and (u_t) a vector field of the kind*

$$u_t = (1-t)\tau_0^t(u_0) + t\tau_1^t(u_1),$$

for some $u_0 \in L^2_{\mu_0}$ and $u_1 \in L^2_{\mu_1}$. Then there exists a sequence of vector fields $(\xi_t^n) \in \mathcal{V}(M \times [0,1])$ which converges to (u_t). Also, if $u_0 \in \mathcal{V}(M)$, the sequence may be chosen to satisfy $\xi_0^n = u_0$ for every $n \in \mathbb{N}$.

Proof Straightforward. □

We are now ready to prove our main approximation result.

Proposition 3.20 (Approximation with vector fields regular in space) *Let (μ_t) be a regular curve and (u_t) an absolutely continuous vector field defined along it. Then there exists a sequence of vector fields $n \mapsto (\xi_t^n) \in \mathcal{V}(M \times [0,1])$ which converges to (u_t).*

Proof Choose $m \in \mathbb{N}$ and define the vector field (u_t^m) as

$$u_t^m := (1-t)\tau_{\frac{i}{m}}^t(u_{\frac{i}{m}}) + t\tau_{\frac{i+1}{m}}^t(u_{\frac{i+1}{m}}), \quad \forall t \in \left[\frac{i}{m}, \frac{i+1}{m}\right].$$

By lemma 3.17 we know that the sequence of vector fields (u_t^m) converges to (u_t) as $m \to \infty$.

Now fix $m \in \mathbb{N}$. We want to approximate the vector field (u_t^m) with vector fields $(\xi_t^{m,n}) \in \mathcal{V}(M \times [0,1])$ using corollary 3.19. We proceed by recursion. Focus on the interval $[0, 1/m]$: here we apply corollary 3.19 to find a sequence $n_1 \mapsto (\xi_t^{m,n_1})$ which converges to (u_t^m) as $n \to \infty$ in $[0, 1/m]$. Up to considering a subsequence, we may assume

$$\left\| \xi_{\frac{1}{m}}^{m,n_1} - u_{\frac{1}{m}}^m \right\|_{\mu_{\frac{1}{m}}} \leq \frac{1}{m^2}, \qquad \forall n_1 \in \mathbb{N}.$$

Now look at the interval $[1/m, 2/m]$ and consider the vector field

$$(2-mt)\,\tau_{\frac{1}{m}}^t\left(\xi_{\frac{1}{m}}^{m,n_1}\right) + (mt-1)\,\tau_{\frac{2}{m}}^t\left(u_{\frac{2}{m}}^m\right). \tag{3.6}$$

Using again corollary 3.19 we get the existence of a sequence of vector fields $n_2 \mapsto (\xi_t^{m,n_2})$ on $[1/m, 2/m]$ which converges to the above vector field as $n_2 \to \infty$. Also, we may choose such a sequence to satisfy

$$\xi_{\frac{1}{m}}^{m,n_2} = \xi_{\frac{1}{m}}^{m,n_1}, \tag{3.7a}$$

$$\left\| \xi_{\frac{2}{m}}^{m,n_2} - u_{\frac{2}{m}}^m \right\|_{\mu_{\frac{2}{m}}} \leq \frac{1}{m^2}, \tag{3.7b}$$

for every $n_2 \in \mathbb{N}$. Equation (3.7b) and the convergence of (ξ_t^{m,n_2}) to the vector field defined on (3.6) implies

$$\begin{aligned}
\int_{\frac{1}{m}}^{\frac{2}{m}} \left\| \frac{\boldsymbol{d}}{dt}\xi_t^{m,n_2} \right\|_{\mu_t} dt &\leq \left\| \xi_{\frac{1}{m}}^{m,n_1} - \tau_{\frac{2}{m}}^{\frac{1}{m}}(u_{\frac{2}{m}}^m) \right\|_{\mu_{\frac{1}{m}}} + f(n_2), \\
&\leq \left\| u_{\frac{1}{m}}^m - \tau_{\frac{2}{m}}^{\frac{1}{m}}(u_{\frac{2}{m}}^m) \right\|_{\mu_{\frac{1}{m}}} + \left\| \xi_{\frac{1}{m}}^{m,n_1} - u_{\frac{1}{m}}^m \right\|_{\mu_{\frac{1}{m}}} + f(n_2) \\
&\leq \int_{\frac{1}{m}}^{\frac{2}{m}} \left\| \frac{\boldsymbol{d}}{dt}u_t^m \right\|_{\mu_t} dt + \frac{1}{m^2} + f(n_2),
\end{aligned} \tag{3.8}$$

where $f(n_2) \to 0$ as $n_2 \to \infty$. Proceeding in this way and then 'gluing' the various vector fields (ξ_t^{m,n_k}), $k = 1, \ldots m$, that we found, we obtain vector fields $(\xi_t^{m,n})$ defined on $[0,1]$ which satisfy for every $m, n \in \mathbb{N}$:

- the map
$$M\times[0,1]\ni(x,t)\mapsto\xi_t^{m,n}(x)\in T_xM$$
is C^∞ in time and space in every interval of the kind $[\frac{i}{m},\frac{i+1}{m}]$, $i=0,\ldots,m-1$.

- For any $t\in[0,1]$ it holds
$$\xi_t^{m,n}\to u_t^m,\qquad \text{in } L^2_{\mu_t},$$
as $n\to\infty$.

- For some function $\mathbb{N}\ni n\mapsto f_m(n)\in\mathbb{R}$ such that $f_m(n)\to 0$ as $n\to\infty$ it holds
$$\int_0^1\left\|\frac{\boldsymbol{d}}{dt}\xi_t^{m,n}\right\|_{\mu_t}\leq\int_0^1\left\|\frac{\boldsymbol{d}}{dt}u_t^m\right\|+\frac{1}{m}+f_m(n).$$

A diagonalization argument shows that there exists a sequence $m\mapsto(\xi_t^{m,n_m})$ which converges to (u_t).

Now observe that from the first of the properties above and using any standard smoothening argument at the finite number of times $t=\frac{i}{m}$, $i=1,\ldots,m-1$, we can modify a bit the vector fields $(\xi_t^{m,n})$ to obtain a new family of vector fields in $\mathcal{V}(M\times[0,1])$ without affecting the convergence to (u_t). We omit the details. □

Proposition 3.21 (Approximation of Lipschitz vector fields) *Let (μ_t) be a regular curve and (u_t) an absolutely continuous vector field defined along it satisfying $\int_0^1\mathrm{L}(u_t)dt<\infty$. Then there exists a sequence of vector fields $n\mapsto(\xi_t^n)\in\mathcal{V}(M\times[0,1])$ which converges to (u_t) such that the sequence of functions $t\mapsto\mathrm{L}(\xi_t^n)$ converge to $t\mapsto\mathrm{L}(u_t)$ in $L^1(0,1)$ as $n\to\infty$.*

Proof The approximation argument is precisely the same that we just used. The only thing we have to add is the (easy to prove) bound:
$$\mathrm{L}\big(\tau_t^s(u)\big)\leq\mathrm{L}(u)\mathrm{Lip}(\mathbf{T}(t,s,\cdot)),\tag{3.9}$$
valid for any $u\in L^2_{\mu_t}$.

Step 1: variant of proposition 3.18. Assume that $\mathrm{L}(u_0)<\infty$. Recall that from part (iii) of theorem 2.6 we can approximate the regular curve (μ_t) with a sequence of transport couples (μ_t^n,v_t^n) such that the sequence of functions $t\mapsto\mathrm{L}(v_t^n)$ converge to $t\mapsto\mathrm{L}(v_t)$ as $n\to\infty$ in $L^1(0,1)$. Now use the same approximation argument of proposition 3.18 to derive that the vector field $(\tau_0^t(u_0))$ may be approximated by vector fields (ξ_t^n) satisfying $\xi_0^n=u_0$ and, by (3.9), the bound
$$\mathrm{L}(\xi_t^n)\leq\mathrm{L}(u_0)\mathrm{Lip}(\mathbf{T}^n(0,t,\cdot))\leq\mathrm{L}(u_0)e^{\int_0^t\mathrm{L}(v_s^n)ds}\leq\Big(e^{\int_0^1 g(s)ds}-1\Big)\mathrm{L}(u_0)\int_0^t g(s)ds,\tag{3.10}$$
where $g:[0,1]\to\mathbb{R}$ dominates the sequence $n\mapsto(\mathrm{L}(v_t^n))$. This inequality shows that the approximation result of 3.18 may be improved to have L^1-convergence of $(\mathrm{L}(\xi_t^n))$ to $(\mathrm{L}(\tau_0^t(u_0)))$.

Step 2: variant of lemma 3.17. Assume that $\sup_{t\in[0,1]}\mathrm{L}(u_t)<\infty$. Then proceed as in the proof of lemma 3.17 and use inequality (3.9) to derive that the sequence of vector fields (u_t^n)

not only converges to (u_t), but also satisfies 'the sequence of functions $n \mapsto (\mathrm{L}(u_t^n))$ converges to $(\mathrm{L}(u_t))$ in $L^1(0,1)$ as $n \to \infty$'.

If $t \mapsto \mathrm{L}(u_t)$ is not bounded, but just L^1, the same argument works, provided, instead of splitting the interval $[0,1]$ with the points $\frac{i}{n}$, we choose the splitting points t_i^n, $i = 0, \ldots, n-1$, such that $\mathrm{L}(u_{t_i^n})$ is close to the essential infimum of $\mathrm{L}(u_t)$ in the interval $[\frac{i}{n}, \frac{i+1}{n}]$ (then we should produce an approximation also in the intervals $[0, t_0^n]$ and $[t_n^n, 1]$, but this presents no difficulties, we omit the details).

Step 3: variant of proposition 3.20. The conclusion follows as in proposition 3.20: the approximation argument is precisely the same, we only need to keep track of the Lipschitz constants of the approximating vector fields. This is done using the first two steps of the proof.

□

4 Parallel transport

Rather than proceeding by first introducing the covariant derivative and then studying the problem of parallel transport, we introduce the parallel transport at first, and then we define the covariant derivative via the analogous of the formula

$$\nabla_{\gamma'(0)} u_t := \frac{d}{dt} T_t^0(u_t)_{|t=0},$$

valid for Riemannian manifolds, where T_t^0 is the parallel transport along the curve $t \mapsto \gamma(t)$ from the point $\gamma(t)$ to $\gamma(0)$.

We chose this approach, because in proving the existence of parallel transport along regular curves we will develop analytical tools that will be useful also in the following.

In order to clarify the idea of the proof, we start proving the existence of parallel transport in the well known case of a Riemannian manifold embedded in $\mathbb{R}^D$. The tools we are going to use for this case have a Wasserstein analogous, and imitating the proof for the Riemannian case we will be able to prove the existence of the parallel transport along regular curves.

It is worth underlying that the results of this chapter, especially those of the first 3 sections, are a direct generalization to the case of manifolds of the results in [1] and [10].

4.1 The case of an embedded Riemannian manifold

Throughout this subsection, $\tilde{M}$ will be a C^∞ manifold embedded in $\mathbb{R}^D$ with the induced Riemannian structure (the ˜ just stands to recall that $\tilde{M}$ has nothing to do with our 'universe' manifold M).

Let $\boldsymbol{\gamma}(t) : [0,1] \to \tilde{M}$ be a fixed C^∞ curve and let $v(t) = \dot{\boldsymbol{\gamma}}(t) \in T_{\boldsymbol{\gamma}(t)}\tilde{M}$, $t \in [0,1]$, be the derivative of $\boldsymbol{\gamma}(t)$. We will think to the tangent space $V_t := T_{\boldsymbol{\gamma}(t)}\tilde{M}$ 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 smooth vector field along the curve. In this setting the Levi-Civita derivative of $u(t)$ along $\gamma(t)$ is given by:

$$\nabla_{v(t)} u(t) := P_t \left(\frac{d}{dt} u(t) \right). \tag{4.1}$$

Thus, the vector field $u(t)$ is the parallel transport of the vector $u(0)$ along $\boldsymbol{\gamma}(t)$ if

$$P_t \left(\frac{du}{dt}(t) \right) = 0. \tag{4.2}$$

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:

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

Therefore the problem is to show the existence of a solution of (4.2) 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 $\mathscr{P}_2(M)$, since we don't have coordinates in such space. 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 4.1 (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

$$\begin{aligned} \sin\theta(V_0, V_1) &= \sup_{\substack{v_0 \in V_0 \\ |v_0|=1}} |v_0 - P_1(v_0)| = \|P_1^\perp|_{V_0}\| \\ &= \sup_{\substack{v_0 \in V_0,\ |v_0|=1 \\ v_1^\perp \in V_1^\perp,\ |v_1^\perp|=1}} \left\langle v_0, v_1^\perp \right\rangle = \sin\theta(V_1^\perp, V_0^\perp), \end{aligned} \tag{4.3}$$

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 $V_0 \subsetneq V_1$ 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:

$$|u - P_s(u)| \le C|u||s - t|, \qquad \forall t,\ s \in [0,1] \text{ and } u \in V_t, \tag{4.4a}$$

$$|P_s(u^\perp)| \le C|u^\perp||s - t|, \qquad \forall t,\ s \in [0,1] \text{ and } u^\perp \in V_t^\perp. \tag{4.4b}$$

The idea of the construction is based on the identity:

$$\nabla_{v(0)} P_t(u) = 0, \quad \forall u \in V_0. \tag{4.5}$$

That is: the vectors $P_t(u)$ are a first order approximation at $t = 0$ of the parallel transport. To prove (4.5) observe that taking (4.1) into account, (4.5) is equivalent to

$$|P_0(u - P_t(u))| = o(t), \qquad u \in V_0. \tag{4.6}$$

Equation (4.6) follows by applying twice the inequalities (4.4) (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 4.2 *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 (4.4). □

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

$$\begin{aligned} & \left|P_{s_N}(u) - P_{s_N}(P_{s_{N-1}}(\cdots(P_{s_2}(u))\cdots))\right| \\ \le\ & \left|P_{s_N}(u) - P_{s_N}(P_{s_{N-1}}(u))\right| + \left|P_{s_{N-1}}(u) - P_{s_{N-1}}(\cdots(P_{s_2}(u)))\right| \\ \le\ & \cdots \\ \le\ & C^2|u| \sum_{i=2}^{N-1} |s_1 - s_i||s_i - s_{i+1}| \le C^2 |u||s_1 - s_N|^2. \end{aligned} \tag{4.7}$$

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

Theorem 4.3 *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

$$|\mathcal{P}(u) - \mathcal{Q}(u)| \le |u|\varepsilon, \quad \forall \mathcal{Q} \ge \mathcal{P}. \tag{4.8}$$

In order to do so, it is sufficient to find $0 = t_0 < t_1 < \cdots < t_N = 1$ such that $\sum_i |t_{i+1} - t_i|^2 \le \varepsilon / C^2$, and repeatedly apply equation (4.7) to all partitions induced by $\mathcal{Q}$ in the intervals (t_i, t_{i+1}) (see the next section 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 4.4 *For any $t_1 \leq t_2 \leq t_3 \in [0, 1]$ it holds*

$$T_{t_2}^{t_3} \circ T_{t_1}^{t_2} = T_{t_1}^{t_3}. \tag{4.9}$$

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 Equation (4.9) follows by considering 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 (4.9) it is sufficient to check that the covariant derivative vanishes at 0. From (4.7) it follows that $|P_t(u) - u_t| \leq C^2 t^2$, thus the thesis follows from (4.5). □

4.2 Parallel transport along regular curves

Definition 4.5 (Parallel transport) *Let (μ_t) be a regular curve, (v_t) its velocity vector field and (u_t) an absolutely continuous tangent vector field. We say that (u_t) is a parallel transport if*

$$\mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt} u_t\right) = 0, \qquad a.e.\ t \in [0, 1].$$

It is clear from the linearity of the covariant derivative that linear combinations of parallel transports are parallel transports; furthermore, from the equality

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

valid for any parallel transport (u_t), we get that the parallel transports have constant norm. By linearity, it follows the uniqueness and the fact that the parallel transport preserves the scalar product.

The proof of existence is more delicate. From what we saw on smooth manifolds embedded in $\mathbb{R}^D$, it seems that the key step which allows to prove the existence of the parallel transport is the Lipschitz property of the angle between tangent spaces. As shown by theorem 2.13 and by inequality (2.4b), a similar property holds true even for in the space $\mathscr{P}_2(M)$ along a regular curve, in the sense that it holds:

$$\|\mathrm{P}_{\mu_t}^{\perp}(\tau_s^t(u_s))\|_{\mu_t} \leq \left(e^{\left|\int_t^s \mathrm{L}(v_r)dr\right|} - 1\right)\|u_s\|_{\mu_s}, \qquad \forall u_s \in \mathrm{Tan}_{\mu_s}(\mathscr{P}_2(M)).$$

Thus hopefully we can replicate the proof given for the case of Riemannian manifolds: we are going to show that this is actually possible. Observe that the convexity of $r \mapsto e^r - 1$ gives that the above inequality may be written as

$$\|\mathrm{P}_{\mu_t}^{\perp}(\tau_s^t(u_s))\|_{\mu_t} \leq C\left|\int_t^s \mathrm{L}(v_r)dr\right|\|u_s\|_{\mu_s}, \qquad \forall u_s \in \mathrm{Tan}_{\mu_s}(\mathscr{P}_2(M)), \tag{4.10}$$

with $C := e^{\int_0^1 \mathrm{L}(v_r)dr} - 1$.

To shorten a bit the notation, define

$$\mathscr{P}_t^s(u) := P_{\mu_s}\Big(\tau_t^s(u)\Big), \qquad \forall u \in \mathrm{L}^2_{\mu_t}$$

Observe that the maps $\mathscr{P}_t^s$ are non-expansive and that, by inequality (4.10) and the analogous of (4.3) we get:

$$\|\mathscr{P}_s^t(w)\|_{\mu_t} \leq C\left|\int_t^s \mathrm{L}(v_r)dr\right| \|w\|_{\mu_s}, \qquad t,\, s \in [0,1],\ w \in \mathrm{Tan}^\perp_{\mu_s}(\mathscr{P}_2(M)), \tag{4.11a}$$

$$\|\tau_t^s(u) - \mathscr{P}_t^s(u)\|_{\mu_s} \leq C\left|\int_t^s \mathrm{L}(v_r)dr\right| \|u\|_{\mu_t}, \qquad t,\, s \in [0,1],\ u \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M)). \tag{4.11b}$$

As for the case of manifolds, let $\mathfrak{P}$ be the direct set of all partitions of $[0,1]$, where, for $\mathcal{P}$, $\mathcal{Q} \in \mathfrak{P}$, $\mathcal{Q} \geq \mathcal{P}$ 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 $u \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ define $\mathcal{P}(u) \in \mathrm{Tan}_{\mu_1}(\mathscr{P}_2(M))$ as:

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

We will use several times the fact that:

$$\lim_{\mathcal{P}\in\mathfrak{P}} \sum_{i=0}^{n-1}\left(\int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr\right)^2 = 0, \tag{4.12}$$

which follows from

$$\begin{aligned}\sum_{i=0}^{n-1}\left(\int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr\right)^2 &\leq \max_{i=0,\dots,n-1}\left\{\int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr\right\}\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr\\ &= \max_{i=0,\dots,n-1}\left\{\int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr\right\}\int_0^1 \mathrm{L}(v_r)dr.\end{aligned}$$

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

The following lemma corresponds to Lemma 4.2:

Lemma 4.6 *Let $0 \leq s_1 \leq s_2 \leq s_3 \leq 1$ and let $u \in \mathrm{Tan}_{\mu_{s_1}}(\mathscr{P}_2(M))$. Then:*

$$\left\|\mathscr{P}^{s_3}_{s_1}(u) - \mathscr{P}^{s_3}_{s_2}(\mathscr{P}^{s_2}_{s_1}(u))\right\|_{\mu_{s_3}} \leq C^2\int_{s_1}^{s_2} \mathrm{L}(v_r)dr\int_{s_2}^{s_3} \mathrm{L}(v_r)dr\|u\|_{\mu_{s_1}}. \tag{4.13}$$

Proof Observe that, thanks to the group property of the translation maps, we have

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

and that $\tau^{s_2}_{s_1}(u) - \mathscr{P}^{s_2}_{s_1}(u) \in \mathrm{Tan}^\perp_{\mu_{s_2}}(\mathscr{P}_2(M))$. Therefore the thesis follows by a direct application of inequalities (4.11). □

Corollary 4.7 *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:*

$$\|\mathcal{P}(u) - \mathcal{Q}(u)\|_{\mu_s} \leq C^2 \|u\|_{\mu_t} \sum_{i=0}^{n-1} \left(\int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr \right)^2 \qquad \forall u \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M)). \tag{4.14}$$

Proof Without loss of generality we may assume $[t,s] = [0,1]$. Fix $i < n$ 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 $k(i) \geq 1$. Now, we claim that

$$\left\| \mathscr{P}_{s_{i,0}}^{s_{i,k(i)}}(u_{t_i}) - \mathscr{P}_{s_{i,k(i)-1}}^{s_{i,k(i)}} \left(\mathscr{P}_{s_{i,k(i)-2}}^{s_{i,k(i)-1}} \left(\cdots \left(\mathscr{P}_{s_{i,0}}^{s_{i,1}}(u_{t_i}) \right) \right) \right) \right\|_{\mu_{t_{i+1}}} \leq C^2 \|u_{t_i}\|_{\mu_{t_i}} \left(\int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr \right)^2 \tag{4.15}$$

for all $u_{t_i} \in \mathrm{Tan}_{\mu_{t_i}}(\mathscr{P}_2(M))$. Indeed, the left hand side of (4.15) can be estimated by

$$\begin{aligned}
& \left\| \mathscr{P}_{s_{i,0}}^{s_{i,k(i)}}(u_{t_i}) - \mathscr{P}_{s_{i,k(i)-1}}^{s_{i_{k(i)}}} \left(\mathscr{P}_{s_{i,0}}^{s_{i,k(i)-1}}(u_{t_i}) \right) \right\|_{\mu_{t_{i+1}}} \\
& \qquad + \left\| \mathscr{P}_{s_{i,k(i)-1}}^{s_{i,k(i)}} \left(\mathscr{P}_{s_{i,0}}^{s_{i,k(i)-1}}(u_{t_i}) \right) - \mathscr{P}_{s_{i,k(i)-1}}^{s_{i,k(i)}} \left(\mathscr{P}_{s_{i,k(i)-2}}^{s_{i,k(i)-1}} \left(\cdots \left(\mathscr{P}_{s_{i,0}}^{s_{i,1}}(u_{t_i}) \right) \right) \right) \right\|_{\mu_{t_{i+1}}} \\
\leq \;& C^2 \|u_{t_i}\|_{\mu_{t_i}} \int_{s_{i,0}}^{s_{i,k(i)-1}} \mathrm{L}(v_r)dr \int_{s_{i,k(i)-1}}^{s_{i,k(i)}} \mathrm{L}(v_r)dr \\
& \qquad + \left\| \mathscr{P}_{s_{i,0}}^{s_{i,k(i)-1}}(u_{t_i}) - \mathscr{P}_{s_{i,k(i)-2}}^{s_{i,k(i)-1}} \left(\mathscr{P}_{s_{i,k(i)-3}}^{s_{i,k(i)-2}} \left(\cdots \left(\mathscr{P}_{t_i}^{s_{i,0}}(u_{t_i}) \right) \right) \right) \right\|_{\mu_{t_{i+1}}} \\
\leq \;& \cdots \\
\leq \;& C^2 \|u_{t_i}\|_{\mu_{t_i}} \sum_{j=0}^{k(i)-1} \int_{s_{i,0}}^{s_{i,j}} \mathrm{L}(v_r)dr \int_{s_{i,j}}^{s_{i,j+1}} \mathrm{L}(v_r)dr \\
\leq \;& C^2 \|u_{t_i}\|_{\mu_{t_i}} \int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr \sum_{j=0}^{k(i)-1} \int_{s_{i,j}}^{s_{i,j+1}} \mathrm{L}(v_r)dr \\
= \;& C^2 \|u_{t_i}\|_{\mu_{t_i}} \left(\int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr \right)^2 .
\end{aligned}$$

Now let $\mathcal{P}' = [t_1, 1] \cap \mathcal{P}$, $\mathcal{Q}' = [t_1, 1] \cap \mathcal{Q}$, choose $u \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ and define $v, w \in \mathrm{Tan}_{\mu_{t_1}}(\mathscr{P}_2(M))$ by

$$\begin{aligned}
v &:= \mathscr{P}_{t_0}^{t_1}(u), \\
w &:= \mathscr{P}_{s_{0,k(0)-1}}^{s_{0,k(0)}} \left(\cdots \left(\mathscr{P}_{s_{0,0}}^{s_{0,1}}(u) \right) \right),
\end{aligned}$$

so that $\mathcal{P}(u) = \mathcal{P}'(v)$ and $\mathcal{Q}(u) = \mathcal{Q}'(w)$. Then, the inequality (4.15) with $i = 0$ reads

$$\|v - w\|_{\mu_{t_1}} \leq C^2 \|u\|_{t_0} \left(\int_{t_0}^{t_1} \mathrm{L}(v_r)dr \right)^2 ,$$

so that

$$
\begin{aligned}
\|\mathcal{P}(u) - \mathcal{Q}(u)\|_{\mu_1} &\leq \|\mathcal{P}'(v) - \mathcal{Q}'(v)\|_{\mu_1} + \|\mathcal{Q}'(v) - \mathcal{Q}'(w)\|_{\mu_1} \\
&\leq \|\mathcal{P}'(v) - \mathcal{Q}'(v)\|_{\mu_1} + \|v - w\|_{\mu_{t_1}} \\
&\leq \|\mathcal{P}'(v) - \mathcal{Q}'(v)\|_{\mu_1} + C^2 \|u\|_{t_0} \left(\int_{t_0}^{t_1} \mathrm{L}(v_r) dr \right)^2 .
\end{aligned}
$$

Since $\|v\|_{t_1} \leq \|u\|_{t_0}$ we can apply repeatedly (4.15) in the intervals (t_i, t_{i+1}) to obtain

$$
\|\mathcal{P}(u) - \mathcal{Q}(u)\|_{\mu_1} \leq C^2 \|u\|_{\mu_0} \sum_{i=0}^{n-1} \left(\int_{t_i}^{t_{i+1}} \mathrm{L}(v_r) dr \right)^2
$$

. □

The following result follows directly from the previous corollary and (4.12).

Theorem 4.8 (Existence of the limit of $\mathcal{P}(u_0)$) *Let (μ_t) be a regular curve and let $u_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$. Then $\lim_{\mathcal{P} \in \mathfrak{P}} \mathcal{P}(u_0)$ exists.*

Proof Straightforward. □

Define $\mathcal{T}_0^1(u_0)$ 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 $\mathcal{T}_t^s : \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M)) \to \mathrm{Tan}_{\mu_s}(\mathscr{P}_2(M))$ whenever $t \leq s$. Furthermore, by considering the curve $t \to \mu_{1-t}$, we can define the maps $\mathcal{T}_t^s$ even for $t > s$.

Proposition 4.9 (Group property) *Let (μ_t) be a regular curve and let $\mathcal{T}_t^s : \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M)) \to \mathrm{Tan}_{\mu_s}(\mathscr{P}_2(M))$ be defined as above. Then*

$$
\mathcal{T}_t^s \circ \mathcal{T}_r^t = \mathcal{T}_r^s, \qquad \forall r,\, s,\, t \in [0,1]. \tag{4.16}
$$

Proof Let us first assume $r \leq t \leq s$. 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 $\mathcal{T}_t^s = (\mathcal{T}_s^t)^{-1}$, or, without loss of generality, that $\mathcal{T}_0^1 = (\mathcal{T}_1^0)^{-1}$. The latter equation will follow if we show that

$$
\lim_{\mathcal{P} \in \mathfrak{P}} \|u - \mathcal{P}^{-1}(\mathcal{P}(u))\|_{\mu_0} = 0 \qquad \forall u \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M)), \tag{4.17}
$$

where $\mathcal{P}^{-1} : \mathrm{Tan}_{\mu_1}(\mathscr{P}_2(M)) \to \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ 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, is *not* the functional inverse of $u \to \mathcal{P}(u)$). Observe that for any $u \in \mathrm{Tan}_{\mu_{t_i}}(\mathscr{P}_2(M))$ the identities $u = \mathscr{P}_{t_{i+1}}^{t_i}(\tau_{t_i}^{t_{i+1}}(u))$

and $\mathscr{P}_{t_i}^{t_{i+1}}(u) - \tau_{t_i}^{t_{i+1}}(u) \in \mathrm{Tan}^{\perp}_{\mu_{t_{i+1}}}(\mathscr{P}_2(M))$, in conjunction with inequalities (4.11), yield

$$\begin{aligned}\left\|\mathscr{P}_{t_{i+1}}^{t_i}\left(\mathscr{P}_{t_i}^{t_{i+1}}(u)\right) - u\right\|_{\mu_{t_i}} &= \left\|\mathscr{P}_{t_{i+1}}^{t_i}\left(\mathscr{P}_{t_i}^{t_{i+1}}(u) - \tau_{t_i}^{t_{i+1}}(u)\right)\right\|_{\mu_{t_i}} \\ &\leq C\|\mathscr{P}_{t_i}^{t_{i+1}}(u) - \tau_{t_i}^{t_{i+1}}(u)\|_{\mu_{t_i}} \int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr \\ &\leq C^2\|u\|_{\mu_{t_i}} \left(\int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr\right)^2.\end{aligned}$$

For any $u \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ we obtain

$$\begin{aligned}&\left\|u - \mathscr{P}_{t_1}^0\left(\cdots\left(\mathscr{P}_1^{t_{n-1}}(\mathcal{P}(u))\right)\cdots\right)\right\|_{\mu_0} \\ &\leq \left\|u - \mathscr{P}_{t_1}^0\left(\mathscr{P}_0^{t_1}(u)\right)\right\|_{\mu_0} + \left\|\mathscr{P}_{t_1}^0\left(\mathscr{P}_0^{t_1}(u)\right) - \mathscr{P}_{t_1}^0\left(\mathscr{P}_{t_2}^{t_1}\left(\cdots\left(\mathscr{P}_1^{t_{n-1}}(\mathcal{P}(u))\right)\right)\right)\right\|_{\mu_0} \\ &\leq C^2\|u\|_{\mu_0}\left(\int_{t_0}^{t_1} \mathrm{L}(v_r)dr\right)^2 + \left\|v - \mathscr{P}_{t_2}^{t_1}\left(\cdots\left(\mathscr{P}_{t_{n-1}}^1(\mathcal{P}'(v))\right)\right)\right\|_{\mu_{t_1}},\end{aligned}$$

where $v = \mathscr{P}_0^{t_1}(u)$ and $\mathcal{P}' = \{t_1 < \cdots < t_n = 1\}$ (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

$$\left\|u - \mathscr{P}_{t_1}^0\left(\cdots\left(\mathscr{P}_1^{t_{n-1}}(\mathcal{P}(u))\right)\right)\right\|_{\mu_0} \leq C^2\|u\|_{\mu_0}\sum_{i=0}^{n-1}\left(\int_{t_i}^{t_{i+1}} \mathrm{L}(v_r)dr\right)^2$$

and this, taking (4.12) into account, leads to (4.17). □

Theorem 4.10 (The limit maps produce the parallel transport) *Let* (μ_t) *be a regular curve,* $u_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ *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 (4.14) and passing to the limit on $\mathcal{Q} \in \mathfrak{P}$ we get

$$\|\mathscr{P}_t^s(u) - \mathcal{T}_t^s(u)\|_{\mu_s} \leq C^2\|u\|_{\mu_t}\left(\int_t^s \mathrm{L}(v_r)dr\right)^2 \qquad \forall u \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M)). \tag{4.18}$$

Coupling this equation with inequality (4.11b) we get

$$\begin{aligned}\|\tau_t^s(u) - \mathcal{T}_t^s(u)\|_{\mu_s} &\leq \|\tau_t^s(u) - \mathscr{P}_t^s(u)\|_{\mu_s} + \|\mathscr{P}_t^s(u) - \mathcal{T}_t^s(u)\|_{\mu_s} \\ &\leq C\,(1+C)\,\|u\|_{\mu_t}\int_t^s \mathrm{L}(v_r)dr \qquad \forall u \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M)),\end{aligned} \tag{4.19}$$

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

Now, pick a Lebesgue point t of the function $t \mapsto \mathrm{L}(v_t)$ and observe that inequality (4.18) gives

$$\|\mathscr{P}_t^s(u_t) - u_s\|_{\mu_s} = o(s-t),$$

therefore, to conclude it is sufficient to prove that

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

Observe that $\mathscr{P}_t^s(u) - \tau_t^s(u) \in \mathrm{Tan}_{\mu_s}^{\perp}(\mathscr{P}_2(M))$, therefore from inequalities (4.11) we get

$$\begin{aligned} \|P_{\mu_t}(\tau_s^t(\mathscr{P}_t^s(u)) - u)\|_{\mu_t} &= \|\mathscr{P}_s^t(\mathscr{P}_t^s(u) - \tau_t^s(u))\|_{\mu_t} \\ &\leq C\|\mathscr{P}_t^s(u) - \tau_t^s(u)\|_{\mu_t} \int_t^s \mathrm{L}(v_r)dr \leq C^2\|u\|_{\mu_t} \left(\int_t^s \mathrm{L}(v_r)dr \right)^2. \end{aligned}$$

□

It is worth to underline that from this proof we get also an estimate on the modulus of absolute continuity of the parallel transport, which we state apart in the following proposition:

Proposition 4.11 (Modulus of absolute continuity of the parallel transport) *Let (μ_t) be a regular curve, (v_t) its velocity vector field and (u_t) a parallel transport along it. Then it holds*

$$\|\tau_t^s(u_t) - u_s\|_{\mu_s} \leq C(1+C)\|u_0\|_{\mu_0} \int_t^s \mathrm{L}(v_r)dr.$$

Proof The thesis is equivalent to equation (4.19). □

We will see later (theorem 5.15) that this bound is non optimal: actually it can be proved that it holds $\|\tau_t^s(u_t) - u_s\|_{\mu_s} \leq \|u_0\|_{\mu_0} \int_t^s \mathrm{L}(v_r)dr$.

Remark 4.12 (Parallel transport along a different flow) In this proof of existence we never used the fact that (v_t) was a tangent vector field, but just the fact that it was Lipschitz. This means that the same construction works as well if the curve (μ_t), rather then being regular, has a velocity field $(\tilde{v}_t)$ - not necessarily tangent - satisfying

$$\int_0^1 \mathrm{L}(\tilde{v}_t)dt < \infty.$$

The vector field which (u_t) which is produced by the limiting process has properties similar to that of the parallel transport. Indeed, defining the maps $\tilde{\mathbf{T}}(t,s,\cdot)$ as the flow maps of $(\tilde{v}_t)$, and $\tilde{\tau}_t^s$ as the associate translation maps (in the same spirit of definition 2.11), we have that the vector field (u_t) satisfies: $t \mapsto \tilde{\tau}_t^0(u_t) \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ is absolutely continuous. Also, this transport preserves the scalar product.

However, if the vector field $(\tilde{v}_t)$ is not a.e. tangent, this transport is not the natural parallel transport on $(\mathscr{P}_2(M), W_2)$, since the covariant derivative which it induces does not satisfy the torsion free identity (while the fact that this new transport preserves the norm is equivalent to the fact that the induced covariant derivative is compatible with the metric). See also remark 5.4. ■

4.3 Forward and Backward parallel transport

The purpose of this section is twofold: on one hand we show how the existence result for the parallel transport along regular curves admits a slight generalization to the case of *forward* parallel transport, on the other we show by an explicit example that the *backward* parallel transport may not exist. In particular, we will see that the parallel transport may fail to exist if the underlying curve is not regular.

The approach here is the following. Consider an absolutely continuous curve (μ_t) on $[0,1]$ such that the function $t \to \mathrm{L}(v_t)$ belongs to $L^1_{\mathrm{loc}}((0,1])$. We say that (u_t) is a parallel transport along (μ_t) if it is tangent for every $t \in [0,1]$, is a parallel transport on the interval $(0,1]$ (which makes sense, due to the locality of the definition of parallel transport) and u_t strongly converge to u_0 as $t \to 0$. Having this definition in mind, two questions come out naturally: the first one is whether there exists the parallel transport along (μ_t) of a vector in $\mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$, which we call forward parallel transport, the second one is whether there exists the parallel transport of a vector in $\mathrm{Tan}_{\mu_1}(\mathscr{P}_2(M))$, which we call backward parallel transport.

For the proof of existence of the forward parallel transport, we will need the following technical result which is of its own interest.

Lemma 4.13 *Let $t \to \mu_t$ be a regular curve and $\mathcal{T}_t^s$ the optimal transport maps along it. Then it holds*

$$\|\mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi\|_{\mu_s} \leq \mathrm{L}(\nabla\varphi)\int_t^s \|v_r\|_{\mu_r}dr, \quad \forall \varphi \in C_c^\infty(\mathbb{R}^d), \tag{4.20}$$

Proof Observe that $s \to \mathcal{T}_t^s(\nabla\varphi) - \nabla\varphi \in \mathrm{Tan}_{\mu_s}(\mathscr{P}_2(M))$ is an absolutely continuous vector field along (μ_t). The conclusion follows from the differential inequality:

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

□

Proposition 4.14 (Existence of forward parallel transport) *Let (μ_t) be an absolutely continuous curve such that the function $t \mapsto \mathrm{L}(v_t)$ belongs to $L^1_{\mathrm{loc}}((0,1])$ and let $u_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$. Then there exists a forward parallel transport of u_0 along (μ_t).*

Proof 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 $\mathrm{Tan}_{\mu_\varepsilon}(\mathscr{P}_2(M))$ and define the vectors $u_t^\varepsilon := \mathcal{T}_\varepsilon^t(\nabla\varphi)$ for any $t \in [\varepsilon, 1]$, so that we have $u_\varepsilon^\varepsilon = \nabla\varphi$. From

$$\|u_t^{\varepsilon'} - u_t^\varepsilon\|_{\mu_t} = \|u_\varepsilon^{\varepsilon'} - u_\varepsilon^\varepsilon\|_{\mu_\varepsilon} = \|\mathcal{T}_{\varepsilon'}^\varepsilon(\nabla\varphi) - \nabla\varphi\|_{\mu_\varepsilon} \leq \mathrm{L}(\nabla\varphi)\omega(\varepsilon) \qquad 0 < \varepsilon' \leq \varepsilon \leq t \leq 1,$$

with $\omega(\varepsilon) := \int_0^\varepsilon \|v_t\|_{\mu_t} dt$, we get that for any t, the family $\{u_t^\varepsilon\}$ converges in $\mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$, 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 given in proposition 4.11 we get its local absolute continuity, and we conclude by the stability of the solutions of

$$\frac{d}{dt} \langle u_t, \nabla\varphi \rangle_{\mu_t} = \left\langle u_t, \nabla^2\varphi \cdot v_t \right\rangle_{\mu_t}, \qquad a.e.\ t \in [0,1], \tag{4.21}$$

which characterizes the parallel transport among the set of all absolutely continuous and tangent vector fields.

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

$$\langle u_\varepsilon, \eta \rangle_{\mu_\varepsilon} = \langle u_\varepsilon - u_\varepsilon^\varepsilon, \eta \rangle_{\mu_\varepsilon} + \langle u_\varepsilon^\varepsilon, \eta \rangle_{\mu_\varepsilon} = R_\varepsilon + \langle \nabla\varphi, \eta \rangle_{\mu_\varepsilon} \qquad \forall \eta \in \mathcal{V}(M),$$

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

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 $t \to u_t^n$ 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 such that $\|u_t\|_{\mu_t} = \|u_0\|_{\mu_0}$. By the stability argument used above we get that $t \to u_t$ is a parallel transport on $(0, 1]$, so we need just to prove that u_t weakly converges to u_0 as $t \to 0$. To prove this, observe that since $[0, 1] \ni t \to u_t^n$ is a forward parallel transport, equation (4.21) gives

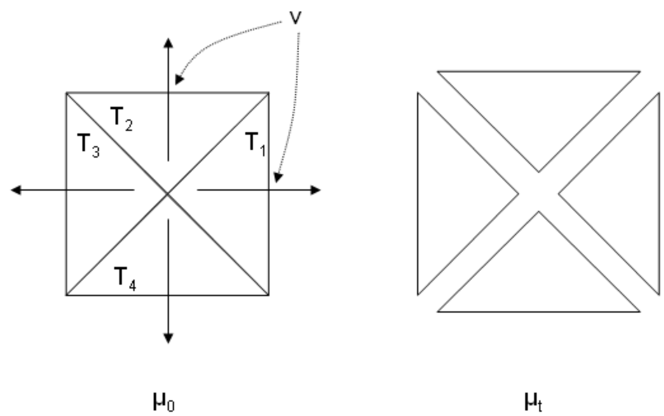
$$\left| \langle u_t^n, \nabla\psi \rangle_{\mu_t} - \langle u_0^n, \nabla\psi \rangle_{\mu_0} \right| = \left| \int_0^t \left\langle u_r^n, \nabla^2\psi \cdot v_r \right\rangle_{\mu_r} dr \right| \le \|u_t^n\|_{\mu_t} \mathrm{L}(\nabla\psi)\omega(t).$$

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

Now we turn to the counterexample to the existence of the backward parallel transport. As we will see, it is possible that a parallel transport u_t exists for positive t, that the vectors u_t converge strongly to some vector u_0 as $t \to 0$, while having that the vector u_0 is *not* a tangent vector: this shows that the problem of existence of the parallel transport is, in general, intrinsically prohibited by the geometry of $\mathscr{P}_2(\mathbb{R}^d)$. Observe that the curve considered is a geodesic.

Example 4.15 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 $\mu_0 := \chi_Q \mathcal{L}^2$ 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 $u = v^\perp$, the rotation by $\pi/2$ of v, in Q and $u = 0$ out of Q. Notice that u is orthogonal to $\mathrm{Tan}_\mu(\mathscr{P}_2(M))$, since it holds $\nabla \cdot (u\mu) = 0$.

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 $\mu_t = \chi_{Q_t} \mathcal{L}^2$.



It is immediate to check that the velocity vectors of μ_t are given by $v_t := v \circ (Id + tv)^{-1}$, so that $\mathrm{Lip}(v_t) = t^{-1}$ and μ_t is locally regular in $(0,1)$, and that the flow maps of μ_t in $(0,1]$ are given by

$$\mathbf{T}(t,s,\cdot) = (Id + sv) \circ (Id + tv)^{-1}, \quad \forall t,s \in (0,1].$$

Now, set $u_t := \tau_0^t(u) = u \circ \mathbf{T}(t,0,\cdot)$ and notice that u_t is tangent at μ_t, because u_t is constant in the connected components of the support of μ_t. Since $u_{t+h} \circ \mathbf{T}(t,t+h,\cdot) = u_t$, we obtain that u_t is a parallel transport in $(0,1]$. Furthermore, since u_t converges to u as $t \to 0$ and $u \notin \mathrm{Tan}_\mu(\mathscr{P}_2(M))$, there is no way to extend u_t to a continuous *tangent* vector field on the whole $[0,1]$. ∎

4.4 On the question of stability and the continuity of $\mu \mapsto \mathrm{P}_\mu$

A natural question which arises is the following: suppose that we have a sequence of regular curves converging (in a sense to be specified) to a limit regular curve, and assume also that a sequence of parallel transports along these curves is given and that these vector fields converge to a limit vector field. Is that true that the limit vector field is a parallel transport?

The general answer is no, as we show with an example.

Example 4.16 (The limit vector field may fail to be tangent) Choose any regular curve (μ_t) whose velocity vector field (v_t) is well defined and smooth in time and space on the whole M. Also, choose $u_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ such that $\tau_0^1(u_0) \notin \mathrm{Tan}_{\mu_1}(\mathscr{P}_2(M))$ (it is clear that such a couple curve-vector exists: that's why we had to take sequences of projections to be able to define the parallel transport). Given that $(v_t) \in \mathcal{V}(M \times [0,1])$, the flow maps $\mathbf{T}(t,s,\cdot)$ of (μ_t) are actually well defined and smooth on the whole M. Now choose any sequence $n \mapsto \mu_0^n$ of measures with finite support which converges to μ_0 w.r.t. W_2 and define

$$\mu_t^n := \mathbf{T}(0,t,\cdot)_\# \mu_0^n.$$

It is clear that the v_t's are the velocity vector fields of the curves (μ_t^n), so that these curves are regular. Also, it is immediate to check that $W_2(\mu_t^n, \mu_t) \to 0$ as $n \to \infty$ uniformly in t.

Now find a sequence of vectors $u_0^n \in L^2_{\mu_0^n}$ which converges to u_0 (this is always possible - see e.g. proposition 4.17 below for our case $u_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$), and define

$$u_t^n := \tau_0^{n,t}(u_0^n),$$

where $\tau_t^{n,s}$ are the translation maps along (μ_t^n). Since μ_t^n has finite support for any n, t, remark 1.33 gives that $u_t^n \in \mathrm{Tan}_{\mu_t^n}(\mathscr{P}_2(M))$, thus (u_t^n) is the parallel transport of u_0^n along (μ_t^n).

From the definition of (μ_t^n) and the fact that u_0^n strongly converges to u_0, it is not hard to see that that $u_t^n = \tau_0^{n,t}(u_0)$ strongly converges to $\tau_0^t(u_0)$ as $n \to \infty$.

Therefore we got a sequence of parallel transports which converges to a limit vector field which is *not* a parallel transport, as $\tau_0^1(u_0)$ is not a tangent vector by hypothesis. ■

On the other hand, we already noticed in the proof of 4.14 that the parallel transport may be characterized as those vector fields absolutely continuous, *tangent* and satisfying

$$\frac{d}{dt} \langle u_t, \nabla\varphi \rangle_{\mu_t} = \left\langle u_t, \nabla^2\varphi \cdot v_t \right\rangle_{\mu_t}, \qquad a.e.\ t \in [0,1].$$

Therefore it is clear from the stability of this condition that once a limit vector field exists, is absolutely continuous and tangent, it has to be a parallel transport.

Thus the big point behind the question of stability is the following: when is it true that limit of tangent vector fields is tangent?

Let us start with the following general result:

Proposition 4.17 ('Lower semicontinuity' of tangent spaces) *Let $n \mapsto \mu_n$ be a sequence in $\mathscr{P}_2(M)$ W_2-converging to some μ. Also, let $u_n \in \mathrm{Tan}^\perp_{\mu_n}(\mathscr{P}_2(M))$, $n \in \mathbb{N}$ and assume that $n \mapsto u_n$ weakly converges to some $u \in L^2_\mu$. Then $u \in \mathrm{Tan}^\perp_\mu(\mathscr{P}_2(M))$.*

Conversely, for any $u \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$, there exists a sequence $n \mapsto u_n \in \mathrm{Tan}_{\mu_n}(\mathscr{P}_2(M))$ strongly converging to it.

Proof To check that $u \in \mathrm{Tan}^\perp_\mu(\mathscr{P}_2(M))$ is equivalent to check that $\langle u, \nabla\varphi \rangle_\mu = 0$ for any $\varphi \in C_c^\infty(M)$. Since we know that $\langle u_n, \nabla\varphi \rangle_{\mu_n} = 0$ for every n, the convergence of $n \mapsto u_n$ to u guarantees that we can pass to the limit in this equality and get the claim.

For the converse, just pick $\boldsymbol{\gamma}_n \in \mathit{OptTan}(\mu_n, \mu)$ and define

$$L^2_{\mu_n} \ni (u \circ \boldsymbol{\gamma}_n)(x) := \int \tau_{\mathrm{v}}\big(u(\exp_x(\mathrm{v}))\big) d(\boldsymbol{\gamma}_n)_x(\mathrm{v}),$$

where $\{(\boldsymbol{\gamma}_n)_x\}$ is the disintegration of $\boldsymbol{\gamma}$ w.r.t. the projection π^M and $\tau_{\mathrm{v}} : T_{\exp_x(v)}M \to T_xM$ is the parallel transport map along the curve $t \mapsto \exp_x((1-t)\mathrm{v})$. Also, let $u_n := \mathrm{P}_{\mu_n}(u \circ \boldsymbol{\gamma}_n)$. It is easy to check that $n \mapsto u \circ \boldsymbol{\gamma}_n$ strongly converges to u as $n \to +\infty$, so we only have to prove that

$$\|u \circ \boldsymbol{\gamma}_n - \mathrm{P}_{\mu_n}(u \circ \boldsymbol{\gamma}_n)\|_{\mu_n} \to 0.$$

To this aim, assume for a moment that $u = \nabla\varphi$ for some $\varphi \in C_c^\infty(M)$ and observe that:

$$\begin{aligned}
\|\nabla\varphi \circ \boldsymbol{\gamma}_n - \mathrm{P}_{\mu_n}(\nabla\varphi \circ \boldsymbol{\gamma}_n)\|_{\mu_n} &\le \|\nabla\varphi \circ \boldsymbol{\gamma}_n - \nabla\varphi\|_{\mu_n} \\
&\le \sqrt{\int \left|\tau_{\mathrm{v}}\big(\nabla\varphi(\exp_x(\mathrm{v}))\big) - \nabla\varphi(x)\right|^2 d\boldsymbol{\gamma}_n(x,\mathrm{v})} \\
&\le \mathrm{L}(\nabla\varphi)\sqrt{\int \mathrm{d}^2(\exp_x(\mathrm{v}), x)^2 d\boldsymbol{\gamma}_n(x,\mathrm{v})} \\
&= \mathrm{L}(\nabla\varphi) W_2(\mu,\mu_n) \to 0
\end{aligned}$$

The general case follows by approximation using the fact the maps $u \mapsto u \circ \boldsymbol{\gamma}_n - \mathrm{P}_{\mu_n}(u \circ \boldsymbol{\gamma}_n)$ are equicontinuous (their norm is bounded by 1). □

Remark 4.18 Approximating a measure μ with measures with finite support μ_n, it is not hard to see that for any $v \in L^2_\mu$ we can find a sequence $v_n \in L^2_{\mu_n} = \mathrm{Tan}_{\mu_n}(\mathscr{P}_2(M))$ strongly converging to v. Also, by a standard smoothening argument, it is possible to check that the approximating measures may be chosen absolutely continuous w.r.t the volume measure and with smooth density (but still having the support made of several connected components).

Thus, in general, tangent vectors may converge to anything. ■

The fact that it may happen that a sequence of tangent vectors may converge to a non-tangent vector, may be seen as the discontinuity of the projection operator: indeed, choose any $\xi \in \mathcal{V}(M)$ and consider the map $\mu \mapsto \mathrm{P}_\mu(\xi)$. From the examples we saw it should be clear that this map is discontinuous if we endow the starting space with the natural topology induced by W_2 and the arriving space with the topology of weak convergence of maps. Before analyzing in which cases we have continuity, we observe that it holds the following sort of stability result:

Proposition 4.19 *Let $n \mapsto \mu_n$ be a sequence W_2 converging to some μ and let $u_n \in L^2_{\mu_n}$ be a sequence of vector fields weakly converging to some $u \in L^2_\mu$. Assume that $n \mapsto \mathrm{P}_{\mu_n}(u_n)$ weakly converges to some $\tilde{u}$. Then it holds*

$$\mathrm{P}_\mu(u) = \mathrm{P}_\mu(\tilde{\mu}).$$

Proof Just observe that

$$\langle u, \nabla\varphi\rangle_\mu = \lim_{n\to\infty} \langle u_n, \nabla\varphi\rangle_\mu = \lim_{n\to\infty} \langle \mathrm{P}_{\mu_n}(u_n), \nabla\varphi\rangle_\mu = \langle \tilde{u}, \nabla\varphi\rangle_\mu, \qquad \forall \varphi \in C_c^\infty(M).$$

□

Let us now focus on finding some conditions to ensure that limit of tangent vectors is a tangent vector. From the examples that we saw, we are allowed to guess that in order to avoid this discontinuity of the projection operator, we should be sure, at least, that the support of the limit measure is not split along the approximating sequence. While it is clear that we have continuity of the projection operator on the range of a regular curve (actually we are going to see that in this case it is absolutely continuous - see theorem 5.29), the discussion below covers more general situations.t along the approximating sequence. While it is clear that we have continuity of the projection operator on the range of a regular curve (actually we are going to see that in this case it is absolutely continuous - see theorem 5.29), the discussion below covers more general situations.

We are going to give two sufficient conditions, both of them are phrased for simplicity in the case $M = \mathbb{R}^d$. It's not hard to generalize them to the general case of Riemannian manifold, but a bit wordy, therefore we preferred to avoid it.

Proposition 4.20 (By convolution) *Let $n \mapsto \mu_n \in \mathscr{P}_2(\mathbb{R}^d)$ be a sequence W_2-converging to some $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. Assume that it holds $\mu_n = \mu * \rho_n$ for some sequence of functions $\rho_n : \mathbb{R}^d \to [0, +\infty)$ satisfying $\int \rho_n = 1$ which weakly converge to δ_0 in duality with continuous and bounded functions (smoothness is not an hypothesis). Let $u_n \in \mathrm{Tan}_{\mu_n}(\mathscr{P}_2(M))$ be a sequence weakly converging to some $u \in L^2_\mu$. Then $u \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$.*

Proof We will argue by duality. Let $w \in \mathrm{Tan}^\perp_\mu(\mathscr{P}_2(M))$ and define

$$w_n := \frac{(w\mu) * \rho_n}{\mu_n},$$

having identified the measure μ_n with its density w.r.t. the volume measure. It is known that $\|w_n\|_{\mu_n} \leq \|w\|_\mu$ (see lemma 8.1.10 in [2]) and thus it's not hard to check that $n \mapsto w_n$ strongly converges to w.

Observe that for any $\varphi \in C^\infty_c(\mathbb{R}^d)$ it holds

$$\langle w_n, \nabla\varphi \rangle_{\mu_n} = \int \langle (w\mu) * \rho_n, \nabla\varphi \rangle \, dx = \int \langle w, (\nabla\varphi) * \rho_n \rangle \, d\mu = \int \langle w, \nabla(\varphi * \rho_n) \rangle \, d\mu = 0,$$

which means $w_n \in \mathrm{Tan}^\perp_{\mu_n}(\mathscr{P}_2(\mathbb{R}^d))$. It is not hard to prove that the strong convergence of the w_n's plus the weak convergence of the u_n's implies

$$\lim_{n\to\infty} \langle w_n, u_n \rangle_{\mu_n} = \langle w, u \rangle_\mu .$$

From $w_n \in \mathrm{Tan}^\perp_{\mu_n}(\mathscr{P}_2(\mathbb{R}^d))$ and $u_n \in \mathrm{Tan}_{\mu_n}(\mathscr{P}_2(\mathbb{R}^d))$ we get $\langle w_n, u_n \rangle_{\mu_n} = 0$ for every $n \in \mathbb{N}$. Therefore

$$\langle w, u \rangle_\mu = 0, \qquad \forall w \in \mathrm{Tan}^\perp_{,}(\mathscr{P}_2(\mathbb{R}^d))$$

which is the thesis. □

The second condition was proven was proven in [1] and [10]. It can be regarded as a generalization of theorem 2.13.

Proposition 4.21 (By Lipschitz maps) *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 (not necessarily optimal). Then it holds*

$$\|v \circ T - \mathrm{P}_\mu(v \circ T)\|_\mu \leq \mathrm{Lip}(T - Id)\|v\|_\nu, \qquad \forall v \in \mathrm{Tan}_\nu(\mathscr{P}_2(\mathbb{R}^d)). \tag{4.22}$$

In particular, assume that $n \mapsto \mu_n$ *is a sequence in* $\mathscr{P}_2(\mathbb{R}^d)$ *which converges to some* μ *w.r.t.* W_2. *Assume also that there exists a sequence of maps (not necessarily optimal)* $T_n : \mathbb{R}^d \to \mathbb{R}^d$ *such that* $(T_n)_\#\mu = \mu_n$ *for every* $n \in \mathbb{N}$, $\|T_n - Id\|_\mu \to 0$ *and* $\mathrm{Lip}(T_n - Id) \to 0$ *as* $n \to \infty$. *Then for every sequence* $n \mapsto u_n \in \mathrm{Tan}_{\mu_n}(\mathscr{P}_2(M))$ *weakly converging to some* $u \in L^2_\mu$, *we have* $u \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$.

Proof The first claim is equivalent to

$$\|\nabla\varphi \circ T - P_\mu(\nabla\varphi \circ T)\|_\mu \le \|\nabla\varphi\|_\nu \mathrm{Lip}(T - Id), \qquad \forall \varphi \in C_c^\infty(\mathbb{R}^d). \tag{4.23}$$

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

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

Now turn to the general case. Find a sequence $(T_n - Id) \subset C_c^\infty(\mathbb{R}^d)$ 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 $\|T - T_n\|_\mu \to 0$, 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.

Let us turn to the second statement. Let $\boldsymbol{\gamma}_n := (Id, T_n)_\#\mu$ and fix $\xi \in \mathcal{V}(\mathbb{R}^d)$. Observe that

$$\langle u_n \circ T_n, \xi \rangle_\mu = \int \langle u_n(y), \xi(x) \rangle \, d\boldsymbol{\gamma}_n(x, y) = \int \langle u_n(y), \xi(y) \rangle \, d\boldsymbol{\gamma}_n(x, y) + R_n = \langle u_n, \xi \rangle_{\mu_n} + R_n,$$

where the reminder term R_n is bounded by

$$\int |u_n(y)||\xi(x) - \xi(y)| d\boldsymbol{\gamma}_n(x, y) \le \|u_n\|_{\mu_n} \mathrm{Lip}(\xi)\|T_n - Id\|_\mu,$$

and therefore,since $\sup_n \|u_n\|_{\mu_n} < \infty$ and $\|T_n - Id\|_\mu \to 0$, converges to 0. Thus we know that $u_n \circ T_n$ weakly converges to u in L^2_μ. From the first statement and the hypothesis $\mathrm{Lip}(T_n - Id) \to 0$ we know that $\|u_n \circ T_n - \mathrm{P}_\mu(u_n \circ T_n)\|_\mu \to 0$ as $n \to \infty$. Therefore $\mathrm{P}_\mu(u_n \circ T_n) \in \mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d))$ weakly converges to u as well. This implies $u \in \mathrm{Tan}_\mu(\mathscr{P}_2(\mathbb{R}^d))$. □

5 Covariant derivative

In this chapter we introduce the Levi-Civita connection on $(\mathscr{P}_2(M), W_2)$, analyze its properties and develop a calculus for vector fields along regular curves.

In the first section we show how the notion of parallel transport lead to a natural covariant differentiation, and we will see that this differentiation is the unique Levi-Civita connection on $(\mathscr{P}_2(M), W_2)$. In the second one, we introduce the important tensor $\mathcal{N}_\mu$ which, in a certain sense, describes precisely the infinitesimal variation of the tangent spaces. Then we pass to the 'practical' computation of covariant and total derivatives of some classes of absolutely continuous vector fields. Finally, we study the problem of the smoothness in time of the operators like $\mathrm{P}_{\mu_t}(\cdot)$ along a regular curve.

5.1 Levi-Civita connection

From the notion of parallel transport, the one of covariant derivative comes out naturally:

Definition 5.1 (Covariant derivative) *Let (μ_t) be a regular curve, $\mathcal{T}_t^s$ the parallel transport maps along it and (u_t) an absolutely continuous* tangent *vector field along (μ_t). The covariant derivative $\frac{\boldsymbol{D}}{dt}u_t$ along (μ_t) is defined as:*

$$\frac{\boldsymbol{D}}{dt}u_t := \lim_{h\to 0}\frac{\mathcal{T}_{t+h}^t(u_{t+h}) - u_t}{h},$$

where the limit is intended in $\mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$, or equivalently in $L^2_{\mu_t}$.

Recalling inequality (4.18) it is immediate to verify that the covariant derivative may be equivalently defined as

$$\frac{\boldsymbol{D}}{dt}u_t = \mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t\right),$$

so that it exists for a.e. $t \in [0,1]$. Also, the trivial inequality

$$\left\|\frac{\boldsymbol{D}}{dt}u_t\right\|_{\mu_t} \leq \left\|\frac{\boldsymbol{d}}{dt}u_t\right\|_{\mu_t},$$

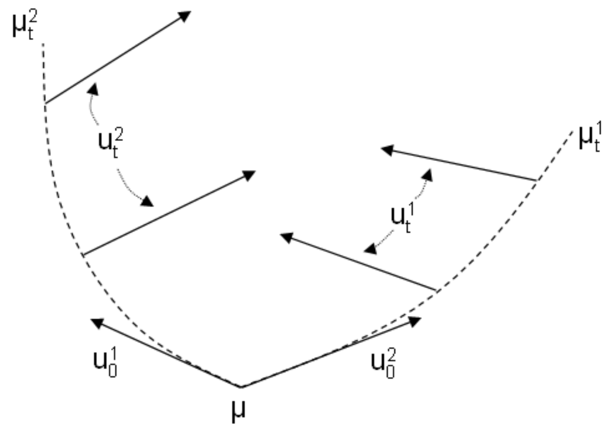
shows that the covariant derivative is an L^1 vector field.

In order to show that this derivative is the Levi-Civita connection on $(\mathscr{P}_2(M), W_2)$, we need to prove its *compatibility with the metric* and the *torsion free identity.*

The compatibility with the metric is a straightforward consequence of the Leibniz rule (3.2). Indeed for any couple of tangent and absolutely continuous vector fields u_t^1, u_t^2 it holds:

$$\begin{aligned}
\frac{d}{dt}\left\langle u_t^1, u_t^2\right\rangle_{\mu_t} &= \left\langle \frac{\boldsymbol{d}}{dt}u_t^1, u_t^2\right\rangle_{\mu_t} + \left\langle u_t^1, \frac{\boldsymbol{d}}{dt}u_t^2\right\rangle_{\mu_t}\\
&= \left\langle P_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t^1\right), u_t^2\right\rangle_{\mu_t} + \left\langle u_t^1, P_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t^2\right)\right\rangle_{\mu_t}\\
&= \left\langle \frac{\boldsymbol{D}}{dt}u_t^1, u_t^2\right\rangle_{\mu_t} + \left\langle u_t^1, \frac{\boldsymbol{D}}{dt}u_t^2\right\rangle_{\mu_t}.
\end{aligned}$$

To prove the torsion-free identity we consider two regular curves μ_t^1, μ_t^2 defined on some right neighborhood of 0 with the same starting point $\mu = \mu_0^1 = \mu_0^2$ whose velocity vector fields (v_t^1) and (v_t^2) are continuous. Along these curves, we consider two C^1 vector fields (u_t^1) and (u_t^2) ((u_t^1) lies along (μ_t^1) and (u_t^2) along (μ_t^2)), such that $u^2 = v_0^1$, $u^1 = v_0^2$ (the continuity assumption on the velocity vector fields is necessary to give a meaning to this initial condition). Observe that with these hypothesis it makes sense to consider the covariant derivative $\frac{\boldsymbol{D}}{dt} u_t^2$ of (u_t^2) along (μ_t^2) at $t = 0$: for this derivative we write $\boldsymbol{\nabla}_{u_0^1} u_t^2$. Similarly for (u_t^1).



Fix $\varphi \in C_c^\infty$ and consider the functional $\mu \to F_\varphi(\mu) := \int \varphi d\mu$. By the continuity equation, the derivative of F_φ along u_t^2 is equal to $\langle \nabla\varphi, u_t^2 \rangle_{\mu_t^2}$, therefore the Leibniz rule (3.2) gives:

$$u_0^1(u^2(F_\varphi))(\mu) = \frac{d}{dt} \langle \nabla\varphi, u_t^2 \rangle_{\mu_t^2} |_{t=0} = \langle \nabla^2\varphi \cdot u^1, u_0^2 \rangle_\mu + \langle \nabla\varphi, \boldsymbol{\nabla}_{u^1} u_t^2 \rangle_\mu .$$

Subtracting the analogous term $u^2(u^1(F_\varphi))(\mu)$ and using the symmetry of $\nabla^2\varphi$ given by equation (1.3) we get

$$[u^1, u^2](F_\varphi)(\mu) = \langle \nabla\varphi, \boldsymbol{\nabla}_{u^1} u_t^2 - \boldsymbol{\nabla}_{u^2} u_t^1 \rangle_\mu .$$

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

$$[u^1, u^2](\mu) = \boldsymbol{\nabla}_{u^1} u_t^2 - \boldsymbol{\nabla}_{u^2} u_t^1, \tag{5.1}$$

which is the torsion free identity, as desired.

The existence of parallel transport allows to prove the existence of the integral of an L^1 tangent vector field, in the same spirit of proposition 3.8:

Proposition 5.2 *Let (μ_t) be a regular curve, (u_t) an L^1 tangent vector field along it and $U^0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$. Then there exists a unique vector field (U_t) tangent and absolutely continuous along (μ_t) which satisfies*

$$\begin{cases} U_0 = U^0, \\ \dfrac{\boldsymbol{D}}{dt} U_t = u_t, \qquad a.e.\ t \in [0,1]. \end{cases}$$

Proof (U_t) solves the problem if and only if the vector field (V_t) defined by $V_t := \mathcal{T}_t^0(U_t) \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ solves

$$\begin{cases} V_0 = U^0, \\ \dfrac{d}{dt}V_t = \mathcal{T}_t^0(u_t), \qquad a.e.\ t \in [0,1], \end{cases}$$

The conclusion follows. □

Remark 5.3 (Equation in the smooth case) If the vector field (u_t) is given by the gradient of a smooth function $\varphi \in C_c^\infty(M \times [0,1])$, i.e. if it holds $u_t = \nabla_x \varphi_t$ for every t, the covariant derivative is given by:

$$\frac{\boldsymbol{D}}{dt}u_t = \mathrm{P}_{\mu_t}\left(\partial_t \nabla_x \varphi_t + \nabla_x^2 \varphi \cdot v_t\right) = \nabla_x \partial_t \varphi_t + \mathrm{P}_{\mu_t}\left(\nabla_x^2 \varphi \cdot v_t\right).$$

In particular, if φ does not depend on time we have:

$$\frac{\boldsymbol{D}}{dt}\nabla\varphi = \mathrm{P}_{\mu_t}\left(\nabla^2 \varphi \cdot v_t\right), \tag{5.2}$$

Remark 5.4 (On the uniqueness of the Levi-Civita connection) The question on whether the covariant derivative is the unique Levi-Civita connection or not, is subtler than it seems at a first glance. Indeed, we defined *a priori* which are the derivable vector fields by using the maps $\mathbf{T}(t,s,\cdot)$, and after that we defined their derivative using again the maps $\mathbf{T}(t,s,\cdot)$. This kind of argument may seem a bit 'circular'.

Actually, we just proved that the covariant derivative we defined is the Levi-Civita connection, so we may be satisfied with this result. Also, from the algebraic point of view, the existence of the flow maps and of the associated translation maps seems necessary to start the machine. Such an existence is ensured by the integrability of the Lipschitz constant of the velocity vector field, which in turn is also a sufficient condition for the geometric result 2.13 to work.

What is possibly unclear, is the role of the choice of *tangent* vectors in the definition of regular curves: a priori it may be non evident where this choice came into play. The short answer to this question is: in the proof of the torsion-free identity.

To make this point clear, let us assume that to every regular curve (μ_t) is associated not only the velocity vector field (v_t), but also another vector field $(\tilde{v}_t)$ (non tangent) for which the continuity equation is satisfied and for which it holds

$$\int_0^1 \|\tilde{v}_t\|_{\mu_t}^2 dt < \infty,$$
$$\int_0^1 \mathrm{L}(\tilde{v}_t) dt < \infty.$$

In this situation, we can naturally associate to this new vector field, the flow maps $\tilde{\mathbf{T}}(t,s,\cdot)$ and the translation maps $\tilde{\tau}_t^s$ by imitating the analogous definition given for the tangent vector field (v_t). Also, we can then define a vector field (u_t) along (μ_t) to be absolutely continuous with respect to the vectors $(\tilde{v}_t)$ if the map $t \mapsto \tilde{\tau}_t^0(u_t) \in L^2_{\mu_0}$ is absolutely continuous. This definition is possibly non coherent with 3.2, but still meaningful: it is naturally associated to the variant of the total derivative defined by

$$\frac{\tilde{\boldsymbol{d}}}{dt} u_t := \lim_{h \to 0} \frac{\tilde{\tau}_{t+h}^t(u_{t+h}) - u_t}{h}.$$

Now observe that in the proof of theorem 2.13 we didn't use the fact that the translation maps where associated to *tangent* vector fields, but only the integrability of the Lipschitz constant. Thus the analogous of such result holds also for the maps $\tilde{\tau}_t^s$. This means that all the results of Chapter 4 are valid in this new setting: in particular, there exists natural parallel transport maps $\tilde{\mathcal{T}}_t^s : \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M)) \to \mathrm{Tan}_{\mu_s}(\mathscr{P}_2(M))$ associated to the $\tilde{v}_t$'s along (μ_t). These maps are characterized by the fact that for any $u \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ the vector field $t \mapsto \tilde{\mathcal{T}}_0^t(u) \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$ is absolutely continuous w.r.t. the vectors $(\tilde{v}_t)$, and

$$\frac{\tilde{\boldsymbol{d}}}{dt} \tilde{\mathcal{T}}_0^t(u) = 0, \qquad a.e.\ t.$$

As for the standard parallel transport, one can prove that these maps preserve the scalar product.

Once one has the parallel transport, he can define the covariant differentiation w.r.t. the $\tilde{v}_t$'s by

$$\frac{\tilde{\boldsymbol{D}}}{dt} u_t := \mathrm{P}_{\mu_t}\left(\frac{\tilde{\boldsymbol{d}}}{dt} u_t\right).$$

Since the maps $\tilde{\mathcal{T}}_t^s$ preserve the scalar product, it is immediate to check that this covariant differentiation is compatible with the metric.

Now assume that the covariant derivative $\frac{\tilde{\boldsymbol{D}}}{dt}$ satisfies also the torsion free identity. Then, by doing the calculations indicated in the Koszul formula, we deduce that the covariant derivatives have to coincide.

From the equality of covariant derivatives, one would like to conclude that the vector fields (v_t) and $(\tilde{v}_t)$ are actually equal. We believe that this is true, but we don't have a proof. Let us just mention that the conclusion holds at least in the case $M = \mathbb{R}^d$. We report the argument, due to L.Ambrosio [1]. Consider a vector field of the kind $(\nabla\varphi)$: by what we just proved we have

$$\mathrm{P}_{\mu_t}\big(\nabla^2\varphi \cdot v_t\big) = \frac{\boldsymbol{D}}{dt}\nabla\varphi = \frac{\tilde{\boldsymbol{D}}}{dt}\nabla\varphi = \mathrm{P}_{\mu_t}\big(\nabla^2\varphi \cdot \tilde{v}_t\big),$$

so that the situation essentially is: we have a measure $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and two vector fields $v, \tilde{v} \in L^2_\mu$ such that

$$\mathrm{P}_\mu\big(\nabla^2\varphi \cdot (v - \tilde{v})\big) = 0, \qquad \forall \varphi \in C_c^\infty(\mathbb{R}^d),$$

and we want to prove that $v - \tilde{v} = 0$. To prove this, rewrite the above equality as

$$\big\langle \nabla^2\varphi \cdot (v - \tilde{v}), \nabla\psi \big\rangle_\mu = 0, \qquad \forall \varphi, \psi \in C_c^\infty(M), \tag{5.3}$$

observe that thanks to the bound $\int |x|^2 d\mu(x) < \infty$ we can take as test functions, functions whose gradient has linear growth, and choose $\varphi(x) := |\langle x, \mathrm{v}\rangle|^2$ for some $\mathrm{v} \in \mathbb{R}^d$. Then equation (5.3) becomes

$$\int \langle \nabla\psi, \mathrm{v}\rangle \langle v - \tilde{v}, \mathrm{v}\rangle \, d\mu = 0, \qquad \forall \psi \in C_c^\infty(\mathbb{R}^d),\ \mathrm{v} \in \mathbb{R}^d,$$

which tells that the symmetric part of the distributional gradient of $(v - \tilde{v})\mu$ is 0. By the Korn inequality, this means that such distribution has to be a constant. By integrability, this constant is 0. ∎

Remark 5.5 (Gradient of a vector field and related notions) Once the covariant derivative is defined, it is natural to define a notion of *gradient* of a smooth vector field defined on the whole $\mathscr{P}_2(M)$. The only problem is that, up to now, we have a smoothness definition only for vector fields defined along a curve, and not for vector fields defined everywhere. Although it would be possible to give such a definition, the analysis results quite heavy: thus we prefer to restrict the attention to some very specific kind of vector fields, namely those of the kind $\mu \mapsto \nabla\varphi$ for some fixed $\varphi \in C_c^\infty(M)$. In this case, equation (5.2) suggests to define $\boldsymbol{\nabla}(\nabla\varphi) : \mathrm{Tan}_\mu(\mathscr{P}_2(M)) \to \mathrm{Tan}_\mu(\mathscr{P}_2(M))$ as:

$$\boldsymbol{\nabla}(\nabla\varphi) \cdot v := \mathrm{P}_\mu(\nabla^2\varphi \cdot v).$$

The trivial inequality

$$\|\mathrm{P}_\mu(\nabla^2\varphi \cdot v)\|_\mu \le \|\nabla^2\varphi \cdot v\|_\mu \le \mathrm{L}(\nabla\varphi)\|v\|_\mu,$$

shows that the linear operator $\boldsymbol{\nabla}(\nabla\varphi)$ is bounded for every $\mu \in \mathscr{P}_2(M)$.

Having defined the gradient of a vector field, we can naturally define its divergence as the trace of the gradient. However, this object is in general not well defined, or attains the value $+\infty$: the problem is not that much related to the smoothness of the vector field, but rather on the fact that dimension of the 'manifold' is infinite. Indeed, if we try to evaluate the divergence $\boldsymbol{\nabla} \cdot \nabla\varphi$ of $\mu \mapsto \nabla\varphi$ we get

$$\boldsymbol{\nabla} \cdot \nabla\varphi := \sum_{i\in\mathbb{N}} \left\langle \nabla^2\varphi \cdot e_i, e_i \right\rangle_\mu,$$

where $\{e_i\}_{i\in\mathbb{N}}$ is an orthonormal base of $\mathrm{Tan}_\mu(\mathscr{P}_2(M))$ (recall that $\mathrm{Tan}_\mu(\mathscr{P}_2(M))$ has infinite dimension as soon as $\operatorname{supp}\mu$ is infinite). Thus, for instance, if $M = \mathbb{R}^D$, $\mu \in \mathscr{P}_2(M)$ has compact support and $\varphi(x) = |x|^2$ on the support of μ the above quantity is

$$\boldsymbol{\nabla} \cdot \nabla\varphi = \sum_{i\in\mathbb{N}} \langle \mathcal{I} \cdot e_i, e_i\rangle_\mu = \sum_{i\in\mathbb{N}} \langle e_i, e_i\rangle_\mu = \sum_{i\in\mathbb{N}} 1 = +\infty.$$

For different choices of φ, in general nothing ensures that the above sum makes sense at all.

For the same reason, while one can hope to define the Hessian of a functional on $\mathscr{P}_2(M)$ (which is perfectly possible, at least, for those functionals of the kind $\mu \mapsto \int \varphi d\mu$, whose gradient is $\mu \mapsto \nabla\varphi$), the Laplacian is in general not well defined.

The fact that the Laplacian is not well defined creates some limits in the analysis over the space $(\mathscr{P}_2(M), W_2)$. The first being: if there is no Laplacian, there is no heat equation and there is no Brownian Motion. Also, the lack of Laplacian implies the lack of volume measure, as we show now. Indeed, recall that in a Riemannian manifold $\tilde{M}$, the volume measure vol is uniquely identified, up to multiplicative constant, by the equation

$$\int_{\tilde{M}} \langle \nabla\varphi, \nabla\psi \rangle \, d\mathrm{vol} = -\int_{\tilde{M}} \psi \Delta\varphi d\mathrm{vol}, \qquad \forall \varphi, \psi \in C_c^\infty(\tilde{M}).$$

We can try to do the same on $(\mathscr{P}_2(M), W_2)$: assume for simplicity that M is compact (so that $\mathscr{P}_2(M)$ is compact as well and we can avoid boundary terms in the integration by parts) and consider the two functionals $\mu \mapsto F_\varphi(\mu) := \int \varphi d\mu$ and the analogous F_ψ, for $\varphi, \psi \in C^\infty(M)$. We are willing to write:

$$\int_{\mathscr{P}_2(M)} \langle \nabla\varphi, \nabla\psi \rangle_\mu \, d\mathbf{vol}(\mu) = -\int_{\mathscr{P}_2(M)} F_\psi(\mu) \mathbf{\Delta} F_\varphi(\mu) d\mathbf{vol}(\mu),$$

where **vol** is the hypothetic volume measure on $(\mathscr{P}_2(M), W_2)$. In this formula, we are able to compute all the terms apart **vol** and $\mathbf{\Delta} F_\varphi = \mathbf{\nabla} \cdot \nabla\varphi$. Thus if the volume measure exists, we would be able to use this equality to gather informations on, and potentially identify, the value of $\mathbf{\Delta} F_\varphi$. However, we already know that $\mathbf{\Delta} F_\varphi$ is not a well defined object. Therefore we must conclude that the same is true for **vol**.

I recently had a conversation with K.T.Sturm, who told me that from a more measure-theoretical approach he came to the same conclusion that both the Laplacian and the volume measure are not well defined (the measures built in [22] and [23] are *not* the volume measures on the space of probability measures).

The fact that two very different approaches (our differential one, and Sturm's measure-theoretical one) lead to the same conclusions, strongly suggests that we should stop searching the Laplacian, the volume measure and the Brownian Motion on $(\mathscr{P}_2(M), W_2)$, and simply admit that these objects cannot exist. ■

Our problem now is how to calculate covariant derivatives. Here we are going to explain heuristically our strategy for vector fields of the kind $(\mathrm{P}_{\mu_t}(u_t))$, where (u_t) is a given absolutely continuous vector fields: in the following two sections we are going to make rigorous the arguments used here, and to see how they can be used to compute other kind of derivatives.

Thus, let (u_t) be an absolutely continuous and not necessarily tangent vector field, and consider the tangent vector field $(\mathrm{P}_{\mu_t}(u_t))$. Using the tools developed in the previous chapter it won't be hard to prove that $(\mathrm{P}_{\mu_t}(u_t))$ is absolutely continuous once (u_t) is: for the moment we skip the proof of this fact, and focus on trying to identify who is its covariant derivative.

We start observing that to know the value of $\frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(u_t)$, it is sufficient to know the value of $\left\langle \frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(u_t), \nabla\varphi \right\rangle_{\mu_t}$ for any $\varphi \in C_c^\infty(M)$. Using the Leibniz rule for covariant derivatives (i.e. the compatibility with the metric) we know that it holds

$$\begin{aligned}\left\langle \frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(u_t), \nabla\varphi \right\rangle_{\mu_t} &= \frac{d}{dt}\left\langle \mathrm{P}_{\mu_t}(u_t), \nabla\varphi\right\rangle_{\mu_t} - \left\langle \mathrm{P}_{\mu_t}(u_t), \frac{\boldsymbol{D}}{dt}\nabla\varphi \right\rangle_{\mu_t} \\ &= \frac{d}{dt}\left\langle \mathrm{P}_{\mu_t}(u_t), \nabla\varphi\right\rangle_{\mu_t} - \left\langle \mathrm{P}_{\mu_t}(u_t), \mathrm{P}_{\mu_t}(\nabla^2\varphi\cdot v_t)\right\rangle_{\mu_t}.\end{aligned}$$

In order to compute the derivative of $\langle \mathrm{P}_{\mu_t}(u_t), \nabla\varphi\rangle_{\mu_t}$, observe that $\nabla\varphi \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$. Therefore it holds $\langle \mathrm{P}_{\mu_t}(u_t), \nabla\varphi\rangle_{\mu_t} = \langle u_t, \nabla\varphi\rangle_{\mu_t}$ for any t. This means that to calculate the derivative of this function, we can use the Leibniz rule for total derivatives (in this way we got rid of the problem of differentiating something where the operator P_{μ_t} appears) and obtain

$$\begin{aligned}\left\langle \frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(u_t), \nabla\varphi \right\rangle_{\mu_t} &= \frac{d}{dt}\left\langle u_t, \nabla\varphi\right\rangle_{\mu_t} - \left\langle \mathrm{P}_{\mu_t}(u_t), \mathrm{P}_{\mu_t}(\nabla^2\varphi\cdot v_t)\right\rangle_{\mu_t} \\ &= \left\langle \frac{\boldsymbol{d}}{dt}u_t, \nabla\varphi \right\rangle_{\mu_t} + \left\langle u_t, \frac{\boldsymbol{d}}{dt}\nabla\varphi \right\rangle_{\mu_t} - \left\langle \mathrm{P}_{\mu_t}(u_t), \mathrm{P}_{\mu_t}(\nabla^2\varphi\cdot v_t)\right\rangle_{\mu_t} \\ &= \left\langle \frac{\boldsymbol{d}}{dt}u_t, \nabla\varphi \right\rangle_{\mu_t} + \left\langle u_t, \nabla^2\varphi\cdot v_t\right\rangle_{\mu_t} - \left\langle \mathrm{P}_{\mu_t}(u_t), \mathrm{P}_{\mu_t}(\nabla^2\varphi\cdot v_t)\right\rangle_{\mu_t} \\ &= \left\langle \frac{\boldsymbol{d}}{dt}u_t, \nabla\varphi \right\rangle_{\mu_t} + \left\langle \mathrm{P}^{\perp}_{\mu_t}(u_t), \mathrm{P}^{\perp}_{\mu_t}(\nabla^2\varphi\cdot v_t)\right\rangle_{\mu_t}.\end{aligned}$$

This formula identifies the value of $\frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(u_t)$ tested against gradient of smooth functions. However it presents a problem: the appearance of the term $\nabla^2\varphi$. Indeed, if $\frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(u_t)$ is vector in $\mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$ (which it is, because $(\mathrm{P}_{\mu_t}(u_t)$ is absolutely continuous), it has to have a meaning its scalar product with any tangent vector, not just with gradients of smooth functions. From the expression above it is unclear the meaning of $\mathrm{P}^{\perp}_{\mu_t}(\nabla^2\varphi\cdot v_t)$ if we want to substitute $\nabla\varphi$ with a generic $w \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$.

Here it comes an important observation. Suppose that v_t is smooth. Then it holds the formula

$$\nabla(\langle\nabla\varphi, v_t\rangle) = \nabla^2\varphi\cdot v_t + (\nabla v_t)^{\mathrm{t}}\cdot\nabla\varphi,$$

and considering this equality as an equality between vectors in $\mathrm{L}^2_{\mu_t}$ and taking the projection onto the normal space, we get

$$0 = \mathrm{P}^{\perp}_{\mu_t}\big(\nabla^2\varphi\cdot v_t\big) + \mathrm{P}^{\perp}_{\mu_t}\big((\nabla v_t)^{\mathrm{t}}\cdot\nabla\varphi\big). \tag{5.4}$$

In this way we 'transferred' the spatial derivative from the factor $\nabla\varphi$ to v_t. Using this equality in the formula for the covariant derivative of $\mathrm{P}_{\mu_t}(u_t)$, we get

$$\left\langle \frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(u_t), \nabla\varphi \right\rangle_{\mu_t} = \left\langle \frac{\boldsymbol{d}}{dt}u_t, \nabla\varphi \right\rangle_{\mu_t} - \left\langle \mathrm{P}^{\perp}_{\mu_t}(u_t), \mathrm{P}^{\perp}_{\mu_t}\big((\nabla v_t)^{\mathrm{t}}\cdot\nabla\varphi\big)\right\rangle_{\mu_t},$$

which identifies $\frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(u_t)$ as

$$\frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(u_t) = \mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t - \nabla v_t\cdot\mathrm{P}^{\perp}_{\mu_t}(u_t)\right).$$

In the following section, we are going to show how to make this calculation rigorous when the v_t's are not smooth.

5.2 The tensor $\mathcal{N}_\mu$

Let us forget for a moment the problem of calculating covariant derivatives, and let us focus on understanding when the 'transfer of the spatial derivative' that we did in formula (5.4) is well defined.

Motivated by the previous heuristic calculation we introduce the tensor $\mathcal{N}_\mu$ as the bilinear operator on $[\mathcal{V}(M)]^2$ with values in $\mathrm{Tan}_\mu^\perp(\mathscr{P}_2(M))$ given by

$$\mathcal{N}_\mu(\xi,\eta) := \mathrm{P}_\mu^\perp\left((\nabla\xi)^t \cdot \eta\right), \tag{5.5}$$

(the letter '$\mathcal{N}$' comes from the fact that the tensor is defined in terms of a projection onto the Normal space). We want to understand if we can extend this operator to some larger class of vector fields. A trivial observation is that the smoothness of the second variable is not required to give a meaning to the expression above; thus certainly $\mathcal{N}_\mu$ can be defined on $\mathcal{V}(M) \times L^2_\mu$. Now considering the gradient of $\langle \xi(x), \eta(x)\rangle$ along a certain test vector field $\tilde{\eta} \in \mathcal{V}(M)$:

$$\nabla \langle \xi, \eta\rangle \cdot \tilde{\eta} = \langle \nabla\xi \cdot \tilde{\eta}, \eta\rangle + \langle \xi, \nabla\eta \cdot \tilde{\eta}\rangle,$$

we identify $\nabla \langle \xi, \eta\rangle$ as

$$\nabla \langle \xi, \eta\rangle = \nabla\xi^{\mathrm{t}} \cdot \eta + \nabla\eta^{\mathrm{t}} \cdot \xi.$$

Thus, as before, taking the projection onto the normal space at μ we get

$$\mathcal{N}_\mu(\xi,\eta) = -\mathcal{N}_\mu(\eta,\xi).$$

Therefore by (anti-)symmetry $\mathcal{N}_\mu$ is certainly well defined also on $L^2_\mu \times \mathcal{V}(M)$. In summary, we proved that there is a natural extension of $\mathcal{N}_\mu(\xi,\eta)$ at least to the set of couples of vector fields in L^2_μ for which at least one vector field is smooth. Actually, we are going to prove now that it is possible to give a meaning to $\mathcal{N}_\mu(\xi,\eta)$ also in the case in which one component is just Lipschitz and the other in L^2_μ. This is quite surprising if we think that if the support of μ is small it may happen that the gradient of a Lipschitz vector field is never defined in the support of μ.

Now we turn to the rigorous definitions.

Definition 5.6 (The Lipschitz-non-Lipschitz set) *Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. The set LnL_μ is the subset of $[L^2_\mu]^2$ of those couples of vector fields (u,w) such that at least one between u and w is Lipschitz.*

We are going to show that $\mathcal{N}_\mu$ can be extend by continuity to the whole LnL_μ: what we need to clarify is the topology of LnL_μ.

To this aim, we introduce the set $\mathcal{S}$ of those sequences $n \mapsto (u_n, w_n)$ in LnL_μ which are Cauchy sequences in $[\mathrm{L}^2_\mu]^2$ and satisfy

$$\sup_{n\in\mathbb{N}} \Big\{ \min\{\mathrm{L}(u_n), \mathrm{L}(w_n)\} \Big\} < +\infty.$$

It is easy to check that a sequence $n \mapsto (u_n, w_n)$ belongs to $\mathcal{S}$ if and only if it is a Cauchy sequence in $[\mathrm{L}^2_\mu]^2$ and for every subsequence, not relabeled, there exists a further extraction $k \mapsto (u_{n_k}, w_{n_k})$ such that

$$\sup_{k\in\mathbb{N}} \mathrm{L}(u_{n_k}) < +\infty \quad \text{or} \quad \sup_{k\in\mathbb{N}} \mathrm{L}(w_{n_k}) < +\infty,$$

thus converging sequences are basically Cauchy sequences in $[\mathrm{L}^2_\mu]^2$ for which at least one of the components is uniformly Lipschitz.

Our aim now is to show that there exists a topology on LNL_μ for which converging sequences are exactly those in $\mathcal{S}$. To do this define, for every $n \in \mathbb{N}$, the set

$$X^n := \Big\{(u, w) \in \mathrm{L}^2_\mu \ : \ \mathrm{L}(u) \le n \text{ or } \mathrm{L}(w) \le n\Big\},$$

and endow it with the topology τ^n induced by its inclusion in $[\mathrm{L}^2_\mu]^2$. Clearly $\mathrm{LNL}_\mu = \cup_n X^n$.

Definition 5.7 (The topology of LNL_μ) *We endow LNL_μ with the inductive limit τ of the topologies τ^n on X^n. This means that τ is the strongest topology that lets all of the embeddings $\iota^n : X^n \hookrightarrow \mathrm{LNL}_\mu$ be continuous.*

It holds the following result.

Proposition 5.8 *The space (LNL_μ, τ) is an Hausdorff space for which converging sequences are those belonging to $\mathcal{S}$. Also, a function $f : (\mathrm{LNL}_\mu, \tau) \to (Y, \tilde\tau)$ is continuous if and only if so are the functions $\iota^n \circ f = f|_{X^n} : (X^n, \tau^n) \to (Y, \tilde\tau)$.*

Proof Let σ be the topology on LNL_μ induced by its inclusion in $[\mathrm{L}^2_\mu]^2$. We claim that τ is finer than σ (and thus in particular is Hausdorff), and that $A \in \tau$ if and only if $A \cap X^n \in \tau^n$ for every $n \in \mathbb{N}$. The first claim follows from the fact that the embeddings $(X^n, \tau^n) \hookrightarrow (\mathrm{LNL}_\mu, \sigma)$ are continuous, and from the fact that τ is by definition the strongest topology for which these embeddings are continuous. Now take $A \in \tau$. Since $(\iota^n)^{-1}(A) = A \cap X^n$ and ι^n is continuous, we get $A \cap X^n \in \tau^n$ for every $n \in \mathbb{N}$. Conversely the set identities

$$\begin{aligned}\Big(\bigcup_{i\in I} A_i\Big) \cap X^n &= \bigcup_{i\in I} (A_i \cap X^n) \qquad &\forall n \in \mathbb{N},\\ (A_1 \cap A_2) \cap X^n &= (A_1 \cap X^n) \cap (A_2 \cap X^n) \qquad &\forall n \in \mathbb{N},\end{aligned}$$

imply that the set of A's for which $A \cap X^n \in \tau^n$ for every $n \in \mathbb{N}$ is a topology. Thus this topology is τ. From this, it immediately follows that a function on (LNL_μ, τ) is continuous if and only if so are its restrictions to the spaces (X^n, τ^n).

Now we claim that $n \mapsto (u_n, w_n)$ is a τ-converging sequence if and only if it is a $\sigma-$converging sequence and there exists $N \in \mathbb{N}$ such that $(u_n, w_n) \in X^N$ for every $n \in \mathbb{N}$. The 'if' part of the statement is obvious, so we turn to the 'only if'. Pick a $\tau-$converging sequence and observe that since τ is finer than σ, this sequence is $\sigma-$converging; therefore we need only to prove that the sequence is contained in some X^N. Arguing by contradiction, it is enough to show that the image of a sequence $n \mapsto (u_n, w_n) \subset \mathrm{LNL}_\mu$ such that $(u_n, w_n) \notin X^n$ is a τ-closed set. But this is obvious, as the intersection of such image with any X^n is a finite set. The conclusion follows.

□

It is worth underlying that in the space (LnL_μ, τ) the axiom of first countability does *not* hold. In particular, this space is not a metric space.

Theorem 5.9 (The tensor $\mathcal{N}_\mu$) *Let $\mu \in \mathscr{P}_2(M)$. For any $\xi, \eta \in \mathcal{V}(M)$ define*

$$\mathcal{N}_\mu(\xi, \eta) = \mathrm{P}_\mu^\perp\big((\nabla\xi)^{\mathrm{t}} \cdot \eta\big).$$

Then the tensor $\mathcal{N}_\mu$ extends in a unique way to a continuous bilinear antisymmetric tensor, still denoted by $\mathcal{N}_\mu$, to the whole LnL_μ, with values in $\mathrm{Tan}_\mu^\perp(\mathscr{P}_2(M))$. Moreover, it holds the estimate:

$$\|\mathcal{N}_\mu(u, w)\|_\mu \leq \min\Big\{\mathrm{L}(u)\|w\|_\mu, \mathrm{L}(w)\|u\|_\mu\Big\}. \tag{5.6}$$

Proof To prove existence and continuity of the extension of $\mathcal{N}_\mu$ to the whole LnL_μ, it is sufficient, by the characterization of continuous functions on (LnL_μ, τ), to prove existence and uniqueness of the continuous extension of $\mathcal{N}_\mu|_{\mathcal{V}(M)\cap X^N}$ to X^N, for every $N \in \mathbb{N}$. Since (X^N, τ^N) is a metric space, continuity is equivalent to sequential continuity; thus pick a sequence $(\xi_n, \eta_n) \in \mathcal{V}(M) \cap X^N$ converging in the $[L^2_\mu]^2$ topology to some (u, w) and observe that up to taking a subsequence and exchanging, if necessary, the variables, we may assume that $\mathrm{L}(\xi_n) \leq N$ for every $n \in \mathbb{N}$. Choose $\eta \in \mathcal{V}(M)$ and observe that by the bilinearity of $\mathcal{N}_\mu$ on $\mathcal{V}(M)^2$ it holds

$$\begin{aligned}\|\mathcal{N}_\mu(\xi_n, \eta_n) - \mathcal{N}_\mu(\xi_m, \eta_m)\|_\mu &\leq \|\mathcal{N}_\mu(\xi_n, \eta_n - \eta)\|_\mu + \|\mathcal{N}_\mu(\xi_n - \xi_m, \eta)\|_\mu + \|\mathcal{N}_\mu(\xi_m, \eta_m - \eta)\|_\mu \\ &\leq \mathrm{L}(\xi_n)\|\eta_n - \eta\|_\mu + \|\xi_n - \xi_m\|_\mu \mathrm{L}(\eta) + \mathrm{L}(\xi_m)\|\eta_m - \eta\|_\mu,\end{aligned} \tag{5.7}$$

so that we get

$$\varlimsup_{n\to\infty} \varlimsup_{m\to\infty} \|\mathcal{N}_\mu(\xi_n, \eta_n) - \mathcal{N}_\mu(\xi_m, \eta_m)\|_\mu \leq \varlimsup_{n} \mathrm{L}(\xi_n)\|w - \eta\|_\mu \leq N\|w - \eta\|_\mu.$$

Letting η tend to w in the L^2_μ norm, and observing that this argument does not depend on the subsequence chosen, we obtain that $n \mapsto \mathcal{N}_\mu(\xi_n, \eta_n)$ is a Cauchy sequence whenever $n \mapsto (\xi_n, \eta_n) \in X^N$ is. Thus $\mathcal{N}_\mu$ extends uniquely to a continuous function on LnL_μ. The fact that $\mathcal{N}_\mu$ is bilinear, antisymmetric and with values in $\mathrm{Tan}_\mu^\perp(\mathscr{P}_2(M))$ is obvious by continuity.

It remains only to prove (5.6). To this aim, by antisymmetry it is sufficient to prove that

$$\|\mathcal{N}_\mu(u, w)\|_\mu \leq \mathrm{L}(u)\|w\|_\mu,$$

holds whenever $\mathrm{L}(u) < \infty$. Choose $\xi, \eta \in \mathcal{V}(M)$. From the inequality

$$\|\mathcal{N}_\mu(\xi, \eta)\|_\mu = \|\mathrm{P}_\mu^\perp((\nabla\xi)^t \cdot \eta)\|_\mu \leq \|(\nabla\xi)^{\mathrm{t}} \cdot \eta\|_\mu \leq \mathrm{L}(\xi)\|\eta\|_\mu,$$

we get that the bound (5.6) holds for smooth vector fields. Now pick $u, w \in L^2_\mu$ such that $\mathrm{L}(u) < \infty$. By definition of $\mathrm{L}(u)$, we know that there exists a sequence $(\xi_n) \subset \mathcal{V}(M)$ such that $\mathrm{L}(\xi_n) \to \mathrm{L}(u)$ and $\|\xi_n - u\|_\mu \to 0$ as $n \to \infty$. Consider also a sequence $(\eta_n) \subset \mathcal{V}(M)$ such that $\|\eta_n - w\|_\mu \to 0$. The conclusion follows by letting n go to infinity in

$$\|\mathcal{N}_\mu(\xi_n, \eta_n)\|_\mu \leq \mathrm{L}(\xi_n)\|\eta_n\|_\mu.$$

□

Remark 5.10 (Something has to be Lipschitz) A priori, it is possible to think that the tensor $\mathcal{N}_\mu$ can be rigorously defined, or at least makes heuristically sense in the whole $[L^2_\mu]^2$, as it may be that there is some approximation process, smarter than the one proposed in inequalities (5.7), that lead to a good definition of $\mathcal{N}_\mu$ on $[L^2_\mu]^2$. Although we cannot disprove this, we strongly believe that such a generalization is not possible. We believe this, because some kind of Lipschitz condition on the vector fields comes out from two very different situations:

- in theorem 2.13 (based on proposition 2.10) to give a quantitative bound on the 'variation of tangent spaces',

- here, in the definition of $\mathcal{N}_\mu$.

Also, we already know that when we forget about Lipschitz conditions on vector fields 'something goes wrong': namely, the parallel transport doesn't exist.

Therefore, in the author's humble opinion, this Lipschitz conditions that appear here and there are not ad hoc hypothesis, but rather symptomatic of true geometrical obstructions[2]. ■

A standard approximation argument shows that that the equality

$$\mathcal{N}_\mu(u, w) = \mathrm{P}^\perp_\mu\big(\nabla u \cdot w\big),$$

holds for any $u, w \in L^2_\mu$ such that $\mathrm{L}(u) < \infty$ and u is differentiable $\mu-$a.e.. However, as already noticed, since we are making no assumption on μ, it may happen that a Lipschitz vector field u is *not* $\mu-$a.e. differentiable: still, it makes sense the object $\mathcal{N}_\mu(u, w)$ for any $w \in L^2_\mu$.

Definition 5.11 (The operators $\mathcal{O}_v(\cdot)$ and $\mathcal{O}^*_v(\cdot)$) *Given $v \in L^2_\mu$ Lipschitz, the map $w \mapsto \mathcal{N}_\mu(v, w)$ is a continuous map from L^2_μ into itself with norm bounded by $\mathrm{L}(v)$. We will denote this map with $\mathcal{O}_v(\cdot)$, so that $\mathcal{N}_\mu(v, w) = \mathcal{O}_v(w)$ for any $v, w \in L^2_\mu$ with $\mathrm{L}(v) < \infty$. We will use the notation $\mathcal{N}_\mu(v, w)$ when we want to highlight the (anti-)symmetry of the expression, and $\mathcal{O}_v(w)$ when we know that the vector v is Lipschitz. Also, we will denote by $\mathcal{O}^*_v(\cdot) : L^2_\mu \to L^2_\mu$ the adjoint map of $\mathcal{O}_v(\cdot)$: i.e., for any $w \in L^2_\mu$, the vector $\mathcal{O}^*_v(w) \in L^2_\mu$ is defined by*

$$\langle \mathcal{O}^*_v(w), u\rangle_\mu = \langle w, \mathcal{O}_v(u)\rangle_\mu, \qquad \forall u \in L^2_\mu.$$

As for $\mathcal{O}_v(\cdot)$, the norm of $\mathcal{O}^*_v(\cdot)$ is bounded by $\mathrm{L}(v)$. When v is smooth, the operator $\mathcal{O}^*_v(\cdot)$ reads as

$$w \mapsto \nabla v_t \cdot \mathrm{P}^\perp_{\mu_t}(w).$$

[2] up to our understanding, it is possible only a small enlargement of the domain of definition of $\mathcal{N}_\mu$, which passes through a different definition of Lipschitz constant. Define $\tilde{\mathrm{L}}_\mu(\xi) := \sup_{x \in \mathrm{supp}(\mu)} \|\nabla\xi(x)\|_{\mathrm{op}}$ for $\xi \in \mathcal{V}(M)$ and then $\tilde{\mathrm{L}}_\mu(v)$ for arbitrary $v \in L^2_\mu$ by lower semicontinuous relaxation. Then, with the same process it is possible to see that $\mathcal{N}_\mu$ can be defined on the set of couples (u, v) for which $\tilde{\mathrm{L}}(u) < \infty$ or $\tilde{\mathrm{L}}(v) < \infty$. The difference between the constant $\mathrm{L}(\xi)$ and $\tilde{\mathrm{L}}(\xi)$ relies on the set on which we are taking the supremum: in the first case it is the whole M, in the second only the support of μ. It is not hard to check that there exists measures μ and vectors $v \in L^2_\mu$ such that $\mathrm{L}(v) = +\infty$ and $\tilde{\mathrm{L}}(v) = 0$, so that this new approach would actually lead to a generalization of the definition of $\mathcal{N}_\mu$. However, neither the Cauchy-Lipschitz theorem 2.6, nor the proposition 2.10 work with this kind of Lipschitz constant, thus we preferred to avoid its introduction and to keep, also for the definition of $\mathcal{N}_\mu$ the 'classical' Lipschitz constant

There is a little abuse of notation in the definition of the operators $\mathcal{O}_v(\cdot)$ and $\mathcal{O}^*_v(\cdot)$, as we lost the reference to the measure μ, which of course plays a role in the definition. However, hopefully, in what follows the context should always clarify the measure, and leave no ambiguity.

Notice that by the properties of $\mathcal{N}_\mu$, for a fixed $u \in L^2_\mu$, the maps $v \mapsto \mathcal{O}_v(u)$ and $v \mapsto \mathcal{O}^*_v(u)$ are continuous in v w.r.t. convergence in L^2_μ plus uniform bound on the Lipschitz constant.

Observe that the adjoint of the operator $\mathcal{O}_v\left(\mathrm{P}^\perp_{\mu_t}(\cdot)\right) = \mathrm{P}^\perp_{\mu_t}(\mathcal{O}_v\left(\mathrm{P}^\perp_{\mu_t}(\cdot)\right))$ is the operator $\mathrm{P}^\perp_{\mu_t}(\mathcal{O}^*_v\left(\mathrm{P}^\perp_{\mu_t}(\cdot)\right)) = \mathrm{P}^\perp_{\mu_t}(\mathcal{O}^*_v(\cdot))$. In general, these two operators are different, however, if the vector field v is of the form $\nabla\varphi$, from the symmetry of $\nabla^2\varphi$ we have

$$\left\langle \mathrm{P}^\perp_\mu(\nabla^2\varphi \cdot \mathrm{P}^\perp_\mu(u_1)), u_2 \right\rangle_\mu = \left\langle u_1, \mathrm{P}^\perp_\mu(\nabla^2\varphi \cdot \mathrm{P}^\perp_\mu(u_2)) \right\rangle_\mu, \quad \forall u_1, u_2 \in L^2_\mu,$$

so that actually it holds

$$\mathcal{O}_{\nabla\varphi}\left(\mathrm{P}^\perp_\mu(\cdot)\right) = \mathrm{P}^\perp_\mu(\mathcal{O}^*_{\nabla\varphi}(\cdot)).$$

It is then natural to think that the vectors $v \in L^2_\mu$ such that

$$\mathcal{O}_v\left(\mathrm{P}^\perp_\mu(\cdot)\right) = \mathrm{P}^\perp_\mu(\mathcal{O}^*_v(\cdot)), \tag{5.8}$$

are precisely the vectors which are both Lipschitz and tangent: we will call vectors v for which the above is true, vectors with symmetric gradient.

We don't know whether Lipschitz tangent vector fields have symmetric gradients or not. What creates difficulties are situations like the one of example 2.4: indeed in that case the tangent vector field may be smooth, but its gradient can be not symmetric. Still, in that case the tangent space coincides with the L^2_μ space, so that the tensor $\mathcal{N}_\mu$ is identically 0 and equation (5.8) is still satisfied.

What we can prove is that the velocity vector field of a geodesic in $\mathcal{P}_c(M)$ have symmetric gradient. The result heavily relies on corollary 2.22.

Proposition 5.12 (Symmetry of ∇v_t along a geodesic) *Let $(\mu_t) \subset \mathcal{P}_c(M)$ be a geodesic. Then for every $t \in (0,1)$ it holds*

$$\mathcal{O}_{v_t}\left(\mathrm{P}^\perp_{\mu_t}(\cdot)\right) = \mathrm{P}^\perp_{\mu_t}(\mathcal{O}^*_{v_t}(\cdot)).$$

Proof Fix $t \in (0,1)$. By corollary 2.22 we know that there exists a sequence of smooth functions $(\phi^n_t) \subset C^\infty_c(M)$ such that

$$\lim_{n\to\infty} \|\nabla\phi^n_t - v_t\|_{\mu_t} = 0,$$
$$\sup_{n\to\infty} \mathrm{L}(\nabla\phi^n_t) < \infty.$$

Now choose $u_1, u_2 \in L^2_{\mu_t}$. Since ϕ^n_t is smooth for every n, we know that it holds

$$\left\langle \mathrm{P}^\perp_{\mu_t}\left(\nabla^2\phi^n_t \cdot \mathrm{P}^\perp_{\mu_t}(u^1)\right), \mathrm{P}^\perp_{\mu_t}(u_2) \right\rangle_{\mu_t} = \left\langle \mathrm{P}^\perp_{\mu_t}(u^1), \mathrm{P}^\perp_{\mu_t}\left(\nabla^2\phi^n_t \cdot \mathrm{P}^\perp_{\mu_t}(u_2)\right) \right\rangle_{\mu_t},$$

and since the sequence $(\nabla\phi^n_t)$ converges to v_t and has uniformly bounded Lipschitz constant, the above equation passes to limit as n goes to ∞ and we get the thesis. $\square$

Remark 5.13 (The case $M = \mathbb{R}^d$) Thanks to remark 2.23, the same conclusion of the above proposition holds in the Euclidean case, regardless of any compactness assumption on the supports of μ_t. ■

5.3 Calculus of derivatives

Here we calculate the total and covariant derivative of some basic kind of vector fields. Our first goal is to study the vector field $(\mathrm{P}_{\mu_t}(u_t))$ for a given (u_t) absolutely continuous and not necessarily tangent. The curve (μ_t) will be a fixed regular curve with velocity vector field (v_t).

Recall that since (μ_t) is regular, for any bounded vector field (u_t), the vector field

$$t \mapsto \mathcal{N}_{\mu_t}(u_t, v_t) = -\mathcal{O}_{v_t}(u_t)$$

is well defined for a.e. $t \in [0,1]$ and, due to inequality $\|\mathcal{N}_{\mu_t}(u_t, v_t)\|_{\mu_t} \leq \mathrm{L}(v_t)\|u_t\|_{\mu_t}$, is an L^1 vector field. The same is true for the vector field

$$t \mapsto \mathcal{O}^*_{v_t}(u_t).$$

The discussion of the previous section allows us to compute the covariant derivative of $(\mathrm{P}_{\mu_t}(u_t))$:

Theorem 5.14 *Let (u_t) be an absolutely continuous (not necessarily tangent) vector field along (μ_t). Then the vector field $(\mathrm{P}_{\mu_t}(u_t))$ is absolutely continuous as well and its covariant derivative is given by*

$$\frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(u_t) = \mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t - \mathcal{O}^*_{v_t}(u_t)\right).$$

Proof We start by proving the absolute continuity of $(\mathrm{P}_{\mu_t}(u_t))$. Choose $t < s \in [0,1]$ and recall equations (4.11) to get:

$$\begin{aligned}
\left\|\tau_s^t(\mathrm{P}_{\mu_s}(u_s)) - \mathrm{P}_{\mu_t}(u_t)\right\|_{\mu_t} &\leq \left\|\tau_s^t(\mathrm{P}_{\mu_s}(u_s)) - \mathrm{P}_{\mu_t}\Big(\tau_s^t(\mathrm{P}_{\mu_s}(u_s))\Big)\right\|_{\mu_t} \\
&\quad + \left\|\mathrm{P}_{\mu_t}\Big(\tau_s^t(\mathrm{P}_{\mu_s}(u_s))\Big) - \mathrm{P}_{\mu_t}(\tau_s^t(u_s))\right\|_{\mu_t} + \left\|\mathrm{P}_{\mu_t}(\tau_s^t(u_s)) - \mathrm{P}_{\mu_t}(u_t)\right\|_{\mu_t} \\
&\leq \left\|\mathrm{P}^{\perp}_{\mu_t}\Big(\tau_s^t(\mathrm{P}_{\mu_s}(u_s))\Big)\right\|_{\mu_t} + \left\|\mathrm{P}_{\mu_t}\Big(\tau_s^t(\mathrm{P}^{\perp}_{\mu_s}(u_s))\Big)\right\|_{\mu_t} + \left\|\tau_s^t(u_s) - u_t\right\|_{\mu_t} \\
&\leq 2NC\int_t^s \mathrm{L}(v_r)dr + \int_t^s \left\|\frac{\boldsymbol{d}}{dr}u_r\right\|_{\mu_r} dr,
\end{aligned}$$

where $N := \sup_{t\in[0,1]} \|u_t\|_{\mu_t}$ and $C := e^{\int_0^1 \mathrm{L}(v_t)dt} - 1$. Thus $(\mathrm{P}_{\mu_t}(u_t))$ is absolutely continuous.

Now we compute the covariant derivative by doing the same calculations we did at the end of the first section of this Chapter: now that we understood the domain of definition and the behavior of $\mathcal{N}_\mu$ we can give a rigorous justification to them. Let $\varphi \in C_c^\infty(M)$, consider the absolutely continuous and tangent vector field $(\nabla\varphi)$ and recall the Leibniz rule for covariant and total derivatives to get

$$\frac{d}{dt}\left\langle \mathrm{P}_{\mu_t}(u_t), \nabla\varphi\right\rangle_{\mu_t} = \left\langle \frac{\boldsymbol{D}}{dt}\big(\mathrm{P}_{\mu_t}(u_t)\big), \nabla\varphi\right\rangle_{\mu_t} + \left\langle \mathrm{P}_{\mu_t}(u_t), \frac{\boldsymbol{D}}{dt}\nabla\varphi\right\rangle_{\mu_t},$$

$$\frac{d}{dt}\left\langle u_t, \nabla\varphi\right\rangle_{\mu_t} = \left\langle \frac{\boldsymbol{d}}{dt}u_t, \nabla\varphi\right\rangle_{\mu_t} + \left\langle u_t, \frac{\boldsymbol{d}}{dt}\nabla\varphi\right\rangle_{\mu_t},$$

Since $(\nabla\varphi)$ is tangent, it holds $\left\langle \mathrm{P}_{\mu_t}(u_t), \nabla\varphi\right\rangle_{\mu_t} = \left\langle u_t, \nabla\varphi\right\rangle_{\mu_t}$ for every $t \in [0,1]$, thus the left hand sides of the above equations are equal for a.e. $t \in [0,1]$. From the equality of the right hand sides we get

$$\begin{aligned}
\left\langle \frac{\boldsymbol{D}}{dt}\big(\mathrm{P}_{\mu_t}(u_t)\big), \nabla\varphi\right\rangle_{\mu_t} &= \left\langle \frac{\boldsymbol{d}}{dt}u_t, \nabla\varphi\right\rangle_{\mu_t} + \left\langle u_t, \frac{\boldsymbol{d}}{dt}\nabla\varphi\right\rangle_{\mu_t} - \left\langle \mathrm{P}_{\mu_t}(u_t), \frac{\boldsymbol{D}}{dt}\nabla\varphi\right\rangle_{\mu_t} \\
&= \left\langle \frac{\boldsymbol{d}}{dt}u_t, \nabla\varphi\right\rangle_{\mu_t} + \left\langle \mathrm{P}^{\perp}_{\mu_t}(u_t), \mathrm{P}^{\perp}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}\nabla\varphi\right)\right\rangle_{\mu_t} \\
&= \left\langle \frac{\boldsymbol{d}}{dt}u_t, \nabla\varphi\right\rangle_{\mu_t} + \left\langle \mathrm{P}^{\perp}_{\mu_t}(u_t), \mathrm{P}^{\perp}_{\mu_t}\left(\nabla^2\varphi\cdot v_t\right)\right\rangle_{\mu_t} \\
&= \left\langle \mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t\right), \nabla\varphi\right\rangle_{\mu_t} - \left\langle \mathrm{P}^{\perp}_{\mu_t}(u_t), \mathcal{O}_{v_t}\left(\nabla\varphi\right)\right\rangle_{\mu_t} \\
&= \left\langle \mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t - \mathcal{O}^*_{v_t}\left(u_t\right)\right), \nabla\varphi\right\rangle_{\mu_t},
\end{aligned}$$

and the conclusion follows from the arbitrariness of $\varphi \in C^\infty_c(M)$. □

This theorem has important consequences in terms of description of the geometry of $(\mathscr{P}_2(M), W_2)$. The first is given in the next theorem, where we compute the difference between the total and the covariant derivative of a tangent and absolutely continuous vector field.

Theorem 5.15 *Let (μ_t) be a regular curve, (v_t) its velocity vector field and (u_t) a tangent and absolutely continuous vector field. Then it holds*

$$\frac{\boldsymbol{d}}{dt}u_t = \frac{\boldsymbol{D}}{dt}u_t - \mathcal{O}_{v_t}\left(u_t\right)$$

Proof Let (w_t) be an arbitrary absolutely continuous vector field, not necessarily tangent. Observe that we have

$$\frac{d}{dt}\left\langle u_t, w_t\right\rangle_{\mu_t} = \left\langle \frac{\boldsymbol{d}}{dt}u_t, w_t\right\rangle_{\mu_t} + \left\langle u_t, \frac{\boldsymbol{d}}{dt}w_t\right\rangle_{\mu_t},$$

$$\frac{d}{dt}\left\langle u_t, \mathrm{P}_{\mu_t}(w_t)\right\rangle_{\mu_t} = \left\langle \frac{\boldsymbol{D}}{dt}u_t, \mathrm{P}_{\mu_t}(w_t)\right\rangle_{\mu_t} + \left\langle u_t, \frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(w_t)\right\rangle_{\mu_t}.$$

Since (u_t) is tangent, we have $\left\langle u_t, w_t\right\rangle_{\mu_t} = \left\langle u_t, \mathrm{P}_{\mu_t}(w_t)\right\rangle_{\mu_t}$ for any $t \in [0,1]$, therefore the left hand sides of the above equations are equal. From the equality of the right hand sides and theorem 5.14 we get

$$\begin{aligned}\left\langle \frac{\boldsymbol{d}}{dt}u_t - \frac{\boldsymbol{D}}{dt}u_t, w_t \right\rangle_{\mu_t} &= \left\langle \frac{\boldsymbol{d}}{dt}u_t - \frac{\boldsymbol{D}}{dt}u_t, \mathrm{P}^{\perp}_{\mu_t}(w_t) \right\rangle_{\mu_t} \\ &= -\left\langle u_t, \frac{\boldsymbol{d}}{dt}w_t - \frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(w_t) \right\rangle_{\mu_t} \\ &= -\left\langle u_t, \mathrm{P}_{\mu_t}\Big(\frac{\boldsymbol{d}}{dt}w_t\Big) - \frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}(w_t) \right\rangle_{\mu_t} \\ &= -\left\langle u_t, \mathcal{O}^*_{v_t}(w_t) \right\rangle_{\mu_t} \\ &= -\left\langle \mathcal{O}_{v_t}(u_t), w_t \right\rangle_{\mu_t}\end{aligned}$$

And the conclusion follows by the arbitrariness of (w_t). □

Remark 5.16 (The operator $\frac{\boldsymbol{d}}{dt} - \frac{\boldsymbol{D}}{dt}$) Observe that what a consequence of the above theorem, is the fact that that the map $(u_t) \mapsto \frac{\boldsymbol{d}}{dt}u_t - \frac{\boldsymbol{D}}{dt}u_t$ is a zero order map, rather then a first order one: indeed the result of this difference does not depend on any derivative of (u_t) but just on its value at the time t. This could have been guessed also from the following calculation: let $f : [0, 1] \to \mathbb{R}$ a smooth function, and consider the vector field $(f(t)u_t)$. Then we have

$$\frac{\boldsymbol{d}}{dt}(f(t)u_t) - \frac{\boldsymbol{D}}{dt}(f(t)u_t) = f'(t)u_t + f(t)\frac{\boldsymbol{d}}{dt}u_t - f'(t)u_t - f(t)\frac{\boldsymbol{D}}{dt}u_t = f(t)\left(\frac{\boldsymbol{d}}{dt}u_t - \frac{\boldsymbol{D}}{dt}u_t\right),$$

which suggest that $\frac{\boldsymbol{d}}{dt}u_t - \frac{\boldsymbol{D}}{dt}u_t$ is a tensor in u_t (as it actually is). This fact is not surprising, as a similar thing happens in the classical case of Riemannian manifolds embedded in some higher dimensional Riemannian structure, where the difference between the derivative in the ambient space and the covariant derivative produces a tensor in the object being derivated.

However, there is an important difference with the classical case: to see which one, let us write the statement of the above theorem as

$$\frac{\boldsymbol{d}}{dt}u_t - \frac{\boldsymbol{D}}{dt}u_t = \mathcal{N}_{\mu_t}(u_t, v_t).$$

Thus the difference between the total derivative and the covariant one is an antisymmetric operator. In the case of a Riemannian manifold embedded in some higher dimensional manifold, the difference between the derivative in the ambient space and the covariant derivative is actually a symmetric (!) operator, not an antisymmetric one (see e.g. Chapter 6.2 of [6]). Given that we built the second order theory on $(\mathscr{P}_2(M), W_2)$ on the analogy of an embedded manifold, this difference is a bit strange.

Up to our understanding, the point here is simply that there is no true ambient manifold on which $(\mathscr{P}_2(M), W_2)$ is embedded: the analogy stops at this point. If one tries, for instance, to follow the proof of proposition 2.1 of Chapter 6 of [6] (the one in which symmetry is proven), he discovers that some of the objects introduced in the calculations are meaningless in our setting: namely, there is no way to define the (analogous of the) Lie Bracket of two vector fields defined in the (non existing) ambient space. ■

Remark 5.17 (Infinitesimal variation of vector fields) If we apply the previous theorem to the vector field $u_t := \mathcal{T}_0^t(u)$, where $u \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ and $\mathcal{T}_t^s$ are the parallel transport maps along (μ_t) we get

$$\frac{\boldsymbol{d}}{dt}\mathcal{T}_0^t(u) = \mathcal{N}_{\mu_t}(\mathcal{T}_0^t(u), v_t).$$

Therefore we see that the tensor $\mathcal{N}_{\mu_t}$ here is describing 'how the tangent space is varying along (μ_t)'. This is why we said at the beginning of the chapter that the tensor $\mathcal{N}_\mu$ describes the infinitesimal variation of the tangent spaces. ■

A direct consequence of the last two theorems are the formulas for the total derivatives of $\mathrm{P}_{\mu_t}(u_t)$ and $\mathrm{P}^\perp_{\mu_t}(u_t)$:

Proposition 5.18 *Let (μ_t) be a regular curve, (v_t) its velocity vector field and (u_t) an absolutely continuous vector field along it. Then both $(\mathrm{P}_{\mu_t}(u_t))$ and $(\mathrm{P}^\perp_{\mu_t}(u_t))$ are absolutely continuous and their total derivatives are given by:*

$$\frac{\boldsymbol{d}}{dt}\mathrm{P}_{\mu_t}(u_t) = \mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t - \mathcal{O}^*_{v_t}(u_t)\right) - \mathcal{O}_{v_t}(\mathrm{P}_{\mu_t}(u_t)), \tag{5.9a}$$

$$\frac{\boldsymbol{d}}{dt}\mathrm{P}^\perp_{\mu_t}(u_t) = \mathrm{P}^\perp_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t\right) + \mathrm{P}_{\mu_t}\big(\mathcal{O}^*_{v_t}(u_t)\big) + \mathcal{O}_{v_t}(\mathrm{P}_{\mu_t}(u_t)). \tag{5.9b}$$

Proof The absolute continuity of $(\mathrm{P}_{\mu_t}(u_t))$ was already established in theorem 5.14; the formula for its total derivative then follows from 5.15.

The fact that $(\mathrm{P}^\perp_{\mu_t}(u_t))$ is absolutely continuous follows from $u_t = \mathrm{P}_{\mu_t}(u_t) + \mathrm{P}^\perp_{\mu_t}(u_t)$; the formula for its derivative comes differentiating this trivial identity. □

Now we pass to calculus of the total derivative of $(\mathcal{N}_{\mu_t}(u_t, w_t))$ for given absolutely continuous vector fields (u_t), (w_t). Let us observe that in order to obtain the absolute continuity of $(\mathcal{N}_{\mu_t}(u_t, w_t))$ we will need to impose some additional condition on either (u_t) or (w_t): to understand why, observe - just at an heuristic level - that for the total derivative of $(\mathcal{N}_{\mu_t}(u_t, w_t))$ we expect something like

$$\frac{\boldsymbol{d}}{dt}\mathcal{N}_{\mu_t}(u_t, w_t) = \mathcal{N}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t, w_t\right) + \mathcal{N}_{\mu_t}\left(u_t, \frac{\boldsymbol{d}}{dt}w_t\right) + \left(\begin{array}{l}\text{some tensor - which we may think}\\ \text{as the derivative of } \mathcal{N}_{\mu_t} \text{ -}\\ \text{applied to the couple } (u_t, w_t)\end{array}\right).$$

Forget about the last object and look at the first two addends: given that the domain of definition of $\mathcal{N}_{\mu_t}$ is not the whole $[L^2_{\mu_t}]^2$, in order for the above formula to make sense, we should ask that in each of the couples $(\frac{\boldsymbol{d}}{dt}u_t, w_t)$ and $(u_t, \frac{\boldsymbol{d}}{dt}w_t)$, at least one vector is Lipschitz. We will assume that $\{\int_0^1 \mathrm{L}(u_t)dt < \infty$ and $\int_0^1 \mathrm{L}(\frac{\boldsymbol{d}}{dt}u_t)dt < +\infty\}$.

Having said this, we may turn to the precise formula. Let us just mention that when everything is smooth, the derivative of $\mathcal{N}_{\mu_t}(\xi_t, \eta_t)$ may be written as

$$\begin{aligned}\frac{\boldsymbol{d}}{dt}\mathrm{P}^\perp_{\mu_t}(\nabla\xi_t \cdot \eta_t) = {} & \mathrm{P}^\perp_{\mu_t}\left(\left(\nabla\frac{\boldsymbol{d}}{dt}\xi_t\right)^{\mathrm{t}} \cdot \eta_t\right) + \mathrm{P}^\perp_{\mu_t}\left(\nabla\xi_t^{\mathrm{t}} \cdot \left(\nabla\frac{\boldsymbol{d}}{dt}\eta_t\right)\right)\\ & - \mathrm{P}^\perp_{\mu_t}\big(R(\xi_t, \eta_t)v_t\big) - \mathrm{P}^\perp_{\mu_t}\left(\nabla v_t^{\mathrm{t}} \cdot \mathrm{P}^\perp_{\mu_t}(\nabla\xi_t^{\mathrm{t}} \cdot \eta_t)\right) + \mathrm{P}_{\mu_t}\Big(\nabla v_t \cdot \mathrm{P}^\perp_{\mu_t}(\nabla\xi_t^{\mathrm{t}} \cdot \eta_t)\Big),\end{aligned}$$

where of course the total derivative of (ξ_t) and (η_t) may be computed by formula (3.3).

In particular, the curvature tensor R appears; let us spend two words on it. Given $\mu \in \mathscr{P}_2(M)$ and three vector fields $u_1, u_2, u_3 \in L^2_\mu$, we may consider the vector field $R(u_1, u_2)u_3$, which is well defined μ-a.e. by $x \mapsto R(u_1(x), u_2(x))u_3(x)$. The question is: how do we know that $R(u_1, u_2)u_3$ belongs to L^2_μ? In general, if $\mu \notin \mathcal{P}_c(M)$ we don't have an answer, as the curvature tensor may be not uniformly bounded on the support of μ. If μ has compact support, then from the inequality

$$|R(u_1(x), u_2(x))u_3(x)| \leq \mathcal{C}(\mathrm{supp}(\mu))|u_1(x)||u_2(x)||u_3(x)|,$$

we can conclude that $R(u_1, u_2)u_3 \in L^2_\mu$ if we know that two of the three vector fields u_1, u_2, u_3 are (essentially) bounded. In the next lemma we show how the essential bound follows from the Lipschitz condition (recall that for Lipschitz vector fields defined on measures with compact support the essential bound is a true bound - see also definition 2.2).

Lemma 5.19 *Let $(\mu_t) \subset \mathcal{P}_c(M)$ be a regular curve and (u_t) an L^1 vector field along it. Then it holds*

$$\mathrm{S}(u_t) \leq \mathrm{L}(u_t)D + \|u_t\|_{\mu_t}, \qquad \forall t \in [0,1], \tag{5.10}$$

D being the diameter of $\cup_t \mathrm{supp}(\mu_t)$.

Also, if (u_t) is absolutely continuous and $\int_0^1 \mathrm{L}(\frac{\boldsymbol{d}}{dt}u_t)dt < \infty$, then

$$\sup_{t\in[0,1]} \mathrm{S}(u_t) \leq \mathrm{S}(u_0) + D\int_0^1 \mathrm{L}\left(\frac{\boldsymbol{d}}{ds}u_s\right) ds + N, \tag{5.11}$$

where $N := \sup_{t\in[0,1]} \|u_t\|_{\mu_t}$.

Proof Recall that for Lipschitz vector field we assume that we are dealing with its continuous representative. Fix $t \in [0,1]$ and $x \in \mathrm{supp}(\mu_t)$. A simple approximation argument shows that

$$|u_t(y)| \leq |u_t(x)| + \mathrm{L}(u_t)\mathrm{d}(x,y),$$

thus we have $\mathrm{S}(u_t) \leq |u_t(x)| + \mathrm{L}(u_t)D$, and hence, by integration

$$\mathrm{S}(u_t) \leq \mathrm{L}(u_t)D + \int |u_t(x)|d\mu(x) \leq \mathrm{L}(u_t)D + \|u_t\|_{\mu_t}.$$

For the second claim, observe that considering the vector field $t \mapsto \tau_t^0(u_t) \in L^2_{\mu_0}$ it is immediate to verify that

$$\mathrm{S}(u_t) \leq \mathrm{S}(u_0) + \int_0^t \mathrm{S}\left(\frac{\boldsymbol{d}}{ds}u_s\right) ds,$$

therefore using (5.10) we obtain

$$\mathrm{S}(u_t) \leq \mathrm{S}(u_0) + \int_0^1 \mathrm{S}\left(\frac{\boldsymbol{d}}{ds}u_t\right) ds \leq \mathrm{S}(u_0) + D\int_0^1 \mathrm{L}\left(\frac{\boldsymbol{d}}{ds}u_s\right) ds + N, \qquad \forall t \in [0,1]$$

□

Lemma 5.20 *Let $(\mu_t) \subset \mathscr{P}_2(M)$ be a regular curve and (u_t) an absolutely continuous vector field satisfying $\int_0^1 \mathrm{L}(u_t)dt < \infty$ and $\int_0^1 \mathrm{L}(\frac{\boldsymbol{d}}{dt}u_t)dt < \infty$. Then it holds*

$$\sup_{t\in[0,1]} \mathrm{L}(u_t) < \infty.$$

Proof Let $\overline{u}_t := \tau_t^0(u_t) \in L^2_{\mu_0}$. It is immediate to verify that

$$\begin{aligned} \mathrm{L}(\overline{u}_t) &\leq \mathrm{L}(u_t)\mathrm{Lip}\big(\mathbf{T}(0,t,\cdot)\big), \\ \mathrm{L}(\frac{d}{dt}\overline{u}_t) &\leq \mathrm{L}\Big(\frac{\boldsymbol{d}}{dt}u_t\Big)\mathrm{Lip}\big(\mathbf{T}(0,t,\cdot)\big), \end{aligned} \tag{5.12}$$

so that $\int_0^1 \mathrm{L}(\overline{u}_t)dt < \infty$ and $\int_0^1 \mathrm{L}\big(\frac{d}{dt}\overline{u}_t\big)dt < \infty$. For the vector field $t \mapsto \overline{u}_t \in L^2_{\mu_0}$ is is obvious that

$$\mathrm{L}(\overline{u}_t) \leq \mathrm{L}(u_s) + \int_t^s \mathrm{L}\Big(\frac{d}{dt}\overline{u}_r\Big)dr \leq \mathrm{L}(u_s) + \int_0^1 \mathrm{L}\Big(\frac{d}{dt}\overline{u}_r\Big)dr, \qquad \forall t < s \in [0,1]$$

and similarly for $t > s$, therefore integrating over s we obtain

$$\mathrm{L}(\overline{u}_t) \leq \int_0^1 \mathrm{L}(u_s)ds + \int_0^1 \mathrm{L}\Big(\frac{d}{dt}\overline{u}_r\Big)dr < \infty,$$

which provides a uniform bound on $\mathrm{L}(\overline{u}_t)$. The conclusion follows from

$$\mathrm{L}(u_t) \leq \mathrm{L}(\overline{u}_t)\mathrm{Lip}\big(\mathbf{T}(t,0,\cdot)\big).$$

□

Theorem 5.21 *Let $(\mu_t) \subset \mathcal{P}_c(M)$ be a regular curve and (u_t), (w_t) two absolutely continuous vector fields along it. Assume that $\int_0^1 \mathrm{L}(u_t)dt < \infty$ and $\int_0^1 \mathrm{L}(\frac{\boldsymbol{d}}{dt}u_t)dt < +\infty$. Then $(\mathcal{N}_{\mu_t}(u_t, w_t))$ is absolutely continuous and it holds:*

$$\begin{aligned} \frac{\boldsymbol{d}}{dt}\mathcal{N}_{\mu_t}(u_t, w_t) =& \mathcal{N}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}u_t, w_t\right) + \mathcal{N}_{\mu_t}\left(u_t, \frac{\boldsymbol{d}}{dt}w_t\right) \\ &- \mathrm{P}^\perp_{\mu_t}\Big(R(u_t, w_t)v_t\Big) + \mathcal{N}_{\mu_t}\Big(\mathcal{N}_{\mu_t}(u_t, w_t), v_t\Big) + \mathrm{P}_{\mu_t}\Big(\mathcal{O}^*_{v_t}\left(\mathcal{N}_{\mu_t}(u_t, w_t)\right)\Big). \end{aligned} \tag{5.13}$$

Proof The delicate point we need to take care of, and which prevents us from applying directly formula (5.9b) to the vector field $\mathrm{P}^\perp_\mu(\nabla u_t^{\mathrm{t}} \cdot w_t)$, is the fact that $\nabla u_t^{\mathrm{t}} \cdot w_t$ is *not* a well defined vector in $L^2_{\mu_t}$. What makes sense is only its projection onto the Normal space. Therefore, we proceed by first proving the formula for smooth vector fields, and then arguing by approximation.

Step 1: why the formula works. Let $\xi, \eta \in \mathcal{V}(M)$, $\mu \in \mathcal{P}_c(M)$ and $v \in L^2_\mu$ a Lipschitz vector field. We claim that it holds:

$$\begin{aligned} &\mathrm{P}^\perp_\mu\Big(\nabla\big(\nabla\xi^{\mathrm{t}} \cdot \eta\big) \cdot v\Big) + \mathcal{O}_v\left(\mathrm{P}_\mu(\nabla\xi^{\mathrm{t}} \cdot \eta)\right) \\ &\quad = \mathcal{N}_\mu\Big(\nabla\xi \cdot v, \eta\Big) + \mathcal{N}_\mu\Big(\xi, \nabla\eta \cdot v\Big) - \mathrm{P}^\perp_\mu\Big(R(\xi, \eta)v\Big) - \mathcal{O}_v\left(\mathcal{N}_\mu(\xi, \eta)\right). \end{aligned} \tag{5.14}$$

Start observing that both sides are continuous in v w.r.t. convergence in L^2_μ plus uniform bound on the Lipschitz constant. Thus we may assume $v \in \mathcal{V}(M)$.

Now, to prove the formula, we need to identify the value of $\nabla\big(\nabla\xi^{\mathrm{t}}\cdot\eta\big)\cdot v$. In order to do so, choose an auxiliary vector field $\tilde\eta \in \mathcal{V}(M)$ and observe that it holds:

$$\begin{aligned}\big\langle\nabla\big(\nabla\xi^{\mathrm{t}}\cdot\eta\big)\cdot v,\tilde\eta\big\rangle &= \nabla\big\langle\nabla\xi^{\mathrm{t}}\cdot\eta,\tilde\eta\big\rangle\cdot v - \big\langle\nabla\xi^{\mathrm{t}}\cdot\eta,\nabla\tilde\eta\cdot v\big\rangle\\ &= \nabla\langle\eta,\nabla\xi\cdot\tilde\eta\rangle\cdot v - \langle\eta,\nabla\xi\cdot\nabla\tilde\eta\cdot v\rangle\\ &= \langle\nabla\eta\cdot v,\nabla\xi\cdot\tilde\eta\rangle + \langle\eta,\nabla(\nabla\xi\cdot\tilde\eta)\cdot v\rangle - \langle\eta,\nabla\xi\cdot\nabla\tilde\eta\cdot v\rangle\\ &= \langle\nabla\eta\cdot v,\nabla\xi\cdot\tilde\eta\rangle + \big\langle\eta,\nabla^2\xi(\tilde\eta,v)\big\rangle.\end{aligned} \tag{5.15}$$

We want to get rid of the term involving $\nabla^2\xi$. Observe that

$$\begin{aligned}\big\langle\nabla\left(\nabla\xi\cdot v\right)^{\mathrm{t}}\cdot\eta,\tilde\eta\big\rangle &= \langle\eta,\nabla\left(\nabla\xi\cdot v\right)\cdot\tilde\eta\rangle\\ &= \big\langle\eta,\nabla^2\xi(v,\tilde\eta)+\nabla\xi\cdot\nabla v\cdot\tilde\eta\big\rangle\\ &= \big\langle\eta,\nabla^2\xi(\tilde\eta,v)\big\rangle + \langle R(v,\tilde\eta)\xi,\eta\rangle + \langle\eta,\nabla\xi\cdot\nabla v\cdot\tilde\eta\rangle.\end{aligned}$$

Substitute the value of $\big\langle\eta,\nabla^2\xi(\tilde\eta,v)\big\rangle$ just found in (5.15) and use the arbitrariness of $\tilde\eta$ to get

$$\nabla\big(\nabla\xi^{\mathrm{t}}\cdot\eta\big)\cdot v = \nabla(\nabla\xi\cdot v)^{\mathrm{t}}\cdot\eta + \nabla\xi^{\mathrm{t}}\cdot\nabla\eta\cdot v - R(\xi,\eta)v - \nabla v^{\mathrm{t}}\cdot\nabla\xi^{\mathrm{t}}\cdot\eta.$$

Taking the projection of both members onto the normal space at μ and observing that

$$\mathrm{P}^\perp_\mu(\nabla v^{\mathrm{t}}\cdot\nabla\xi^{\mathrm{t}}\cdot\eta) = \mathrm{P}^\perp_\mu\Big(\nabla v^{\mathrm{t}}\cdot\mathrm{P}_\mu(\nabla\xi^{\mathrm{t}}\cdot\eta)\Big) + \mathrm{P}^\perp_\mu\Big(\nabla v^{\mathrm{t}}\cdot\mathrm{P}^\perp_\mu(\nabla\xi^{\mathrm{t}}\cdot\eta)\Big)$$

we get equation (5.14).

Step 2: rigorous justification for smooth vector fields. Let us come back to our derivative along the regular curve (μ_t). Assume that the vector fields (u_t) and (w_t) are of the form (ξ), (η) respectively for some $\xi,\eta \in \mathcal{V}(M)$. Since $\nabla\xi^{\mathrm{t}}\cdot\eta \in \mathcal{V}(M)$, by proposition 5.18 we know that $(\mathrm{P}^\perp_{\mu_t}(\nabla\xi^{\mathrm{t}}\cdot\eta))$ is an absolutely continuous vector field along (μ_t). Apply formula (5.9b) to get that for a.e. $t \in [0,1]$ it holds

$$\frac{\boldsymbol{d}}{dt}\mathrm{P}^\perp_{\mu_t}(\nabla\xi^{\mathrm{t}}\cdot\eta) = \mathrm{P}^\perp_{\mu_t}\Big(\nabla(\nabla\xi^{\mathrm{t}}\cdot\eta)\cdot v_t\Big) + \mathrm{P}_{\mu_t}\Big(\mathcal{O}^*_{v_t}\left(\nabla\xi^{\mathrm{t}}\cdot\eta\right)\Big) + \mathcal{O}_{v_t}\Big(\mathrm{P}_\mu(\nabla\xi^{\mathrm{t}}\cdot\eta)\Big).$$

Since v_t is Lipschitz for a.e. $t \in [0,1]$, applying equation (5.14) we get (5.13).

Thus the thesis is proved for vector fields of the kind (ξ), (η). It is immediate to verify that more generally, the formula holds for vector fields of the kind $(\xi_t), (\eta_t) \in \mathcal{V}(M\times[0,1])$ (we avoided time dependency in our calculations just to focus on the most important aspects of the proof).

Step 3: first passage to the limit. Let (u_t) be given as in the hypothesis of the theorem, and $(\eta_t) \in \mathcal{V}(M\times[0,1])$. Use proposition 3.21 to find a sequence $n \mapsto (\xi^n_t) \in \mathcal{V}(M\times[0,1])$ such that $\|\xi^n_t - u_t\|_{\mu_t} \to 0$ as $n\to\infty$ uniformly in t, $t \mapsto \big\|\frac{\boldsymbol{d}}{dt}\xi^n_t\big\|_{\mu_t}$ converges to $t \mapsto \big\|\frac{\boldsymbol{d}}{dt}u_t\big\|_{\mu_t}$ as $n\to\infty$ in $L^1(0,1)$ and $t \mapsto \mathrm{L}(\xi^n_t)$ converges to $t\mapsto \mathrm{L}(u_t)$ as $n\to\infty$ in $L^1(0,1)$. Also, it is easy to check that we can choose ξ^n_0 such that $\mathrm{S}(\xi^n_0)\to\mathrm{S}(u_0)$ (the fact that $\mathrm{S}(u_0)<\infty$ follows from lemmata 5.20 and 5.19).

From the result of the previous step we know that $(\mathcal{N}_{\mu_t}(\xi_t^n, \eta_t))$ is an absolutely continuous vector field for every $n \in \mathbb{N}$. Also, we know that $\mathcal{N}_{\mu_t}(\xi_t^n, \eta_t) \to \mathcal{N}_{\mu_t}(u_t, \eta_t)$ for every $t \in [0,1]$ as $n \to \infty$ and that the derivative of $\mathcal{N}_{\mu_t}(\xi_t^n, \eta_t)$ converges pointwise to the right hand side of (5.13) with $w_t = \eta_t$. Therefore, in order to prove formula (5.13) for the couple (u_t, η_t) we need only to show that the sequence of derivatives $n \mapsto (\frac{\boldsymbol{d}}{dt}\mathcal{N}_{\mu_t}(\xi_t^n, \eta_t))$ is dominated in $L^1(0,1)$.

Let K be a compact set containing $\cup_t \operatorname{supp}(\mu_t)$ and observe that we have the bound

$$\begin{aligned} &\left\| \frac{\boldsymbol{d}}{dt}\mathcal{N}_{\mu_t}(\xi_t^n, \eta_t) \right\|_{\mu_t} \\ &\quad \le \underbrace{\left\| \frac{\boldsymbol{d}}{dt}\xi_t^n \right\|_{\mu_t} \mathrm{L}(\eta_t)}_{(\mathbf{A})} + \underbrace{\|\xi_t^n\|_{\mu_t} \mathrm{L}\left(\frac{\boldsymbol{d}}{dt}\eta_t\right)}_{(\mathbf{B})} + \underbrace{\mathcal{C}(K)\mathrm{S}(\xi_t^n)\|\eta_t\|_{\mu_t}\mathrm{S}(v_t)}_{(\mathbf{C})} + 2\underbrace{\mathrm{L}(v_t)\|\xi_t^n\|_{\mu_t}\mathrm{L}(\eta_t)}_{(\mathbf{D})}. \end{aligned} \tag{5.16}$$

We are going to show that each of the 4 addends is dominated as function of t uniformly in n.

(**A**) This is obvious, as $\sup_t \mathrm{L}(\eta_t) < \infty$ and $n \mapsto (\|\frac{\boldsymbol{d}}{dt}\xi_t^n\|_{\mu_t})$ converges to $(\|\frac{\boldsymbol{d}}{dt}u_t\|)$ in $L^1(0,1)$.

(**B**) The fact that $\|\xi_t^n\|_{\mu_t} \to \|u_t\|_{\mu_t}$ uniformly on t and the continuity of (u_t) gives that first factor is uniformly bounded in t, n. For the second observe that

$$\mathrm{L}\left(\frac{\boldsymbol{d}}{dt}\eta_t\right) = \mathrm{L}\left(\partial_t \eta_t + \nabla \eta_t \cdot v_t\right) \le \mathrm{L}(\partial_t \eta_t) + \tilde{C}\mathrm{S}(v_t) + \mathrm{L}(\eta_t)\mathrm{L}(v_t),$$

where $\tilde{C}$ is a uniform bound on the second order derivatives of η_t. By smoothness we have $\sup_t\{\mathrm{L}(\partial_t \eta_t), \mathrm{L}(\eta_t)\} < +\infty$ and by the bound (5.10) and the fact that $\int_0^1 \mathrm{L}(v_t)dt < \infty$ we have $\int_0^1 \mathrm{S}(v_t)dt < \infty$. Thus we have that $\int_0^1 \mathrm{L}(\frac{\boldsymbol{d}}{dt}\eta_t)dt < \infty$ and the conclusion.

(**C**) By the bound (5.11) we have

$$\sup_{t\in[0,1]} \mathrm{S}(\xi_t^n) \le \mathrm{S}(\xi_0^n) + N + D\int_0^1 \mathrm{L}(\frac{\boldsymbol{d}}{ds}\xi_s^n)ds,$$

for some N, D independent on n (recall that $\|\xi_t^n\|_{\mu_t}$ converges to $\|u_t\|_{\mu_t}$ uniformly on t). By smoothness, it holds

$$\sup_{t\in[0,1]} \|\eta_t\|_{\mu_t} < \infty,$$

and, as before by (5.10), we have

$$\int_0^1 \mathrm{S}(v_t)dt \le D\int_0^1 \mathrm{L}(v_t)dt + \int_0^1 \|v_t\|_{\mu_t}dt < \infty.$$

The conclusion follows from the fact that $(\mathrm{L}(\frac{\boldsymbol{d}}{dt}\xi_t^n))$ converges to $(\mathrm{L}(\frac{\boldsymbol{d}}{dt}u_t))$ in $L^1(0,1)$ and $\mathrm{S}(\xi_0^n) \to \mathrm{S}(u_0)$.

(**D**) We already noticed that $\sup_{n,t}\|\xi_t^n\|_{\mu_t} < \infty$, that $\sup_t \|\eta_t\|_{\mu_t} < \infty$ and that $\int_0^1 \mathrm{L}(v_t)dt < \infty$. There is nothing more to add.

Step 4: second passage to the limit. Use proposition 3.20 to find a sequence $n \mapsto (\eta_t^n) \in \mathcal{V}(M \times [0,1])$ which converges to (w_t) in the sense that: $\|\eta_t^n - w_t\|_{\mu_t} \to 0$ as $n \to \infty$ uniformly on t, and the sequence of functions $t \mapsto \|\frac{\boldsymbol{d}}{dt}\eta_t^n\|_{\mu_t}$ converges to $t \mapsto \|\frac{\boldsymbol{d}}{dt}w_t\|_{\mu_t}$ in $L^1(0,1)$.

Arguing as before, we need only to show that $n \mapsto (\|\frac{\boldsymbol{d}}{dt}\mathcal{N}_{\mu_t}(u_t,\eta_t^n)\|_{\mu_t})$ is a dominated sequence in $L^1(0,1)$. By the result of the previous step we know that $(\mathcal{N}_{\mu_t}(u_t,\eta_t^n))$ is absolutely continuous and that its derivative is given by formula (5.13), thus we know that

$$\begin{aligned}
&\left\|\frac{\boldsymbol{d}}{dt}\mathcal{N}_{\mu_t}(u_t,\eta_t^n)\right\|_{\mu_t} \\
&\qquad \le \mathrm{L}\left(\frac{\boldsymbol{d}}{dt}u_t\right)\|\eta_t^n\|_{\mu_t} + \mathrm{L}(u_t)\left\|\frac{\boldsymbol{d}}{dt}\eta_t^n\right\|_{\mu_t} + \mathcal{C}(K)\mathrm{S}(u_t)\|\eta_t^n\|_{\mu_t}\mathrm{S}(v_t) + 2\mathrm{L}(v_t)\mathrm{L}(u_t)\|\eta_t^n\|_{\mu_t}.
\end{aligned}$$

With arguments similar to those used in the previous step we know that $\sup_{n,t}\|\eta_t^n\|_{\mu_t} < \infty$, that $\sup_t \mathrm{S}(u_t) < \infty$ and that the function $t \mapsto \mathrm{S}(v_t)$ is integrable. Thus to conclude we need only to show that $\sup_t \mathrm{L}(u_t) < \infty$. To prove this, let $\overline{u}_t := \tau_t^0(u_t)$ and observe that $\mathrm{L}(u_t) \le \mathrm{L}(\overline{u}_t)\mathrm{Lip}(\mathbf{T}(t,0,\cdot))$ and $\mathrm{L}(\frac{d}{dt}\overline{u}_t) \le \mathrm{L}(\frac{\boldsymbol{d}}{dt}u_t)\mathrm{Lip}(\mathbf{T}(0,t,\cdot))$. Therefore from $\mathrm{L}(\overline{u}_t) \le \mathrm{L}(\overline{u}_0) + \int_0^t \mathrm{L}(\frac{d}{ds}\overline{u}_s)ds$ (which is obvious) we get

$$\begin{aligned}
\mathrm{L}(u_t) \le \mathrm{L}(\overline{u}_t)\mathrm{Lip}(\mathbf{T}(t,0,\cdot)) &\le e^{\int_0^1 \mathrm{L}(v_s)ds}\left(\mathrm{L}(\overline{u}_0) + \int_0^1 \mathrm{L}\Big(\frac{d}{ds}\overline{u}_s\Big)ds\right) \\
&\le e^{\int_0^1 \mathrm{L}(v_s)ds}\left(\mathrm{L}(u_0) + \int_0^1 \mathrm{L}\Big(\frac{\boldsymbol{d}}{ds}u_s\Big)\mathrm{Lip}\big(\mathbf{T}(0,s,\cdot)\big)ds\right) \\
&\le e^{\int_0^1 \mathrm{L}(v_s)ds}\left(\mathrm{L}(u_0) + e^{\int_0^1 \mathrm{L}(v_s)ds}\int_0^1 \mathrm{L}\Big(\frac{\boldsymbol{d}}{ds}u_s\Big)ds\right), \qquad \forall t \in [0,1].
\end{aligned}$$

The conclusion follows. □

Remark 5.22 (The case $M = \mathbb{R}^d$) In this proposition we needed to assume the measures μ_t's to be compactly supported to be sure that the vector field $R(u_t,w_t)v_t$ belongs to $L^2_{\mu_t}$. If $M = \mathbb{R}^d$, or more generally M is a flat manifold, then it is possible to see, with the same arguments, that everything is true without such compactness assumption. ■

An immediate consequence of this theorem are the formulas for the total derivatives of $(\mathcal{O}_{v_t}(u_t))$ and $(\mathcal{O}^*_{v_t}(u_t))$.

Corollary 5.23 *Let (μ_t) be a regular curve and assume that its velocity vector field (v_t) satisfies:*

$$\int_0^1 \mathrm{L}\left(\frac{\boldsymbol{d}}{dt}v_t\right)dt < \infty.$$

Then for every absolutely continuous vector field (u_t) both $(\mathcal{O}_{v_t}(u_t))$ and $(\mathcal{O}^*_{v_t}(u_t))$ are absolutely continuous and their total derivatives are given by:

$$\begin{aligned}
\frac{\boldsymbol{d}}{dt}\mathcal{O}_{v_t}(u_t) &= \mathcal{O}_{\frac{\boldsymbol{d}}{dt}v_t}(u_t) + \mathcal{O}_{v_t}\left(\frac{\boldsymbol{d}}{dt}u_t\right) - \mathrm{P}^{\perp}_{\mu_t}\Big(R(v_t,u_t)v_t\Big) - \mathcal{O}_{v_t}(\mathcal{O}_{v_t}(u_t)) + \mathrm{P}_{\mu_t}\Big(\mathcal{O}^*_{v_t}(\mathcal{O}_{v_t}(u_t))\Big) \\
\frac{\boldsymbol{d}}{dt}\mathcal{O}^*_{v_t}(u_t) &= \mathcal{O}^*_{\frac{\boldsymbol{d}}{dt}v_t}(u_t) + \mathcal{O}^*_{v_t}\left(\frac{\boldsymbol{d}}{dt}u_t\right) - R\Big(v_t,\mathrm{P}^{\perp}_{\mu_t}(u_t)\Big)v_t - \mathcal{O}^*_{v_t}\big(\mathcal{O}^*_{v_t}(u_t)\big) + \mathcal{O}^*_{v_t}(\mathcal{O}_{v_t}(\mathrm{P}_{\mu_t}(u_t)))
\end{aligned} \tag{5.17}$$

Proof The first formula follows directly from theorem 5.21, the second from the fact that $\mathcal{O}^*_{v_t}(\cdot)$ is the adjoint of $\mathcal{O}_{v_t}(\cdot)$. □

5.4 Smoothness of time dependent operators

The discussion we just had on the calculus of total and covariant derivatives, shows that any time we have an absolutely continuous vector field (μ_t), the vector field $(\mathrm{P}_{\mu_t}(u_t))$ is absolutely continuous as well. This says that in a certain sense the projection operator varies smoothly in time along a regular curve. Here we want to study this question more in detail.

The main result of this section, shortly said, is that along (the restriction of) a geodesic in $\mathcal{P}_c(M)$, the operators $\mathrm{P}_{\mu_t}(\cdot)$, $\mathcal{O}_{v_t}(\cdot)$ and $\mathcal{O}^*_{v_t}(\cdot)$ are C^∞ in the sense that for any $n \in \mathbb{N}$, and any vector field (u_t) of class C^n, the vector fields $(\mathrm{P}_{\mu_t}(u_t))$, $(\mathcal{O}_{v_t}(u_t))$ and $(\mathcal{O}^*_{v_t}(u_t))$ are C^n as well.

As for the study of vector fields, the first thing to do is to give a definition of regularity for a family of operators defined along a regular curve. To this aim, recall that if $A_t : H \to H$ is a given curve of operators on a certain Hilbert space H, we have a natural definition of continuity and absolutely continuity: namely, continuity w.r.t. the operator norm, and existence of a family (A'_t) satisfying $\int_0^1 \|A'_t\|_{\mathrm{op}} dt < \infty$ and

$$A_s - A_t = \int_t^s A'_r \, dr,$$

where the integral is the Bochner integral.

By analogy with the definition of regularity of a vector field we give the following:

Definition 5.24 (Regularity of time dependent operators) *Let (μ_t) be a regular curve and $\mathcal{A}_t : L^2_{\mu_t} \to L^2_{\mu_t}$, $t \in [0,1]$ a given family of operators. Define $A_t : L^2_{\mu_0} \to L^2_{\mu_0}$ as*

$$A_t(u) := \tau_t^0\Big(\mathcal{A}_t\big(\tau_0^t(u)\big)\Big), \qquad \forall u \in L^2_{\mu_0}.$$

We say that $(\mathcal{A}_t)$ is L^1 (or continuous, or absolutely continuous, or C^1 ...) if and only if so is (A_t).

If $(\mathcal{A}_t)$ is absolutely continuous, its derivative can be defined by

$$\frac{\boldsymbol{d}}{dt}\mathcal{A}_t(u) := \tau_0^t\Big(\frac{d}{dt}A_t\big(\tau_t^0(u)\big)\Big) = \lim_{h\to 0}\frac{\mathcal{A}_{t+h}(\tau_t^{t+h}(u)) - \mathcal{A}_t(u)}{h}, \qquad \forall u \in L^2_{\mu_t}, \tag{5.18}$$

and it is clear that $(\frac{\boldsymbol{d}}{dt}\mathcal{A}_t)$ is an L^1 family of operators.

The absolute continuity of a family of operators may be tested in the following weak way.

Proposition 5.25 (Check of absolute continuity for operators) *Let (μ_t) be a regular curve and $(\mathcal{A}_t)$ a family of operators along it. Then $(\mathcal{A}_t)$ is absolutely continuous and its derivative is the L^1 family $(\mathcal{B}_t)$ if and only if for any absolutely continuous vector field (u_t), the vector field $(\mathcal{A}_t(u_t))$ is absolutely continuous and it holds*

$$\frac{\boldsymbol{d}}{dt}\mathcal{A}_t(u_t) = \mathcal{B}_t(u_t) + \mathcal{A}_t\left(u_t\right), \qquad a.e.t \tag{5.19}$$

Proof The 'only if' part is obvious, so we turn to the 'if'. Define $A_t(u) := \tau_t^0\Big(\mathcal{A}_t\big(\tau_0^t(u)\big)\Big)$, $\forall u \in L^2_{\mu_0}$ and similarly (B_t). Now we have that for any absolutely continuous curve $(\overline{u}_t) \subset L^2_{\mu_0}$, the curve $(A_t(\overline{u}_t)) \subset L^2_{\mu_0}$ is absolutely continuous and it holds

$$\frac{d}{dt}A_t(x_t) = B_t(x_t) + A_t\left(\frac{d}{dt}x_t\right).$$

In particular, considering a constant curve $(\overline{u})$, we have that $(A_t(\overline{u}))$ is absolutely continuous and the above equation implies

$$A_s(\overline{u}) - A_t(\overline{u}) = \int_t^s B_r(\overline{u})dr,$$

thus

$$\|(A_s - A_t)(\overline{u})\|_{\mu_0} \leq \|\overline{u}\|_{\mu_0}\int_t^s \|B_r\|_{\mathrm{op}}dr, \qquad \forall \overline{u} \in L^2_{\mu_0}$$

and the thesis follows. □

Remark 5.26 It is clear that a family of operators is C^1 if and only if it is absolutely continuous and its derivative admits a continuous representative. Also, in this case equations (5.18), (5.19) hold for every t. We will use this remark in the following without explicit mention. ■

In the following proposition we will indicate the composition of two operators $\mathcal{A}, \mathcal{B}$ by $\mathcal{A}\cdot\mathcal{B}$.

Proposition 5.27 (Basic properties of the derivation of operators) *Let (μ_t) be a regular curve and $(\mathcal{A}_t), (\mathcal{B}_t)$ two absolutely continuous families of operators. Then for a.e. t it holds:*

i) $\dfrac{\boldsymbol{d}}{dt}(\mathcal{A}_t + \mathcal{B}_t) = \dfrac{\boldsymbol{d}}{dt}(\mathcal{A}_t) + \dfrac{\boldsymbol{d}}{dt}(\mathcal{B}_t)$,

ii) $\dfrac{\boldsymbol{d}}{dt}\big(\phi(t)\mathcal{A}_t\big) = \phi'(t)\mathcal{A}_t + \phi(t)\dfrac{\boldsymbol{d}}{dt}\mathcal{A}_t$, *for every* $\phi \in C^\infty([0,1])$,

iii) $\dfrac{\boldsymbol{d}}{dt}(\mathcal{A}_t \cdot \mathcal{B}_t) = \left(\dfrac{\boldsymbol{d}}{dt}\mathcal{A}_t\right)\cdot\mathcal{B}_t + \mathcal{A}_t\cdot\left(\dfrac{\boldsymbol{d}}{dt}\mathcal{B}_t\right)$,

iv) $\dfrac{\boldsymbol{d}}{dt}(\mathcal{A}_t)^{\mathrm{t}} = \left(\dfrac{\boldsymbol{d}}{dt}\mathcal{A}_t\right)^{\mathrm{t}}$, *where $\mathcal{A}^{\mathrm{t}}$ is the adjoint of the operator $\mathcal{A}$.*

If $(\mathcal{A}_t), (\mathcal{B}_t)$ are C^1 then the above equations are true for every $t \in [0,1]$.

Proof All the properties are immediate consequences of the definition of derivative. □

With this terminology, and thanks to propositions 5.18 and 5.25, we have that the projection operator $\mathrm{P}_{\mu_t}(\cdot)$ is absolutely continuous along a regular curve and that its derivative is given by

$$\frac{\boldsymbol{d}}{dt}\mathrm{P}_{\mu_t}(\cdot) = -\mathrm{P}_{\mu_t}\big(\mathcal{O}^*_{v_t}(\cdot)\big) - \mathcal{O}_{v_t}\left(\mathrm{P}_{\mu_t}(\cdot)\right).$$

Similarly, if the regular curve satisfies $\int_0^1 \mathrm{L}(\frac{\boldsymbol{d}}{dt}v_t)dt < +\infty$, from corollary 5.23 we have that the operators $\mathcal{O}_{v_t}(\cdot)$ and $\mathcal{O}^*_{v_t}(\cdot)$ are absolutely continuous and their derivatives are

$$\begin{aligned}
\frac{\boldsymbol{d}}{dt}\mathcal{O}_{v_t}(\cdot) &= \mathcal{O}_{\frac{\boldsymbol{d}}{dt}v_t}(\cdot) - \mathrm{P}^{\perp}_{\mu_t}\big(R(v_t,\cdot)v_t\big) - \mathcal{O}_{v_t}\left(\mathcal{O}_{v_t}(\cdot)\right) + \mathrm{P}_{\mu_t}\big(\mathcal{O}^*_{v_t}\left(\mathcal{O}_{v_t}(\cdot)\right)\big),\\
\frac{\boldsymbol{d}}{dt}\mathcal{O}^*_{v_t}(\cdot) &= \mathcal{O}^*_{\frac{\boldsymbol{d}}{dt}v_t}(\cdot) - R\big(v_t,\mathrm{P}^{\perp}_{\mu_t}(\cdot)\big)v_t - \mathcal{O}^*_{v_t}\left(\mathcal{O}^*_{v_t}(\cdot)\right) + \mathcal{O}^*_{v_t}\left(\mathcal{O}_{v_t}\left(\mathrm{P}_{\mu_t}(\cdot)\right)\right),
\end{aligned}$$

Lemma 5.28 *Let $(\mu_t) \subset \mathcal{P}_c(M)$ be a regular curve and assume that its velocity vector field (v_t) is absolutely continuous and satisfies $\sup_t \mathrm{L}(v_t) < \infty$. Then the operator $R(v_t,\cdot)v_t$ is absolutely continuous and its derivative is given by*

$$\frac{\boldsymbol{d}}{dt}R(v_t,\cdot)v_t = (\nabla_{v_t}R)(v_t,\cdot)v_t + R\Big(\frac{\boldsymbol{d}}{dt}v_t,\cdot\Big)v_t + R(v_t,\cdot)\frac{\boldsymbol{d}}{dt}v_t,$$

where $(\nabla_{v_t}R)(v_t,u_t)v_t \in L^2_{\mu_t}$ is defined as $((\nabla_{v_t}R)(v_t,u_t)v_t)(x) := (\nabla_{v_t(x)}R)(v_t(x),u_t(x))v_t(x)$.

More generally, if (v_t) is of class $C^{n,1}$ and $\sup_t \mathrm{L}\big(\frac{\boldsymbol{d}^i}{dt^i}v_t\big) < \infty$ for any $i = 0,\dots,n$, then the operator $R(v_t,\cdot)v_t$ is $C^{n,1}$ as well. In particular, if (μ_t) is the restriction of a geodesic, $R(v_t,\cdot)v_t$ is C^∞.

Proof Observe that

$$\|R(v_t,u_t)v_t\|_{\mu_t} \le \mathcal{C}(K)\|u_t\|_{\mu_t}\mathrm{S}^2(v_t) \le 2\mathcal{C}\|u_t\|_{\mu_t}\Big(D^2\mathrm{L}^2(v_t) + \|v_t\|^2_{\mu_t}\Big),$$

where $K := \cup_t \operatorname{supp}(\mu_t)$ and in the second inequality we used lemma 5.19. This inequality shows that for any $t \in [0,1]$ the operator $R(v_t,\cdot)v_t$ maps $L^2_{\mu_t}$ into $L^2_{\mu_t}$.

Now pick an absolutely continuous vector field (u_t) and use proposition 3.13 to get

$$\frac{\boldsymbol{d}}{dt}R(v_t,u_t)v_t = (\nabla_{v_t}R)(v_t,u_t)v_t + R\left(\frac{\boldsymbol{d}}{dt}v_t,u_t\right)v_t + R\left(v_t,\frac{\boldsymbol{d}}{dt}u_t\right)v_t + R(v_t,u_t)\frac{\boldsymbol{d}}{dt}v_t.$$

Therefore the only thing we need to show is that $\int_0^1 \|(\nabla_{v_t}R)(v_t,u_t)v_t\|_{\mu_t}dt < \infty$. This follows from the regularity of the manifold: indeed, since the curvature tensor is C^1, we know that it holds

$$|(\nabla_{\mathrm{v}_1}R)(\mathrm{v}_2,\mathrm{v}_3)\mathrm{v}_4| \le C'_x|\mathrm{v}_1||\mathrm{v}_2||\mathrm{v}_3||\mathrm{v}_4|, \qquad \forall x \in M, \mathrm{v_i} \in T_xM, i = 1,2,3,4,$$

for some constant C'_x depending continuously on x. Therefore there exists $C'(K) \in \mathbb{R}$ which bounds C'_x for all $x \in K$ and we have

$$\|(\nabla_{v_t}R)(v_t,u_t)v_t\|_{\mu_t} \le C'(K)\Big(\mathrm{S}(v_t)\Big)^3\|u_t\|_{\mu_t}.$$

the conclusion follows from 5.25.

The general case follows analogously. □

From this lemma and the results of the previous section we get the following theorem.

Theorem 5.29 (Smoothness of operators) *Let $(\mu_t) \subset \mathcal{P}_c(M)$ be a regular curve. Then the operators $\mathrm{P}_{\mu_t}(\cdot)$ and $\mathrm{P}^{\perp}_{\mu_t}(\cdot)$ are absolutely continuous. Also, assume that its velocity vector field (v_t) is $C^{n,1}$ for some $n \in \mathbb{N}$ and that*

$$\int_0^1 \mathrm{L}\left(\frac{\boldsymbol{d}^i}{dt^i} v_t\right) dt < \infty, \qquad \forall i = 0, \dots, n+1.$$

Then the operators $\mathrm{P}_{\mu_t}(\cdot)$ and $\mathrm{P}^{\perp}_{\mu_t}(\cdot)$ are $C^{n+1,1}$ along (μ_t), and the operators $\mathcal{O}_{v_t}(\cdot)$ and $\mathcal{O}^*_{v_t}(\cdot)$ are $C^{n,1}$.

In particular, if (μ_t) is the restriction of a geodesic in $\mathcal{P}_c(M)$ defined on some larger interval $[-\varepsilon, 1+\varepsilon]$, $\varepsilon > 0$, then all these operators are C^∞.

Proof The first claim follows directly from propositions 5.18 and 5.25. The case $n = 0$ follows taking into account corollary 5.23. Also, since the derivative of $\mathrm{P}_{\mu_t}(\cdot)$ involves only the projection operators and the operators $\mathcal{O}_{v_t}(\cdot)$ and $\mathcal{O}^*_{v_t}(\cdot)$, the thesis follows if we prove the claim for the latter operators. To this aim, it is sufficient to work by induction: we do the explicit calculation for the case $n = 1$ and the operator $\mathcal{O}_{v_t}(\cdot)$. We know that in this case $\mathcal{O}_{v_t}(\cdot)$ is absolutely continuous and that its derivative is given by

$$\frac{\boldsymbol{d}}{dt}\mathcal{O}_{v_t}(\cdot) = \mathcal{O}_{\frac{\boldsymbol{d}}{dt}v_t}(\cdot) - \mathrm{P}^{\perp}_{\mu_t}\big(R(v_t,\cdot)v_t\big) - \mathcal{O}_{v_t}\left(\mathcal{O}_{v_t}(\cdot)\right) + \mathrm{P}_{\mu_t}\big(\mathcal{O}^*_{v_t}\left(\mathcal{O}_{v_t}(\cdot)\right)\big),$$

The last two terms of the above expression are absolutely continuous by inductive hypothesis, the term involving the curvature is absolutely continuous by lemma 5.28, thus we only need to prove that $t \mapsto \mathcal{O}_{\frac{\boldsymbol{d}}{dt}v_t}(\cdot) = \mathcal{N}_{\mu_t}(\frac{\boldsymbol{d}}{dt}v_t, \cdot)$ is absolutely continuous as well. This fact follows from theorem 5.21 and hypothesis $\int_0^1 \mathrm{L}(\frac{\boldsymbol{d}^2}{dt^2}v_t)dt < \infty$. The conclusion follows. □

Remark 5.30 (The regularity of M) The fact that M is a smooth manifold came into play for the first time in lemma 5.28, where we used the fact that the curvature tensor is C^∞. Prior to that, we only used that M was a C^2 manifold, rather than C^∞, as we only needed the existence of the parallel transport along absolutely continuous curves on M to define the translation maps τ_t^s. Observe that even if M is only C^2, we can still speak about, say, C^∞ vector fields along a regular curve: definition 3.2 still makes perfectly sense. The regularity of the manifold does not come into play, because once we translate the vector fields onto a fixed L^2 space (and to do this we only need the manifold to be C^2), the question of regularity trivializes into regularity of a curve in an Hilbert space. Therefore all the results up to section 5.3 are still valid, also those concerning, e.g., the fact that the velocity vector field of a geodesic is C^∞.

If the reader is skeptical about the possibility of defining an highly regular vector field on a space with less regularity, observe that in a certain sense we already did such an operation: think to the C^∞ vector field $(\tau_0^t(u))$ along a generic regular curve (μ_t). For a generic regular curve, we don't have any kind of time-dependent regularity of the vector field, so that we cannot say, for instance, that this curve is C^1. Still, it make perfectly sense to speak about C^∞ vector fields along it.

The regularity of the manifold comes into play only when checking the regularity of the operators like $\mathrm{P}_{\mu_t}(\cdot)$.

Also, let us mention that if M is analytic, that the operator $\mathrm{P}_{\mu_t}(\cdot)$ (and similarly $\mathcal{O}_{v_t}(\cdot)$, $\mathcal{O}^*_{v_t}(\cdot)$) is analytic as well along (the restriction of) a geodesic. This means that for any $t_0 \in [0,1]$, the family of operators $\overline{P}_t : L^2_{\mu_{t_0}} \to L^2_{\mu_{t_0}}$ defined by

$$\overline{P}_t(u) := \tau_t^{t_0}\Big(\mathrm{P}_{\mu_{t_0}}\big(\tau_{t_0}^t(u)\big)\Big),$$

admits the power series expansion around t_0:

$$\overline{P}_t = \sum_{i\geq 0} \frac{(t-t_0)^i}{i!}\frac{d^i}{dt^i}\overline{P}_t|_{t=t_0}. \tag{5.20}$$

Indeed, we already know that $(\mathrm{P}_{\mu_t}(\cdot))$ is C^∞, so that $\overline{P}_t$ is C^∞ as well. Thus, the only thing we need to check is the summability of the series. Now, along the restriction of a geodesic, the Lipschitz constant of the vectors v_t is uniformly bounded by some constant, say L. Thus the norm of the first derivative of $\overline{P}_t$ is bounded, by equation (5.9a), by $2L$. Arguing inductively and using equations (5.17), it is immediate to check that in the n-th derivative of $(\mathrm{P}_{\mu_t}(\cdot))$ they appear up to n consecutive applications of the operators $\mathcal{O}_{v_t}(\cdot)$ and $\mathcal{O}^*_{v_t}(\cdot)$ (plus possibly various projection onto the tangent/normal space, which do not increase the norm), and the derivatives of the operator $R(v_t,\cdot)v_t$ up to the order $n-2$. Since the norm of a sequence of n operators of the kind $\mathcal{O}_{v_t}(\cdot)$, $\mathcal{O}^*_{v_t}(\cdot)$ is bounded by L^n, in order to have the summability of the series in (5.20), we only need to check the summability of the norm of the derivatives of $R(v_t,\cdot)v_t$. This is a consequence of the fact that M is analytic. ■

The theorem just proved has an interesting consequence. Consider a tangent vector field (u_t) along a certain regular curve (μ_t) and consider the maps $t \mapsto \tau_t^0(u_t)$ and $t \mapsto \mathcal{T}_t^0(u_t)$. Both of these maps take value in $L^2_{\mu_0}$. Also, it is not hard to check that from proposition 4.11 it follows that one of them is absolutely continuous if and only if the other is. In general, equivalence between higher order regularity is not true: in order to get it, we need to make some assumptions on the velocity vector field of (μ_t), like in the previous theorem.

Proposition 5.31 *Let $(\mu_t) \subset \mathcal{P}_c(M)$ be a regular curve and assume that its velocity vector field (v_t) is $C^{n,1}$ for some $n \in \mathbb{N}$ and that*

$$\int_0^1 \mathrm{L}\left(\frac{\boldsymbol{d}^i}{dt^i}v_t\right)dt < \infty, \qquad \forall i = 0,\ldots,n+1.$$

Let (u_t) be a tangent vector field along it. Then the map $t \mapsto \tau_t^0(u_t)$ is $C^{n+1,1}$ (i.e. (u_t) is $C^{n+1,1}$) if and only of $t \mapsto \mathcal{T}_t^0(u_t)$ is $C^{n+1,1}$. Equivalently, (u_t) is $C^{n,1}$ if and only if it admits covariant derivatives up to order n and the n-th covariant derivative is absolutely continuous.

Proof Let $n = 0$. Recall that it holds

$$\frac{\boldsymbol{d}}{dt}u_t = \frac{\boldsymbol{D}}{dt}u_t + \mathcal{N}_{\mu_t}(u_t, v_t). \tag{5.21}$$

From the hypothesis on (v_t) we know that $(\mathcal{N}_{\mu_t}(u_t, v_t))$ is absolutely continuous. Thus $(\frac{\boldsymbol{d}}{dt}u_t)$ is absolutely continuous if and only if $(\frac{\boldsymbol{D}}{dt}u_t)$ is. Which is the claim.

The general case follows by induction. We run the explicit calculation for the case $n = 1$. Derivate (5.21) to get

$$\frac{\boldsymbol{d}^2}{dt^2}u_t = \frac{\boldsymbol{d}}{dt}\frac{\boldsymbol{D}}{dt}u_t + \frac{\boldsymbol{d}}{dt}\mathcal{N}_{\mu_t}(u_t, v_t) = \frac{\boldsymbol{D}^2}{dt^2}u_t + \mathcal{N}_{\mu_t}\left(\frac{\boldsymbol{D}}{dt}u_t, v_t\right) + \frac{\boldsymbol{d}}{dt}\mathcal{N}_{\mu_t}(u_t, v_t) \tag{5.22}$$

By the hypothesis on (v_t) and theorem 5.29 the vector field $(\frac{\boldsymbol{d}}{dt}\mathcal{N}_{\mu_t}(u_t, v_t))$ is absolutely continuous. Now assume that $t \mapsto \tau_t^0(u_t)$ is $C^{2,1}$. In particular $(\frac{\boldsymbol{d}}{dt}u_t)$ is absolutely continuous and from the result for the case $n = 0$ we have that $(\frac{\boldsymbol{D}}{dt}u_t)$ is absolutely continuous. By theorem 5.21 we get that $(\mathcal{N}_{\mu_t}(\frac{\boldsymbol{D}}{dt}u_t, v_t))$ is absolutely continuous as well. Thus from equation (5.22) we deduce that $(\frac{\boldsymbol{D}^2}{dt^2}u_t)$ is absolutely continuous.

Arguing analogously and assuming that $(\frac{\boldsymbol{D}^2}{dt^2}u_t)$ is absolutely continuous, we can deduce that $(\frac{\boldsymbol{d}^2}{dt^2}u_t)$ is absolutely continuous. $\square$

6 Curvature

6.1 The curvature tensor

In this section we study the curvature operator in $(\mathscr{P}_2(\mathbb{R}^d), W_2)$ together with its domain of definition. This operator was already introduced in [14] and, in the case $M = \mathbb{R}^d$, in [1] and [10], but in both cases there was no analysis on minimal regularity requirements (Lott considered only smooth positive measures on a compact Riemannian manifold and smooth vector fields, while the author and Ambrosio dropped the smoothness assumption on the measures, but kept the one on vector fields).

Trying to find the minimal regularity requirements is not just an academic exercise: as we will see in Chapter 8, the results proven here allow a precise discussion on existence of Jacobi fields on $(\mathcal{P}_c(M), W_2)$.

Following the analogy with the Riemannian case, we are lead to define the curvature tensor in the following way: given three vector fields $\mu \mapsto \nabla\varphi_\mu^i \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$, $i = 1, \ldots, 3$, the curvature tensor $\mathbf{R}$ calculated on them at the measure μ is defined as:

$$\mathbf{R}(\nabla\varphi_\mu^1, \nabla\varphi_\mu^2)(\nabla\varphi_\mu^3) := \boldsymbol{\nabla}_{\nabla\varphi_\mu^2}(\boldsymbol{\nabla}_{\nabla\varphi_\mu^1}\nabla\varphi_\mu^3) - \boldsymbol{\nabla}_{\nabla\varphi_\mu^1}(\boldsymbol{\nabla}_{\nabla\varphi_\mu^2}\nabla\varphi_\mu^3) + \boldsymbol{\nabla}_{[\nabla\varphi_\mu^1, \nabla\varphi_\mu^2]}\nabla\varphi_\mu^3,$$

where the objects like $\boldsymbol{\nabla}_{\nabla\varphi_\mu}(\nabla\psi_\mu)$, are, heuristically speaking, the covariant derivative of the vector field $\mu \mapsto \nabla\psi_\mu$ along the vector field $\mu \mapsto \nabla\varphi_\mu$.

However, in order to give a precise meaning to the above formula, we should be sure, at least, that the derivatives we are taking exist. Such an approach is possible, but heavy: indeed, consider that we should define what are C^1 and C^2 vector field, and in doing so we cannot just consider derivatives along curves. Indeed we would need to be sure that 'the partial derivatives have the right symmetries', otherwise there won't be those cancellations which let the above operator be a tensor.

Instead, we adopt the following strategy:

- First we calculate the curvature tensor for some very specific kind of vector fields, for which we are able to do and justify the calculations. Specifically, we will consider vector fields of the kind $\mu \mapsto \nabla\varphi$, where the function $\varphi \in C_c^\infty(M)$ does not depend on the measure μ (the calculations we will do are basically the same done by Lott in [14]).
- Then we prove that the object found is actually a tensor, i.e. that its value depends only on the $\mu-$a.e. value of the considered vector fields, and not on the fact that we obtained the formula assuming that the functions φ's were independent on the measure.
- Finally, we discuss the minimal regularity requirements for the object found to be well defined.

We will use the following proposition.

Proposition 6.1 *Let $\mu \in \mathcal{P}_c(M)$ and $\varphi \in C_c^\infty(M)$. Then for $T > 0$ sufficiently small the curve $t \mapsto (\exp(t\nabla\varphi))_\#\mu$ is a regular geodesic on $[0,T]$ whose velocity vector field (v_t) satisfies $\sup_{t\in[0,T]} \mathrm{L}(v_t) < \infty$.*

Proof From theorem 13.5 of [25] we know that for $T > 0$ sufficiently small the map $\exp(T\nabla\varphi)$ is d^2-cyclically monotone; also, up to taking T a bit smaller, we can assume that $\exp(t\nabla\varphi)$ is invertible with C^∞ inverse for any $t \in [0,T]$. In particular, (μ_t) is a geodesic on $[0,T]$. To get the second part of the statement observe that $\tau_0^t(\nabla\varphi)$ is the velocity vector field of (μ_t) on $[0,T]$ (τ_0^t being the natural translation maps along the curves $r \mapsto \exp_x(r\nabla\varphi)$). Since we are in a smooth setting, it is easy to check that from $\mathrm{L}(\nabla\varphi) < \infty$ it follows $\mathrm{L}(\tau_0^t(\nabla\varphi)) < \infty$ for any $t \in [0,T]$. □

The covariant derivative of a vector field of the kind $(\nabla\psi)$ along a curve of the kind $t \mapsto (\exp(t\nabla\varphi))_\#\mu$ is clearly continuous, and its value in 0 is given by $\mathrm{P}_\mu(\nabla^2\psi \cdot \nabla\varphi)$. Thus we write:

$$\boldsymbol{\nabla}_{\nabla\varphi}\nabla\psi := \mathrm{P}_\mu(\nabla^2\psi \cdot \nabla\varphi) \qquad \forall \varphi, \psi \in C_c^\infty(M). \tag{6.1}$$

Proposition 6.2 *Let $\mu \in \mathcal{P}_c(M)$ and $\varphi_1, \varphi_2, \varphi_3 \in C_c^\infty(M)$. The curvature tensor in μ calculated for the 3 vector fields $\nabla\varphi_i$, $i = 1,2,3$ is given by*

$$\begin{aligned}\mathbf{R}(\nabla\varphi_1, \nabla\varphi_2)\nabla\varphi_3 =& \mathrm{P}_\mu\Big(R(\nabla\varphi_1, \nabla\varphi_2)\nabla\varphi_3 + \nabla^2\varphi_2 \cdot \mathcal{N}_\mu(\nabla\varphi_1, \nabla\varphi_3) \\ &- \nabla^2\varphi_1 \cdot \mathcal{N}_\mu(\nabla\varphi_2, \nabla\varphi_3) + 2\nabla^2\varphi_3 \cdot \mathcal{N}_\mu(\nabla\varphi_1, \nabla\varphi_2)\Big).\end{aligned} \tag{6.2}$$

Proof Equation (6.2) is equivalent to

$$\begin{aligned}\langle \mathbf{R}(\nabla\varphi_1, \nabla\varphi_2)\nabla\varphi_3, \nabla\varphi_4\rangle_\mu =& \langle R(\nabla\varphi_1, \nabla\varphi_2)\nabla\varphi_3, \nabla\varphi_4\rangle_\mu \\ &+ \langle \mathcal{N}_\mu(\nabla\varphi_1, \nabla\varphi_3), \mathcal{N}_\mu(\nabla\varphi_2, \nabla\varphi_4)\rangle_\mu \\ &- \langle \mathcal{N}_\mu(\nabla\varphi_2, \nabla\varphi_3), \mathcal{N}_\mu(\nabla\varphi_1, \nabla\varphi_4)\rangle_\mu \\ &+ 2\,\langle \mathcal{N}_\mu(\nabla\varphi_1, \nabla\varphi_2), \mathcal{N}_\mu(\nabla\varphi_3, \nabla\varphi_4)\rangle_\mu,\end{aligned} \tag{6.3}$$

for any $\varphi_4 \in C_c^\infty(M)$.

Define $\mu_t := \exp(t\nabla\varphi_2)_\#\mu$ and $F(\nu) := \int \eta d\nu$ with $\eta := \left\langle \nabla^2\varphi_3 \cdot \nabla\varphi_1, \nabla\varphi_4 \right\rangle$. Evaluate the derivative at $t=0$ of $F(\mu_t)$ to get

$$\frac{d}{dt}F(\mu_t)_{|t=0} = \frac{d}{dt}\int \eta \circ \exp(t\nabla\varphi_2) d\mu_{|t=0} = \langle \nabla\eta, \nabla\varphi_2 \rangle_\mu .$$

On the other hand, using equations (6.1) and the fact that (μ_t) is a regular geodesic on $[0,T]$ for some $T>0$ (proposition 6.1), we have

$$\begin{aligned}\frac{d}{dt}F(\mu_t)_{|t=0} &= \frac{d}{dt}\left\langle \nabla^2\varphi_3 \cdot \nabla\varphi_1, \nabla\varphi_4 \right\rangle_{\mu_t} {}_{|t=0} \\ &= \frac{d}{dt}\left\langle \boldsymbol{\nabla}_{\nabla\varphi_1}\nabla\varphi_3(\mu_t), \nabla\varphi_4 \right\rangle_{\mu_t} {}_{|t=0} \\ &= \left\langle \boldsymbol{\nabla}_{\nabla\varphi_2}(\boldsymbol{\nabla}_{\nabla\varphi_1}\nabla\varphi_3), \nabla\varphi_4 \right\rangle_\mu + \left\langle \boldsymbol{\nabla}_{\nabla\varphi_1}\nabla\varphi_3, \boldsymbol{\nabla}_{\nabla\varphi_2}\nabla\varphi_4 \right\rangle_\mu .\end{aligned}$$

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

$$\begin{aligned}\left\langle \boldsymbol{\nabla}_{\nabla\varphi_2}(\boldsymbol{\nabla}_{\nabla\varphi_1}\nabla\varphi_3), \nabla\varphi_4 \right\rangle_\mu &= \left\langle \nabla\Big(\left\langle \nabla^2\varphi_3 \cdot \nabla\varphi_1, \nabla\varphi_4 \right\rangle \Big), \nabla\varphi_2 \right\rangle_\mu - \left\langle \boldsymbol{\nabla}_{\nabla\varphi_1}\nabla\varphi_3, \boldsymbol{\nabla}_{\nabla\varphi_2}\nabla\varphi_4 \right\rangle_\mu \\ &= \left\langle \nabla\Big(\left\langle \nabla^2\varphi_3 \cdot \nabla\varphi_1, \nabla\varphi_4 \right\rangle \Big), \nabla\varphi_2 \right\rangle_\mu - \left\langle \nabla^2\varphi_3 \cdot \nabla\varphi_1, \nabla^2\varphi_4 \cdot \nabla\varphi_2 \right\rangle_\mu \\ &\quad + \left\langle \mathcal{N}_\mu(\nabla\varphi_3, \nabla\varphi_1), \mathcal{N}_\mu(\nabla\varphi_4, \nabla\varphi_2) \right\rangle_\mu .\end{aligned}$$

The computation of the gradient of $\langle \nabla^2\varphi_3 \cdot \nabla\varphi_1, \nabla\varphi_4 \rangle$ gives

$$\begin{aligned}\left\langle \boldsymbol{\nabla}_{\nabla\varphi_2}(\boldsymbol{\nabla}_{\nabla\varphi_1}\nabla\varphi_3), \nabla\varphi_4 \right\rangle_\mu =& \left\langle \nabla^3\varphi_3(\nabla\varphi_1, \nabla\varphi_2), \nabla\varphi_4 \right\rangle_\mu + \left\langle \nabla^2\varphi_3 \cdot \nabla\varphi_4, \nabla^2\varphi_1 \cdot \nabla\varphi_2 \right\rangle_\mu \\ &+ \left\langle \mathcal{N}_\mu(\nabla\varphi_3, \nabla\varphi_1), \mathcal{N}_\mu(\nabla\varphi_4, \nabla\varphi_2) \right\rangle_\mu .\end{aligned} \tag{6.4}$$

Analogously, it holds:

$$\begin{aligned}\left\langle \boldsymbol{\nabla}_{\nabla\varphi_1}(\boldsymbol{\nabla}_{\nabla\varphi_2}\nabla\varphi_3), \nabla\varphi_4 \right\rangle_\mu =& \left\langle \nabla^3\varphi_3(\nabla\varphi_2, \nabla\varphi_1), \nabla\varphi_4 \right\rangle_\mu + \left\langle \nabla^2\varphi_3 \cdot \nabla\varphi_4, \nabla^2\varphi_2 \cdot \nabla\varphi_1 \right\rangle_\mu \\ &+ \left\langle \mathcal{N}_\mu(\nabla\varphi_3, \nabla\varphi_2), \mathcal{N}_\mu(\nabla\varphi_4, \nabla\varphi_1) \right\rangle_\mu ,\end{aligned} \tag{6.5}$$

so that, subtracting (6.5) from (6.4), from the formula (1.5) with $\xi^i = \nabla\varphi_i$, $i = 1,2,3$ we get

$$\begin{aligned}&\left\langle \boldsymbol{\nabla}_{\nabla\varphi_2}(\boldsymbol{\nabla}_{\nabla\varphi_1}\nabla\varphi_3), \nabla\varphi_4 \right\rangle_\mu - \left\langle \boldsymbol{\nabla}_{\nabla\varphi_1}(\boldsymbol{\nabla}_{\nabla\varphi_2}\nabla\varphi_3), \nabla\varphi_4 \right\rangle_\mu \\ &= \langle R(\nabla\varphi_1, \nabla\varphi_2)\nabla\varphi_3, \nabla\varphi_4 \rangle_\mu + \left\langle \nabla^2\varphi_3 \cdot \nabla\varphi_4, \nabla^2\varphi_1 \cdot \nabla\varphi_2 \right\rangle_\mu - \left\langle \nabla^2\varphi_3 \cdot \nabla\varphi_4, \nabla^2\varphi_2 \cdot \nabla\varphi_1 \right\rangle_\mu \\ &\quad + \left\langle \mathcal{N}_\mu(\nabla\varphi_3, \nabla\varphi_1), \mathcal{N}_\mu(\nabla\varphi_4, \nabla\varphi_2) \right\rangle_\mu - \left\langle \mathcal{N}_\mu(\nabla\varphi_3, \nabla\varphi_2), \mathcal{N}_\mu(\nabla\varphi_4, \nabla\varphi_1) \right\rangle_\mu .\end{aligned} \tag{6.6}$$

Recalling the torsion free identity (5.1) we get

$$\begin{aligned}
\left\langle \boldsymbol{\nabla}_{[\nabla\varphi_1,\nabla\varphi_2]}\nabla\varphi_3, \nabla\varphi_4\right\rangle_\mu &= \left\langle \nabla^2\varphi_3\cdot P_\mu\Big(\nabla^2\varphi_2\cdot\nabla\varphi_1 - \nabla^2\varphi_1\cdot\nabla\varphi_2\Big), \nabla\varphi_4\right\rangle_\mu \\
&= \left\langle P_\mu\Big(\nabla^2\varphi_2\cdot\nabla\varphi_1 - \nabla^2\varphi_1\cdot\nabla\varphi_2\Big), \nabla^2\varphi_3\cdot\nabla\varphi_4\right\rangle_\mu \\
&= \left\langle \nabla^2\varphi_3\cdot\nabla\varphi_4, \nabla^2\varphi_2\cdot\nabla\varphi_1 - \nabla^2\varphi_1\cdot\nabla\varphi_2\right\rangle_\mu \\
&\quad - \left\langle \mathcal{N}_\mu(\nabla\varphi_2,\nabla\varphi_1), \mathcal{N}_\mu(\nabla\varphi_3,\nabla\varphi_4)\right\rangle_\mu \\
&\quad + \left\langle \mathcal{N}_\mu(\nabla\varphi_1,\nabla\varphi_2), \mathcal{N}_\mu(\nabla\varphi_3,\nabla\varphi_4)\right\rangle_\mu .
\end{aligned}$$

Adding this equation to (6.6) we get the thesis. □

In the following, mostly in order to highlight the symmetries of $\mathbf{R}$, we will think the curvature tensor as a map which takes four vector fields and gives a real number, rather than a map which takes 3 vector fields and gives a vector field. In practice, we will mostly use equation (6.3), rather than equation (6.2).

Proposition 6.3 *Let $\mu \in \mathcal{P}_c(M)$. The curvature operator, given by formula* (6.3)*, is a tensor on $[\mathcal{V}(M)]^4 \subset [\mathrm{L}^2_\mu]^4$, i.e. its value depends only on the $\mu-$a.e. value of the 4 vector fields.*

Proof Clearly the left hand side of equation (6.3) is a tensor w.r.t. the fourth coordinate. The conclusion follows from the symmetries of the right hand side. □

Let us now observe that by arguments analogous to those used to describe the tensor $\mathcal{N}_\mu$, it possible to show that the set of 4−ples of vector fields in $\mathrm{Tan}_\mu(\mathscr{P}_2(M))$ such that at least 3 of them are Lipschitz, may be endowed with a topology such that $n \mapsto (v_n^1, v_n^2, v_n^3, v_n^4)$ is converging if and only if it is a Cauchy sequence in $[\mathrm{Tan}_\mu(\mathscr{P}_2(M))]^4$ and for every subsequence, not relabeled, there exists a further extraction $k \mapsto (v_{n_k}^1, v_{n_k}^2, v_{n_k}^3, v_{n_k}^4)$ such that

$$\sup_k \mathrm{L}(v_{n_k}^i) < \infty \text{ for at least 3 indexes } i \in \{1,2,3,4\}.$$

Proposition 6.4 *Let $\mu \in \mathcal{P}_c(M)$. The curvature tensor is well defined and sequentially continuous on the set of 4−ples of vector field in $\mathrm{Tan}_\mu(\mathscr{P}_2(M))$ such that at least 3 of them are Lipschitz, endowed with the topology described above. For a given 4-ple $v_1, v_2, v_3, v_4 \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$ it is given by*

$$\begin{aligned}
\langle \mathbf{R}(v_1,v_2)v_3, v_4\rangle_\mu = \langle R(v_1,v_2)v_3, v_4\rangle_\mu &+ \langle \mathcal{N}_\mu(v_1,v_3), \mathcal{N}_\mu(v_2,v_4)\rangle_\mu \\
&- \langle \mathcal{N}_\mu(v_2,v_3), \mathcal{N}_\mu(v_1,v_4)\rangle_\mu + 2\langle \mathcal{N}_\mu(v_1,v_2), \mathcal{N}_\mu(v_3,v_4)\rangle_\mu
\end{aligned} \tag{6.7}$$

Proof Is a straightforward consequence of the properties of $\mathcal{N}_\mu$, we omit the details. □

Remark 6.5 (On the compactness assumption) The hypothesis $\mu \in \mathcal{P}_c(M)$ plays the same role played in theorem 5.21: it deserves to be sure that $R(v_1,v_2)v_3$ belongs to L^2_μ. When $M = \mathbb{R}^d$, or more generally is a flat manifold, this hypothesis can be dropped. ■

It is important to underline that the fact that the curvature tensor extends continuously to the space described above, does *not* mean that its value can be recovered directly from the definition by taking double derivatives and running the calculations: those calculations are in general not justified under the loose regularity assumptions on the vectors we are making now. What we are saying here, is just that there is a well defined notion of $\langle \mathbf{R}(v_1, v_2)v_3, v_4\rangle_\mu$ as soon as 3 of these vectors are Lipschitz. If we want to recover the usual links between the curvature tensor and the geometry of the space we are still in need to be careful about the regularity assumptions we make.

Proposition 6.6 (Symmetries of R) *Let $\mu \in \mathcal{P}_c(M)$. Then the curvature tensor $\mathbf{R}$ has the following symmetries:*

$$\langle \mathbf{R}(v_1, v_2)v_3, v_4\rangle_\mu = -\langle \mathbf{R}(v_2, v_1)v_3, v_4\rangle_\mu = -\langle \mathbf{R}(v_1, v_2)v_4, v_3\rangle_\mu = \langle \mathbf{R}(v_3, v_4)v_1, v_2\rangle_\mu ,$$
$$\langle \mathbf{R}(v_1, v_2)v_3, v_4\rangle_\mu + \langle \mathbf{R}(v_3, v_1)v_2, v_4\rangle_\mu + \langle \mathbf{R}(v_2, v_3)v_1, v_4\rangle_\mu = 0$$

Proof All the equations are direct consequence of the analogous equations valid for R and of the antisymmetry of $\mathcal{N}_\mu$. □

It possible, making appropriate smoothness assumptions, to prove that the curvature tensor satisfies also the second Bianchi identity, and not just the first one. However, the calculations would be really cumbersome, and we believe at this stage it is not worth to investigate in this direction.

Remark 6.7 (Curvature tensor and of local geometry: handle with care) It is natural to expect that the curvature tensor $\mathbf{R}$ on $(\mathcal{P}_c(M), W_2)$ describes the local geometry like the curvature tensor R does on M. However, this is *not* always the case. To make this point as clear as possible, think at the case $M = \mathbb{R}^d$: we know that even if the origin space is flat, the space $\mathscr{P}_2(\mathbb{R}^d)$ is actually curved. Also, from many points of view, measures with a singular part, and in particular finite combinations of δ's, are measures where the Riemannian structure of $\mathscr{P}_2(\mathbb{R}^d)$ degenerates (e.g. the tangent space is not an Hilbert space - see appendix of [2] and chapter 4 of [10]). Therefore, it would be natural to expect that the curvature tensor degenerates at such measures.

However, this is not the case. Consider for instance a measure $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ with finite support. In this case $\mathrm{Tan}_\mu(\mathscr{P}_2(M)) = L^2_\mu$, and therefore the tensor $\mathcal{N}_\mu$ is identically 0. Furthermore, since $M = \mathbb{R}^d$, we have $R = 0$ and thus we obtain $\mathbf{R} = 0$ at μ.

Also, if the underlying manifold is not $\mathbb{R}^d$, it is still true that $\mathcal{N}_\mu$ is identically 0 on measures with finite support: this means that the curvature tensor $\mathbf{R}$ is actually a bounded operator defined on the whole $[L^2_\mu]^4$. On the other side, Sturm proved in [24] with an example based on geometrical arguments, that when the base space is not flat, the sectional curvature degenerates both to $+\infty$ and $-\infty$ at these measures. So how is it possible that both the arguments are correct? The point is that Sturm proved that there is a sequence of measures μ^n with finite support along which the sectional curvature (Alexandrov curvature, in his setting) becomes unbounded from both above and below near this sequence. But nothing excludes that at each of these measures the curvature is bounded. For better clarity, we explicitate the calculus of the sectional curvature for this kind of measures. Let $\mu = \sum_{i=1}^n a_i \delta_{x_i}$, $a_i > 0$ for every i and $\sum_i a_i = 1$. Then we have

$$| \langle \mathbf{R}(v,w)w, v \rangle_\mu | = \left| \sum_{i=1}^n a_i \langle R(v_i, w_i)w_i, v_i \rangle \right| \le C \sum_{i=1}^n a_i \big(|v_i|^2 |w_i|^2 - \langle v_i, w_i \rangle^2 \big),$$

where C is the maximum value of the sectional curvature of M at the points x_i (so that equality can be reached), $v_i = v(x_i)$ and $w_i = w(x_1)$, $i = 1, \ldots, n$. Now, to ask if the sectional curvature is bounded at μ and who is its best bound, is the same as to ask who is the best constant $\mathbf{C}(\mu)$, if any, for which it holds

$$\begin{aligned} | \langle \mathbf{R}(v,w)w, v \rangle_\mu | &\le \mathbf{C}(\mu) \big(\|v\|_\mu^2 \|w\|_\mu^2 - \langle v, w \rangle_\mu^2 \big) \\ &= \mathbf{C}(\mu) \left(\sum_{i=1}^n a_i |v_i|^2 \sum_{i=1}^n a_i |w_i|^2 - \Big(\sum_{i=1}^n a_i \langle v_i, w_i \rangle \Big)^2 \right). \end{aligned}$$

From the finiteness of the problem, it is immediate to verify that some $\mathbf{C}(\mu)$ satisfying the above inequality always exists. However, this constant *depends* on the values a_i's, and actually it explodes if some of the a_i's goes to 0. This is precisely what happens in Sturm's example, thus the two approaches give the same conclusion.

But if the curvature is bounded at some fixed measure with finite support, why we don't see the degeneracy of the Riemannian structure? The point is the following: $\mathbf{R}$ is defined on a dense subset of $[\mathrm{Tan}_\mu(\mathscr{P}_2(M))]^4$, and therefore can reasonably describe the local geometry only for those measures μ for which the image of the exponential map $\mathrm{Tan}_\mu(\mathscr{P}_2(M)) \ni v \mapsto \mathbf{exp}_\mu(v) \in \mathscr{P}_2(M)$ contains a neighborhood of μ in $\mathscr{P}_2(M)$. It is not hard to see that this is the case if and only if the exponential map is surjective, and this is true if and only if for every $\nu \in \mathscr{P}_2(M)$ there exists an optimal transport map from μ to ν. It is known that all the measures which give 0 mass to all $\dim(M) - 1$ dimensional sets have this property. For a complete characterization of the measures for which the exponential map is surjective, see [12].

For instance, consider a measure μ with finite support on $\mathbb{R}^d$. It is easy to check that for $r > 0$ sufficiently small, the exponential map restricted to $B_r(0) \subset \mathrm{Tan}_\mu(\mathscr{P}_2(M))$ is an isometry. Since $\mathrm{Tan}_\mu(\mathscr{P}_2(M))$ is an Hilbert space (in this case it also has finite dimension), it is flat. Thus there is no surprise that the curvature tensor $\mathbf{R}$ is 0 at μ, since it is 'interested' only on those measures in the image of $B_r(0)$, which is a set isometric to a flat ball in $\mathrm{Tan}_\mu(\mathscr{P}_2(M))$. ■

6.2 Related notions of curvature

Even if the curvature tensor requires 3 vector fields to be Lipschitz to be well defined, other important notions of curvature make sense with less regularity assumptions. Consider for instance the sectional curvature: if we evaluate the formula of the curvature tensor for a 4-ple of vector fields of the form (v, w, w, v), the formula reduces to:

$$\int K(v,w)(|v|^2|w|^2 - \langle v, w \rangle^2) d\mu + 3 \|\mathcal{N}_\mu(v,w)\|_\mu^2,$$

where $K(v(x), w(x))$ is the sectional curvature of M at the point x along the plane generate by the vectors $v(x), w(x) \in T_xM$. This formula makes sense as soon as just one of the vectors is Lipschitz. Thus we give the following definition.

Definition 6.8 (Sectional curvature) *Let $\mu \in \mathcal{P}_c(\mathbb{R}^d)$ and $(v, w) \in \mathrm{L}\textsc{n}\mathrm{L}_\mu$ be two non proportional vector fields. The sectional curvature $\mathbf{K}(v, w)$ is:*

$$\mathbf{K}(v, w) := \frac{\langle \mathbf{R}(v, w)w, v\rangle_\mu}{\|v\|_\mu^2 \|w\|_\mu^2 - \langle v, w\rangle_\mu^2} = \frac{\displaystyle\int K(v, w)\Big(|v|^2|w|^2 - \langle v, w\rangle^2\Big)d\mu + 3\|\mathcal{N}_\mu(v, w)\|_\mu^2}{\|v\|_\mu^2 \|w\|_\mu^2 - \langle v, w\rangle_\mu^2} \tag{6.8}$$

As already noticed by Lott ([14]) (who found the formula (6.8) - below we briefly recall his arguments), the expression for the sectional curvature confirms that if the manifold M has nonnegative sectional curvature, the same is true for $\mathscr{P}_2(M)$, while in general other bounds on the sectional curvature are not inherited by $\mathscr{P}_2(M)$. Indeed, it is clear that from $K(v(x), w(x)) \geq 0$ for μ-a.e. x we deduce $\mathbf{K}(v, w) \geq 0$. On the other hand, from a bound of the kind $K(v(x), w(x)) \leq c$ we cannot deduce anything on $\mathbf{K}(v, w)$, because of the presence of the term $3\|\mathcal{N}_\mu(v, w)\|_\mu^2$. Finally, from $K(v(x), w(x)) \geq c > 0$, it is not possible, due to normalization issues, to deduce $\mathbf{K}(v, w) \leq c$, and if $K(v(x), w(x)) < 0$ for a μ-positive (and infinite) set of x's, then the same normalization issues lead to $\inf_{(v,w)\in \mathrm{L}\textsc{n}\mathrm{L}_\mu} \mathbf{K}(v, w) = -\infty$, in complete analogy with Sturm's example ([24]), as already discussed in remark 6.7.

Observe that the formula (6.8) is consistent with the O'Neill formula already present in [20].

To conclude, consider two vector fields $u, v \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$ and assume that $\mathrm{L}(v) < \infty$. Then there is a good definition of the vector field $\mathbf{R}(v, u)v \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$, as there are simplifications in the formula for the curvature tensor similar to those appearing in the definition of the sectional curvature. From formula (6.7) we can write it as:

$$\mathbf{R}(v, u)v := \mathrm{P}_\mu\Big(R(v, u)v + 3\mathcal{O}_v^*\left(\mathcal{O}_v\left(u\right)\right)\Big).$$

The operator $\mathrm{Tan}_\mu(\mathscr{P}_2(M)) \ni u \mapsto \mathbf{R}(v, u)v \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$ is continuous and its norm is bounded by $\mathcal{C}(\mathrm{supp}(\mu))\mathrm{S}^2(v)^+(3\mathrm{L}(v))^2$. When $v \in \mathcal{V}(M)$, the above formula may be written as

$$\mathbf{R}(v, u)v = \mathrm{P}_\mu\Big(R(v, u)v + 3\nabla v \cdot \mathrm{P}_\mu^\perp(\nabla v^{\mathrm{t}} \cdot u)\Big),$$

A direct consequence of theorem 5.29 is the smoothness of $\mathbf{R}(v_t, \cdot)v_t$ along a geodesic:

Proposition 6.9 ($\mathbf{R}(v_t, \cdot)v_t$ along a geodesic) *Let $(\mu_t) \subset \mathcal{P}_c(M)$ be the restriction to $[0, 1]$ of a geodesic defined in some larger interval $[-\varepsilon, 1 + \varepsilon]$ and (v_t) its velocity vector field. Then the operator $\mathbf{R}(v_t, \cdot)v_t$ is C^∞.*

7 Differentiability of the exponential map

7.1 Introduction to the problem

Recall that given a Riemannian manifold $\tilde{M}$, a way to define the differential of the exponential map is the following. We pick a smooth curve $t \mapsto (x_t, u_t) \in TM$ and define the curve $t \mapsto y_t :=$

$\exp_{x_t}(u_t)$. Then we say that the differential of the exponential map $\exp : TM \to M$ at the point (x_t, u_t) along the direction $(\dot{x}_t, \nabla_{\dot{x}_t} u_t)$ is the vector $\dot{y}_t \in T_{y_t}M$.

It is not hard to imitate this definition in the case of the 'manifold' $(\mathscr{P}_2(M), W_2)$. Let us start recalling that given $\mu \in \mathscr{P}_2(M)$ and $u \in L^2_\mu$, the exponential $\mathbf{exp}_\mu(u) \in \mathscr{P}_2(M)$ is defined as

$$\mathbf{exp}_\mu(u) := (\exp(u))_\# \mu. \tag{7.1}$$

This definition is justified by the fact that if φ is a smooth Kantorovich potential, then $[0,1] \ni t \mapsto \mathbf{exp}_\mu(t\nabla\varphi)$ is a constant speed geodesic. Since the formula (7.1) makes sense for general vectors $u \in L^2_\mu$, it is customary to keep the definition at this level of generality, and not, for instance, to restrict it to tangent vectors.

Now, we didn't define a smooth structure on TM, and we won't, but still we can try to do the following. Pick a regular curve (μ_t) and an absolutely continuous vector field (u_t) along it; then define the curve $t \mapsto \nu_t := \mathbf{exp}_{\mu_t}(u_t)$. Hopefully, the curve (ν_t) is absolutely continuous, and if this is the case we may consider its velocity vector field (w_t). Then, it is reasonable to affirm that the differential of the exponential map at the point (μ_t, u_t) along the direction $(v_t, \frac{\boldsymbol{d}}{dt} u_t)$ is given by $w_t \in \mathrm{Tan}_{\nu_t}(\mathscr{P}_2(M))$ for a.e. t (here (v_t) is the velocity vector field of (μ_t)).

We will see in a moment that this heuristic argument actually works under only minimal boundedness assumptions on (μ_t) and (u_t), but before turning to the technical details, let us take some time to understand 'who' is the velocity vector field (w_t) of (ν_t) in terms of the data of the problems, i.e. (μ_t) and (u_t).

We are going to be sloppy about regularity/integrability assumptions here, the rigorous result will be given in the next section (theorem 7.2). In order to check if (ν_t) is absolutely continuous and to identify its velocity vector field, let us check directly the validity of the continuity equation. As usual, let $\mathbf{T}(t,s,\cdot)$ be the flow maps of (μ_t) and (v_t) its velocity vector field. Fix $\varphi \in C^\infty_c(M)$ and consider the derivative $\frac{d}{dt}\int \varphi d\nu_t$. Since

$$\nu_t = \mathbf{exp}_{\mu_t}(u_t) = (\exp(u_t))_\# \mu_t = \big(\exp(u_t) \circ \mathbf{T}(0,t,\cdot)\big)_\# \mu_0,$$

we have

$$\begin{aligned}\frac{d}{dt}\int \varphi d\nu_t &= \frac{d}{dt}\int \varphi \circ \exp(u_t) \circ \mathbf{T}(0,t,\cdot) d\mu_0 \\ &= \int \Big\langle (\nabla\varphi) \circ \exp(u_t) \circ \mathbf{T}(0,t,\cdot), \frac{d}{dt}\Big(\exp(u_t) \circ \mathbf{T}(0,t,\cdot)\Big)\Big\rangle d\mu_0.\end{aligned}$$

To express the value of $\frac{d}{dt}\Big(\exp(u_t) \circ \mathbf{T}(0,t,\cdot)\Big)$ a bit more explicitly, let us introduce the following notation. Let $\mu \in \mathscr{P}_2(M)$, and $u, \tilde{u}^1, \tilde{u}^2 \in L^2_\mu$ and assume that $\exp(u) : \mathrm{supp}(M) \to M$ is μ-essentially invertible. Then we can define the vector field $j_{\mu,u}(\tilde{u}^1, \tilde{u}^2)$ as

$$j_{\mu,u}(\tilde{u}^1, \tilde{u}^2)(\exp_x(u(x))) := \begin{cases} \text{the value at } t = 1 \text{ of the Jacobi field } t \mapsto j_t \in T_{\exp_x(tu(x))}M \\ \text{along the geodesic } t \mapsto \exp_x(tu(x)) \\ \text{having initial conditions } j_0(x) = \tilde{u}^1(x),\ j'_0(x) = \tilde{u}^2(x), \end{cases}$$

the assumption on the invertibility of $\exp(u)$ guarantees that $j_{\mu,u}(\tilde{u}^1, \tilde{u}^2)$ is well defined $\mathbf{exp}_\mu(u)$-a.e.. Let us assume that $j_{\mu,u}(\tilde{u}^1, \tilde{u}^2)$ belongs to $L^2_{\mathbf{exp}_\mu(u)}$ (this will actually require some hypothesis on μ, u).

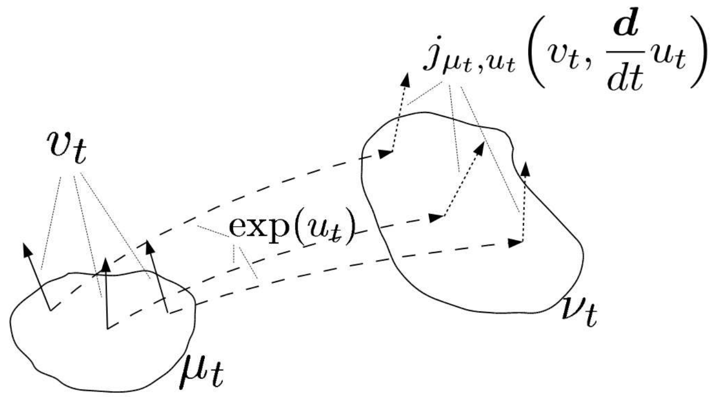


Figure 2:
(i) Along the curve (μ_t), each 'atom' x of the mass of μ_t is moving in the direction $v_t(x)$
(ii) In order to produce the distribution ν_t, each 'atom' x of μ_t is moved to $\exp_x(u_t(x))$
(iii) When t varies, the infinitesimal variation of $\exp_x(u_t(x))$ (which is the infinitesimal variation of mass along the curve (ν_t)) is given by the Jacobi field $j_{\mu_t,u_t}(v_t, \frac{\boldsymbol{d}}{dt}u_t)$, whose initial conditions depend on v_t and the variation of u_t
(iv) The vector field $j_{\mu_t,u_t}(v_t, \frac{\boldsymbol{d}}{dt}u_t)$ may be not optimal in terms of kinetic energy spent to produce the variation of (ν_t): the optimal one is found by taking the projection of $j_{\mu_t,u_t}(v_t, \frac{\boldsymbol{d}}{dt}u_t)$ onto the tangent space at ν_t.

With this notation and recalling that by proposition 3.13 we have that

$$\frac{d}{dt}\Big(u_t\big(\mathbf{T}(0,t,x)\big)\Big) = \Big(\frac{\boldsymbol{d}}{dt}u_t\Big)\big(\mathbf{T}(0,t,x)\big), \qquad \mu_0 - a.e.\ x,\ a.e.\ t,$$

it is easy to check that

$$\frac{d}{dt}\Big(\exp(u_t)\circ\mathbf{T}(0,t,\cdot)\Big) = j_{\mu_t,u_t}\Big(v_t, \frac{\boldsymbol{d}}{dt}u_t\Big)\Big(\exp(u_t)\circ\mathbf{T}(0,t,\cdot)\Big), \qquad a.e.\ t,$$

and therefore

$$\begin{aligned}\frac{d}{dt}\int \varphi d\nu_t &= \int \Big\langle (\nabla\varphi)\circ\exp(u_t)\circ\mathbf{T}(0,t,\cdot), \frac{d}{dt}\Big(\exp(u_t)\circ\mathbf{T}(0,t,\cdot)\Big)\Big\rangle d\mu_0 \\ &= \int \Big\langle \nabla\varphi, j_{\mu_t,u_t}\Big(v_t, \frac{\boldsymbol{d}}{dt}u_t\Big)\Big\rangle d\nu_t, \qquad \forall \varphi \in C^\infty_c(M).\end{aligned}$$

This means that (ν_t) is an absolutely continuous curve, and that an admissible choice of vector fields satisfying the continuity equation with (ν_t) is given by $\Big(j_{\mu_t,u_t}\Big(v_t, \frac{\boldsymbol{d}}{dt}u_t\Big)\Big)$. Therefore, by theorem 1.28, the velocity vector field (w_t) of (ν_t) (i.e. the tangent one), is given by

$$w_t = \mathrm{P}_{\nu_t}\left(j_{\mu_t,u_t}\Big(v_t, \frac{\boldsymbol{d}}{dt}u_t\Big)\right).$$

7.2 Rigorous result

In order to turn the discussion of the previous section into a rigorous result, the only thing we need to check is the fact that the vector field $(j_{\mu_t,u_t})(v_t, \frac{\boldsymbol{d}}{dt}u_t)$ belongs to $L^2_{\nu_t}$ and its norm is integrable. Also, it would be nice to drop the hypothesis on the invertibility of $\exp(u)$ in the definition of $j_{\mu,u}$, at least because is a condition difficult to check in practice.

Let us address the first issue. From inequality (1.8b) we get the bound

$$\|j_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\|^2_\nu \leq 2\|\tilde{u}^1\|^2_\mu \cosh^2\left(S\sqrt{\mathcal{C}}\right) + 2\frac{\|\tilde{u}^2\|^2_\mu}{S^2\mathcal{C}}\sinh^2\left(S\sqrt{\mathcal{C}}\right),$$

where $\nu := \mathbf{exp}_\mu(u)$, S is the essential supremum of u and $\mathcal{C}$ is a bound of the curvature of M on the set $\{\exp_x(t(u(x)))\}_{x\in\mathrm{supp}(\mu),t\in[0,1]}$. Therefore in general, in order to be sure that $\|j_{\mu,u}(\tilde{u}_1,\tilde{u}_2)\|_\nu < \infty$ we must impose that u is essentially bounded and that μ has compact support, so that $\mathcal{S} < \infty$ and the set $\{\exp_x(t(u(x)))\}_{x\in\mathrm{supp}(\mu),t\in[0,1]}$ is relatively compact. For general manifolds, we can't do better than this, however, if M has non negative sectional curvatures, then inequality (1.9) provides a much better bound on the norm of the Jacobi fields. Under this assumption it holds:

$$\|j_{\mu,u}(\tilde{u}_1,\tilde{u}_2)\|^2_\nu \leq 2\|\tilde{u}^1\|_\mu + 2\|\tilde{u}^2\|_\mu,$$

and therefore it always hold $\|j_{\mu,u}(\tilde{u}_1,\tilde{u}_2)\|^2_\nu < \infty$ (we will always assume $\tilde{u}^1,\tilde{u}^2 \in L^2_\mu$).

To drop the hypothesis of μ-essential invertibility of $\exp(u)$ we do the following: for any $y \in \mathrm{supp}(\nu)$ for which there is more than one $x \in \mathrm{supp}(\mu)$ such that $\exp(u(x)) = y$, we define the value of $j_{\mu,u}(\tilde{u}^1,\tilde{u}^2)$ as average of the values given by each of these x's, the average being taken w.r.t. the disintegration of μ with respect to the map $\exp(u)$.

The rigorous definition is the following.

Definition 7.1 *Let $\mu \in \mathscr{P}_2(M)$ and $u, \tilde{u}^1, \tilde{u}^2 \in L^2_\mu$. Assume that μ has compact support and u is essentially bounded. Define the map $J_{\mu,u}(\tilde{u}^1,\tilde{u}^2) : \mathrm{supp}(\mu) \to TM$ satisfying $J_{\mu,u}(\tilde{u}^1,\tilde{u}^2)(x) \in T_{\exp(u(x))}M$ by*[3]

$$J_{\mu,u}(\tilde{u}^1,\tilde{u}^2)(x) := \begin{cases} \text{the value at } t=1 \text{ of the Jacobi field } t\mapsto j_t \in T_{\exp_x(tu(x))}M \\ \text{along the geodesic } t\mapsto \exp_x(tu(x)) \\ \text{having initial conditions } j_0(x) = \tilde{u}^1(x),\ j_0'(x) = \tilde{u}^2(x), \end{cases}$$

and the plan $\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2) \in \mathscr{P}(TM)$ by

$$\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2) := \big(\exp(u), J_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\big),$$

[3]if $\exp(u)$ is μ-essentially invertible it holds $J_{\mu,u}(\tilde{u}^1,\tilde{u}^2) = j_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\circ\exp(u)$, with $j_{\mu,u(\tilde{u}^1,\tilde{u}^2)}$ defined as in the previous section.

(observe that for $\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2)$-a.e. (x,v) it holds $\mathrm{v}\in T_xM$). Then the vector field $j_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\in L^2_{\mathbf{exp}_\mu(u)}$ is defined as the barycentric projection of $\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2)$, i.e.:

$$j_{\mu,u}(\tilde{u}^1,\tilde{u}^2)(x):=\int \mathrm{v}d\big(\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\big)_x(\mathrm{v}),$$

where $\big(\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\big)_x$ is the disintegration of $\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2)$ w.r.t. the natural projection $\pi^M: TM\to M$.

If the manifold M has non negative sectional curvatures, we keep this definition also in the case $\mu\notin\mathcal{P}_c(M)$ and u not essentially bounded.

In order to prove that this is a good definition we need to show that $\big(\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\big)_x$ has finite first moment for $\mathbf{exp}_\mu(u)$-a.e. x and that $j_{\mu,u}(\tilde{u}^1,\tilde{u}^2)$ belongs to $L^2_{\mathbf{exp}_\mu(u)}$. Both these facts are obvious, since the bound (1.8b) gives

$$\|J_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\|^2_\mu\le 2\|\tilde{u}^1\|^2_\mu\cosh^2\left(S\sqrt{\mathcal{C}}\right)+2\frac{\|\tilde{u}^2\|^2_\mu}{S^2\mathcal{C}}\sinh^2\left(S\sqrt{\mathcal{C}}\right)<\infty,$$

by the hypothesis on μ,u, and therefore

$$\begin{aligned}\int\left(\int|\mathrm{v}|d\big(\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\big)_x\right)^2d\mathbf{exp}_\mu(u)&\le\int\int|\mathrm{v}|^2d\big(\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\big)_x d\mathbf{exp}_\mu(u)\\&=\int|\mathrm{v}|^2d\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2)=\|J_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\|^2_\mu<\infty,\end{aligned}$$

this shows that $j_{\mu,u}(\tilde{u}^1,\tilde{u}^2)$ is well defined $\mathbf{exp}_\mu(u)$-a.e., to show that it belongs to $L^2_{\mathbf{exp}_\mu(u)}(\tilde{u}^1,\tilde{u}^2)$ it is enough to observe

$$\begin{aligned}\|j_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\|^2_{\mathbf{exp}_\mu(u)}&=\int\left(\int\mathrm{v}d\big(\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\big)_x\right)^2\\&\le\int\left(\int|\mathrm{v}|d\big(\boldsymbol{\gamma}_{\mu,u}(\tilde{u}^1,\tilde{u}^2)\big)_x\right)^2<\infty.\end{aligned}$$

Similar arguments work for the case of M with non negative sectional curvatures.

Theorem 7.2 (Differentiability of exponential map) *Let $(\mu_t)\subset\mathcal{P}_c(M)$ be a regular curve, (v_t) its velocity vector field, (u_t) an absolutely continuous vector field along it and define $\nu_t:=(\exp(u_t))_\#\mu_t$. Assume that the vectors u_t are equibounded in $L^\infty_{\mu_t}$ and that*

$$\int_0^1\left\|\frac{\boldsymbol{d}}{dt}u_t\right\|^2_{\mu_t}dt<\infty.$$

Then (ν_t) is absolutely continuous and its velocity vector field (w_t) is given by

$$w_t:=\mathrm{P}_{\nu_t}\left(j_{\mu_t,u_t}\Big(v_t,\frac{\boldsymbol{d}}{dt}u_t\Big)\right). \tag{7.2}$$

If the manifold M has non-negative sectional curvatures, then the conclusion is true also dropping the hypothesis of uniform bound of L^∞ norms and the compactness hypothesis on the supports of the μ_t's.

Proof As usual, denote by $\mathbf{T}(t,s,\cdot)$ the flow maps of the curve (μ_t) and by τ_t^s the associated translation maps. For any $t,s\in[0,1]$ define

$$\boldsymbol{\gamma}_t^s := \Big(\exp(u_t)\circ\mathbf{T}(0,t,\cdot),\exp(u_s)\circ\mathbf{T}(0,s,\cdot)\Big)_{\#}\mu_0,$$

so that $\boldsymbol{\gamma}_t^s\in \mathcal{Adm}(\nu_t,\nu_s)$. By equation (1.11), the cost of $\boldsymbol{\gamma}_t^s$ is bounded by

$$\int \mathrm{d}^2(X,Y)d\boldsymbol{\gamma}_t^s(X,Y)\leq\left(\cosh(S\sqrt{\mathcal{C}})+\frac{\sinh S\sqrt{\mathcal{C}}}{S\sqrt{\mathcal{C}}}\right)^2\int \mathrm{D}^2\big((x,u_t(x)),(y,u_s(y))\big)d\boldsymbol{\sigma}_t^s(x,y),$$

where $\boldsymbol{\sigma}_t^s:=(\mathbf{T}(0,t,\cdot),\mathbf{T}(0,s,\cdot))_{\#}\mu_0$, S is the essential bound of all the u_t's in $L^\infty_{\mu_t}$ and $\mathcal{C}$ is a bound of the curvature of M on the relatively compact set $\{\exp_x(tu_s(x))\}_{t,s\in[0,1],x\in\mathrm{supp}(\mu_s)}$.

The value of $\mathrm{D}^2\big((x,u_t(x)),(y,u_s(y))\big)$ can be bounded from above by $\mathrm{d}^2(x,y)+|\tau_s^t(u_s(y))-u_t(x)|^2$, therefore it holds

$$\begin{aligned}\int \mathrm{D}^2\big((x,v_t(x)),&(y,v_s(y))\big)d\boldsymbol{\sigma}(x,y)\\ &\leq\int \mathrm{d}^2(x,y)+|\tau_s^t(u_s(y))-u_t(x)|^2 d\boldsymbol{\sigma}_t^s(x,t)\\ &\leq\int\left(\int_t^s\Big|v_r\big(\mathbf{T}(t,r,x)\big)\Big|dr\right)^2+\left(\int\Big|\frac{\boldsymbol{d}}{dr}u_r\big(\mathbf{T}(t,r,x)\big)\Big|\right)^2d\mu_t(x)\\ &\leq|s-t|\left(\int_t^s\|v_r\|^2_{\mu_r}dr+\int_t^s\left\|\frac{\boldsymbol{d}}{dr}u_r\right\|^2_{\mu_r}dr\right),\end{aligned}$$

which shows that (ν_t) is absolutely continuous. The fact that a choice of velocity vector field (not necessarily tangent) is given by the vectors $j_{\mu_t,u_t}\Big(v_t,\frac{\boldsymbol{d}}{dt}u_t\Big)$ defined in 7.1 comes observing that from

$$\nu_t:=\big(\exp(u_t)\circ\mathbf{T}(0,t,\cdot)\big)_{\#}\mu_0,$$

we have

$$\begin{aligned}\frac{d}{dt}\int\varphi d\nu_t=&\frac{d}{dt}\int\varphi\circ\exp(u_t)\circ\mathbf{T}(0,t,\cdot)d\mu_0\\ =&\int\big\langle(\nabla\varphi)\circ\big(\exp(u_t)\circ\mathbf{T}(0,t,\cdot)\big),J_{\mu,u}(\tilde u^1,\tilde u^2)\circ\mathbf{T}(0,t,\cdot)\big\rangle\,d\mu_0\\ =&\int\langle\nabla\varphi(x),\mathrm{v}\rangle\,d\boldsymbol{\gamma}_{\mu,u}(\tilde u^1,\tilde u^2)\\ =&\int\left\langle\nabla\varphi(x),\int\mathrm{v}d\big(\boldsymbol{\gamma}_{\mu,u}(\tilde u^1,\tilde u^2)\big)_x\right\rangle d\nu_t(x)\\ =&\int\big\langle\nabla\varphi,j_{\mu,u}(\tilde u^1,\tilde u^2)\big\rangle\,d\nu_t.\end{aligned}$$

The rest is obvious. □

Remark 7.3 (The result read as a result on infinitesimal perturbation) Recalling theorem 1.31, we can deduce from the result just proven the following fact. For any $t, s \in [0, 1]$ choose a plan $\boldsymbol{\gamma}_t^s \in \mathit{OptTan}(\nu_t, \nu_s)$ and define the rescaled plans

$$\boldsymbol{\sigma}_t^s := \frac{1}{s-t} \cdot \boldsymbol{\gamma}_t^s,$$

where the rescalation of a plan in $\mathscr{P}(TM)$ is defined in 1.5. Then for almost every $t \in [0, 1]$ the plans $\boldsymbol{\sigma}_t^s$ converge to $(Id, w_t)_\# \nu_t$ in $(\mathscr{P}_2(TM), W_2)$, where the w_t are given in equation (7.2) ■

Remark 7.4 (Tangent variation) It is worth underlying that the derivative of (u_t) appearing in the above theorem is the *total* one, and not the covariant one. This means that if (u_t) is a *tangent* and absolutely continuous vector field (a situation which fits better into the intrinsic Riemannian structure of $(\mathscr{P}_2(M), W_2)$ - see in particular the next chapter), the velocity vectors (w_t) of (ν_t) may be written as

$$w_t = \mathrm{P}_{\nu_t}\left(j_{\mu_t, u_t}\left(v_t, \frac{\boldsymbol{D}}{dt} u_t + \mathcal{N}_{\mu_t}(u_t, v_t) \right) \right) \tag{7.3}$$

■

Remark 7.5 (Non tangent velocities) Observe that in formula (7.2) the velocity vector field (v_t) is always tangent, while the total derivative $\frac{\boldsymbol{d}}{dt} u_t$ is generally not. This 'asymmetry' is due to the fact that we preferred to state the result considering regular curves. A more general version of the theorem (which can be proven analogously) works for curves (μ_t) for which there is a vector field $(\tilde{v}_t)$ compatible with it via the continuity equation (but not necessarily tangent) for which it holds

$$\int_0^1 \mathrm{L}(v_t) dt < \infty,$$
$$\int_0^1 \|v_t\|_{\mu_t}^2 dt < \infty.$$

In this situation, we have to ask for the vector field (u_t) to be absolutely continuous w.r.t. $(\tilde{v}_t)$ (plus uniform essential bound in $L^\infty_{\mu_t}$ if the curve has not compact support), and the result of the theorem would read as

$$w_t = \mathrm{P}_{\nu_t}\left(j_{\mu_t, u_t}\left(\tilde{v}_t, \frac{\tilde{\boldsymbol{d}}}{dt} u_t \right) \right),$$

see also remark 5.4. ■

7.3 A pointwise result

Having proved theorem 7.2 and taking into account remark 7.3, it is natural to ask whether the results expressed, which are valid for a.e. t, are valid also for *every* $t \in [0, 1]$, possibly under some additional hypothesis. Or, which is more or less the same, we may question whether remark 7.3 remains valid for the time $t = 0$.

The aim of this section is twofold: on one side we want to point out that this question is not trivial at all, in the sense that a pointwise result cannot be deduced by the 'a.e. one' by adding some kind of continuity assumptions (see example 7.13). The convergence that we prove here, is *much weaker* that the one expressed in remark 7.3. On the other hand, proposition 7.8 below, although weak, is enough to prove inequality (8.7) in the last section of the work. Therefore the result proven here deserves to produce a 'practical' consequence of the existence of Jacobi fields.

The following definition is quite natural:

Definition 7.6 (Perturbation of (μ, u)) *Let $\mu \in \mathscr{P}_2(M)$ and $u, \tilde{u}^1, \tilde{u}^2 \in L^2_\mu$ with $\tilde{u}^1 \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$. A perturbation of (μ, u) along the direction $(\tilde{u}^1, \tilde{u}^2)$ is a couple (μ_t), (u_t), where:*

- *(μ_t) is a regular curve defined on some interval $[0, T]$, $T > 0$, such that its velocity vector field (v_t) admits a continuous representative (not relabeled) satisfying $v_0 = \tilde{u}^1$,*

- *(u_t) is a C^1 vector field along (μ_t) satisfying*

$$
\begin{aligned}
u_0 &= u, \\
\frac{\boldsymbol{d}}{dt} u_t|_{t=0} &= \tilde{u}^2.
\end{aligned}
$$

Remark 7.7 It is not clear to us whether a perturbation exists for any $\tilde{u}^1 \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$. By proposition 6.1 we know that existence is ensured if $\tilde{u}^1 = \nabla\varphi$ for some $\varphi \in C_c^\infty(M)$: in this case an admissible choice of regular curve is simply given by $\mu_t := \mathbf{exp}_\mu(t\nabla\varphi)$. Then it is easy to produce a C^1 vector field along (μ_t) with the prescribed derivative in 0. ■

In the next proposition, we are going to consider the *barycentric projection* of a plan $\boldsymbol{\gamma} \in \mathscr{P}_2(TM)$, which is the function $\mathscr{B}(\boldsymbol{\gamma}) \in L^2_{\pi^M_\# \boldsymbol{\gamma}}$ defined by

$$
\mathscr{B}(\boldsymbol{\gamma})(x) := \int_{T_x M} \mathrm{v} d\boldsymbol{\gamma}_x(\mathrm{v}),
$$

where $\boldsymbol{\gamma}_x$ is the disintegration of $\boldsymbol{\gamma}$ w.r.t. the natural projection $\pi^M : TM \to M$. The trivial inequality

$$
\int_M \left|\mathscr{B}(\boldsymbol{\gamma})(x)\right|^2 d\pi^M_\# \boldsymbol{\gamma}(x) = \int_M \left| \int_{T_x M} \mathrm{v} d\boldsymbol{\gamma}_x(\mathrm{v}) \right|^2 d\pi^M_\# \boldsymbol{\gamma} \le \int_{TM} |\mathrm{v}|^2 d\boldsymbol{\gamma}(x, \mathrm{v}) < \infty,
$$

shows that $\mathscr{B}(\boldsymbol{\gamma}) \in L^2_{\pi^M_\# \boldsymbol{\gamma}}$.

Proposition 7.8 (Directional derivative of the exponential map) *Let $\mu \in \mathscr{P}_2(M)$, $u, \tilde{u}^1, \tilde{u}^2 \in L^2_\mu$ with $\tilde{u}^1 \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$, and (μ_t), (u_t) a perturbation of (μ, u) along the direction $(\tilde{u}^1, \tilde{u}^2)$. Define $\nu_t := \mathbf{exp}_{\mu_t}(u_t)$ and choose $\boldsymbol{\gamma}_0^t \in \mathit{OptTan}(\nu_0, \nu_t)$. Consider the rescaled plans*

$$
\boldsymbol{\sigma}_0^t := \frac{1}{t} \cdot \boldsymbol{\gamma}_0^t.
$$

Then:

A) Assume that $(\mu_t) \subset \mathcal{P}_c(M)$ *and that the vectors* u_t*'s are uniformly bounded in* $L^\infty_{\mu_t}$*. Then the maps* $\mathrm{P}_{\nu_0}(\mathscr{B}(\boldsymbol{\sigma}^t_0))$ *weakly converge to* $\mathrm{P}_{\nu_0}(j_{\mu,u}(\tilde{u}^1, \tilde{u}^2))$ *in* $L^2_{\nu_0}$ *as* $t \to 0$*.*

B) Assume that M *has non negative sectional curvatures. Then the same result as above is true without the compactness assumptions on the supports of* (μ_t) *and the uniform bound on the* u_t*'s.*

Proof For both the cases (A) and (B), the same arguments used in proof of theorem 7.2 together with the fact that (v_t) is continuous ((v_t) being the velocity vector field of (μ_t)) and (u_t) is C^1, give that (ν_t) is a Lipschitz curve. Therefore, letting L be its Lipschitz constant, we have

$$\|\mathrm{P}_{\nu_0}(\mathscr{B}(\boldsymbol{\sigma}^t_0))\|_{\nu_0} \le \|\mathscr{B}(\boldsymbol{\sigma}^t_0)\|_{\nu_0} \le \|\mathrm{v}\|_{\boldsymbol{\sigma}^t_0} = \frac{1}{t}\|\mathrm{v}\|_{\boldsymbol{\gamma}^s_t} = \frac{W_2(\nu_0, \nu_t)}{t} \le L.$$

Thus to conclude it is enough to check that

$$\lim_{t \to 0} \left\langle \mathscr{B}(\boldsymbol{\sigma}^t_0), \nabla\varphi \right\rangle_{\nu_0} = \left\langle j_{\mu,u}(\tilde{u}^1, \tilde{u}^2), \nabla\varphi \right\rangle_{\nu_0}, \qquad \forall \varphi \in C^\infty_c(M).$$

To prove this, fix $\varphi \in C^\infty_c(M)$ and define the function $t \mapsto f(t) := \int \varphi d\nu_t$. By theorem 7.2 we know that

$$f'(t) = \left\langle j_{\mu_t,u_t}\Big(v_t, \frac{\boldsymbol{d}}{dt}u_t\Big), \nabla\varphi \right\rangle_{\nu_t}, \qquad a.e.\ t. \tag{7.4}$$

Now we use the continuity hypothesis in the definition of perturbation: since both (v_t) and $\frac{\boldsymbol{d}}{dt}u_t$ are continuous, by the regularity of the Jacobi fields on M it is easy to check that vector field $(j_{\mu_t,u_t}(v_t, \frac{\boldsymbol{d}}{dt}u_t))$ is continuous along (ν_t). Therefore the right hand side of equation (7.4) is continuous. This means that f is C^1 and that (7.4) actually holds for every t. In particular it holds

$$\lim_{t \downarrow 0} \frac{f(t) - f(0)}{t} = \left\langle j_{\mu,u}(\tilde{u}^1, \tilde{u}^2), \nabla\varphi \right\rangle_{\nu_0}. \tag{7.5}$$

The same limit can be computed from

$$\begin{aligned}
\frac{f(t) - f(0)}{t} &= \frac{\int \varphi d\nu_t - \int \varphi d\nu_0}{t} = \frac{1}{t}\int \varphi(y) - \varphi(x) d(\pi^M, \exp)_\# \boldsymbol{\gamma}^t_0(x, y) \\
&= \frac{1}{t}\int \langle \nabla\varphi(x), \mathrm{v} \rangle \, d\boldsymbol{\gamma}^t_0(x, \mathrm{v}) + R(t) \\
&= \int \langle \nabla\varphi(x), \mathrm{v} \rangle \, d\boldsymbol{\sigma}^t_0(x, \mathrm{v}) + R(t) \\
&= \int \left\langle \nabla\varphi, \int \mathrm{v} d(\boldsymbol{\sigma}^t_0)_x(\mathrm{v}) \right\rangle d\nu_0(x) + R(t) \\
&= \left\langle \nabla\varphi, \mathscr{B}(\boldsymbol{\sigma}^t_0) \right\rangle_{\nu_0} + R(t).
\end{aligned}$$

So that to conclude we need only to show that the reminder term $R(t)$ goes to 0 with t. To see this, just integrate the trivial inequality

$$|\varphi(\exp_x(\mathrm{v})) - \varphi(x) - \langle \nabla\varphi(x), \mathrm{v} \rangle| \le \mathrm{Lip}(\nabla\varphi)\frac{|\mathrm{v}|^2}{2},$$

to obtain

$$|R(t)| \leq \frac{1}{t}\int |\varphi(\exp_x(\mathrm{v})) - \varphi(x) - \langle\nabla\varphi(x), \mathrm{v}\rangle| d\boldsymbol{\gamma}_0^t(x, \mathrm{v})$$

$$\leq \mathrm{Lip}(\nabla\varphi)\frac{\|\mathrm{v}\|^2_{\boldsymbol{\gamma}_0^t(\mathrm{t})}}{2t} \leq \mathrm{Lip}(\nabla\varphi)\frac{W_2^2(\nu_0, \nu_t)}{2t} \leq t\mathrm{Lip}(\nabla\varphi)\frac{L^2}{2}.$$

The thesis follows. □

Remark 7.9 It can be proved that, with the same notation as above, $\mathscr{B}(\boldsymbol{\sigma}_t^s)$ actually belongs to $\mathrm{Tan}_{\nu_t}(\mathscr{P}_2(M))$ (see [12]). Therefore the above proposition actually tells that $\mathscr{B}(\boldsymbol{\sigma}_0^t)$ weakly converges to $\mathrm{P}_{\nu_0}(j_{\mu,u}(\tilde{u}^1, \tilde{u}^2))$. ■

Remark 7.10 Observe that we didn't prove any kind of converging result at the level of plans, but only at the level of barycentric projections. Our conjecture here is the following. If the map $\exp(u)$ is μ-essentially invertible or ν_0 is absolutely continuous w.r.t. the volume measure, then the family of plans $\boldsymbol{\sigma}_0^t$ converge to $(Id, \mathrm{P}_{\nu_0}(j_{\mu,u}(\tilde{u}^1, \tilde{u}^2)))_{\#}\nu_0$ in $\mathscr{P}_2(TM)$ as $t \downarrow 0$. If the map $\exp(u)$ is not essentially invertible, then the family of plans $\boldsymbol{\sigma}_0^t$ should still converge in an appropriate sense to the 'projection onto the tangent space at ν_0' of the plan $(\exp(u), J_{\mu,u}(\tilde{u}^1, \tilde{u}^2))_{\#}\mu$ (see Chapter 4 of [10] for a discussion about the projection operator on plans). ■

Remark 7.11 In the proof of the above proposition, we never used the fact that $(\pi^M, \exp)_{\#}\boldsymbol{\gamma}_0^t$ is optimal. The only properties used, which are therefore sufficient to have the same conclusion, are

$$(\pi^M, \exp)_{\#}\boldsymbol{\gamma}_0^t \in \mathscr{Adm}(\nu_0, \nu_t),$$

$$\overline{\lim_{t\to 0}} \frac{\|\mathrm{v}\|_{\boldsymbol{\gamma}_0^t}}{t} < \infty$$

■

Remark 7.12 (How to read the result for tangent variations) Observe that in the definition of perturbation we asked for the vector $\tilde{u}^1$ to be tangent, while the same requirement is not imposed on $\tilde{u}^2$. This is motivated by the look for the maximal generality: the same actually happens in theorem 7.2, where the base measure μ_t is moved accordingly to the tangent vector field v_t, while the vector field u_t is moved along the general direction $\frac{\boldsymbol{d}}{dt}u_t$ (clearly we can perturb the mass also along non tangent directions, like in remark 7.5, but we won't stress this point any further). Also, the result on differentiation of the exponential map looks simpler - and to some extent more natural - when considering general vector fields, rather than tangent ones (compare formula (7.2) with (7.3)).

If $u \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$, then one can speak of tangent perturbations, like in remark 7.4. In this case the correct formulation of perturbation for the vector field (u_t) is the following: (u_t) is a C^1 and tangent vector field along (μ_t) such that $\mathrm{P}_{\nu_0}\big(\frac{\boldsymbol{d}}{dt}u_{t|_{t=0}}\big) = \tilde{u}^2$ (observe that if (u_t) is C^1 and (v_t) is continuous, then not necessarily $(\frac{\boldsymbol{D}}{dt}u_t)$ is continuous as well - see example 7.13 below).

In this case, if we assume also that $\tilde{u}^1$ is Lipschitz, the limit vector in proposition 7.8 can be written in terms of $\tilde{u}^1, \tilde{u}^2$ as

$$\mathrm{P}_{\nu_0}\Big(j_{\mu,u}(\tilde{u}^1, \tilde{u}^2 + \mathcal{N}_{\nu_0}(u, \tilde{u}^1))\Big).$$

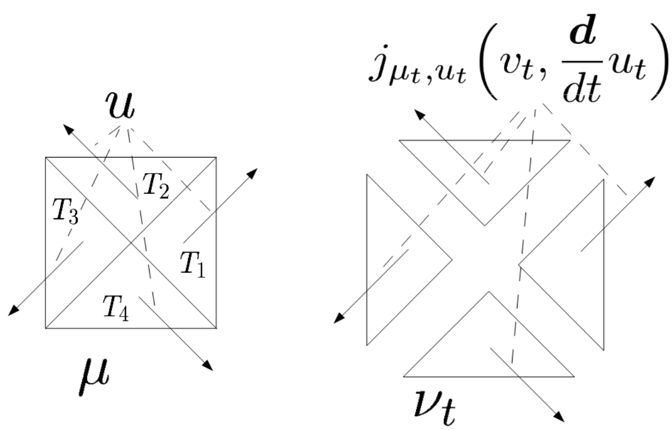
In the next chapter, we are going to see that vector fields of this kind are the (only) solutions of the Jacobi equation, thus closing the gap between the work done here in the identification of the differential of the exponential map, and the description of the curvature tensor done in the previous chapter. ■

We conclude this section by showing that the pointwise result cannot be derived by continuity arguments from theorem 7.2. The main issue here is the fact that there are no reasonable assumptions on (μ_t) and (u_t) which guarantees that the vector field

$$\mathrm{P}_{\nu_t}\left(j_{\mu_t,u_t}\Big(v_t, \frac{\boldsymbol{d}}{dt}u_t\Big)\right)$$

is continuous: indeed, the continuity of $j_{\mu_t,u_t}(v_t, \frac{\boldsymbol{d}}{dt}u_t)$, that we used in the proof of 7.8, is not enough to conclude, since the projection operator may very well be discontinuous. The following example should clarify this fact. Observe the similarity with the example of non-existence of parallel transport 4.15.

Example 7.13 Let $M = \mathbb{R}^2$ and $\mu_t \equiv \mu := \mathcal{L}^2|_{[0,1]^2}$ for any t. Divide the unit square into four triangles by its diagonals, as shown in the picture.



Define the vector field $u : [0,1]^2 \to \mathbb{R}^2$ as:

$$u(x) := \begin{cases} (1,1), & \text{if } x \in T_1, \\ (-1,1), & \text{if } x \in T_2, \\ (-1,-1), & \text{if } x \in T_3, \\ (1,-1), & \text{if } x \in T_4, \end{cases}$$

and $u_t := tu$. Then the curve (μ_t) is clearly regular and its velocity vector field is continuous (because it is constantly equal to 0), and the vector field (u_t) along it is C^∞. Also, it is immediate to check that it holds

$$j_{\mu_t,u_t}\Big(v_t, \frac{\boldsymbol{d}}{dt}u_t\Big) \circ \exp(u_t) = j_{\mu,tu}(0,u) \circ (Id + tu) = u,$$

and that the support of $\nu_t := \mathbf{exp}_{\mu_t}(u_t) = (Id + tu)_\# \mu$ is, for positive times, made of four different connected components, like in the picture above.

Since for positive times the vector field $j_{\mu_t,u_t}(v_t, \frac{\boldsymbol{d}}{dt}u_t)$ is constant on each connected component of $\mathrm{supp}(\nu_t)$, it is easy to check that it is tangent. Then, since $(j_{\mu_t,u_t}(v_t, \frac{\boldsymbol{d}}{dt}u_t))$ is continuous, our example will be concluded if we show that $j_{\mu_0,u_0}(v_0, \frac{\boldsymbol{d}}{dt}u_0) = u \notin \mathrm{Tan}_\mu(\mathscr{P}_2(M))$. A quick way to see this, is to check that the vector field $w \in L^2_\mu$ defined by

$$w(x) := \begin{cases} (0,1), & \text{if } x \in T_1, \\ (-1,0), & \text{if } x \in T_2, \\ (0,-1), & \text{if } x \in T_3, \\ (1,0), & \text{if } x \in T_4, \end{cases}$$

belongs to $\mathrm{Tan}^\perp_\mu(\mathscr{P}_2(M))$ (since by direct calculation one verifies that $\langle w, \nabla\varphi\rangle_\mu = 0$ for any $\varphi \in C^\infty_c(\mathbb{R}^2)$). Then, since it holds $\langle w, u\rangle_\mu \neq 0$ it cannot hold $u \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$. ■

8 Jacobi fields

In this chapter we study the Jacobi fields on our 'manifold' $(\mathcal{P}_c(M), W_2)$. What we know about basic Riemannian geometry, is that the differential of the exponential map satisfies the Jacobi equation: we are going to see that the same is true in the Wasserstein space.

The chapter is organized as follows. In the first section, we introduce the Jacobi equation: we will see when it is meaningful and study, from an abstract point of view, which are the main properties of its solutions. Then, in the second section, we will see that the differential of the exponential map that we calculated in the previous chapter, actually identifies all the solution of the Jacobi equation. Finally, in the last section, we will use the knowledge we gathered on the Jacobi fields to derive some quantitative estimate on the regularity of the map $\nu \mapsto \{\text{optimal transport map from a fixed measure } \mu \text{ to } \nu\}$.

8.1 The Jacobi equation

We have a curvature tensor, so we can write down the Jacobi equation. Let $(\mu_t) \subset \mathcal{P}_c(M)$ be a geodesic and (v_t) its velocity vector field: the Jacobi equation along (μ_t) is:

$$\frac{\boldsymbol{D}^2}{dt^2}\mathbf{J}_t + \mathbf{R}(v_t, \mathbf{J}_t)v_t = 0. \tag{8.1}$$

Let us understand when this equation makes sense. The first term presents no problem: it simply asks for the solution $(\mathbf{J}_t)$ to be a $C^{1,1}$ vector field along (μ_t). In the second term, however, it appears the velocity vector field of the geodesic, which exists for every time except, possibly, 0 and 1. This means that the equation may be not meaningful outside the open interval $(0,1)$. For $0 < t < 1$, the vector v_t is well defined and Lipschitz, thus from the discussion made at the end of section 6.2, we know that the operator $\mathbf{R}(v_t, \cdot)v_t : \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M)) \to \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$ is well defined and continuous. So the Jacobi equation is always well defined in the interval $(0,1)$.

The fact that the equation is possibly meaningless for $t = 0$, implies that we cannot prescribe the initial condition on its solution by imposing the value of $\mathbf{J}_t$ and its covariant derivative at $t = 0$ (like we are used to do in common Riemannian manifolds). Rather, we can only give 'intermediate' conditions of the kind

$$\begin{cases} \mathbf{J}_{t_0} & = u_{t_0}, \\ \dfrac{\boldsymbol{D}}{dt}\mathbf{J}_t \ |_{t=t_0} & = u'_0, \end{cases}$$

for some $t_0 \in (0,1)$ and certain given $u_{t_0}, u'_{t_0} \in \mathrm{Tan}_{\mu_{t_0}}(\mathscr{P}_2(M))$.

In order to simplify the exposition, we will use the following approach: we will assume that (μ_t) is the restriction to $[0,1]$ of some geodesic defined on a larger interval $[-\varepsilon, 1+\varepsilon]$, so that the Jacobi equation will be meaningful in the whole $[0,1]$ and we can prescribe the initial conditions in the classical way at $t = 0$. It is obvious that once we will be able to solve the equation in this context, a simple step-by-step argument allows to produce a solution in the internal part of any geodesic in $\mathcal{P}_c(M)$.

Thus throughout this Chapter we will use the following assumption and notation:

Assumption and notation: (μ_t) is a fixed geodesic on $[0,1]$ that is the restriction of a geodesic defined on a larger interval $[-\varepsilon, 1+\varepsilon]$, (v_t) is its velocity vector field, and the constants $L, S, \mathcal{C} \in \mathbb{R}$ are defined by:

$$L := \sup_{t\in[0,1]} \mathrm{L}(v_t),$$

$$S := \sup_{t\in[0,1]} \mathrm{S}(v_t) = \mathrm{S}(v_0),$$

$\mathcal{C}$:= a bound on the curvature of M on the compact set $\cup_{t\in[0,1]} \mathrm{supp}(\mu_t)$.

Definition 8.1 (Solutions of the Jacobi equation) *We say that $(\mathbf{J}_t)$ is a solution of the Jacobi equation with initial conditions $u_0, u'_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ if it is a $C^{1,1}$ vector field along (μ_t), satisfies*

$$\begin{cases} \mathbf{J}_0 & = u_0, \\ \dfrac{\boldsymbol{D}}{dt}\mathbf{J}_t \ |_{t=0} & = u'_0, \end{cases} \tag{8.2}$$

and the equation (8.1) *is fulfilled for a.e. $t \in [0,1]$.*

Note that since $(\mathbf{J}_t)$ is $C^{1,1}$, by proposition 5.31 its covariant derivative is absolutely continuous and thus the initial condition makes sense.

We have the following result:

Proposition 8.2 (Existence, uniqueness and regularity) *For any* $u_0, u_0' \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ *there exists a unique solution* $(\mathbf{J}_t)$ *of the Jacobi equation satisfying the initial conditions* (8.2). *Such solution is a* C^∞ *vector field along* (μ_t).

Proof Observe that $(\mathbf{J}_t)$ solves the Jacobi equation with the prescribed initial conditions if and only if the curve $t \mapsto \overline{J}_t := \mathcal{T}_t^0(\mathbf{J}_t) \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ solves

$$\begin{cases} \dfrac{d^2}{dt^2}\overline{J}_t + A_t(\overline{J}_t) = 0, \\ \overline{J}_0 = u_0, \\ \dfrac{d}{dt}\overline{J}_t|_{t=0} = u_0', \end{cases}$$

where $A_t : \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M)) \to \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ is defined, for every $t \in [0,1]$, as

$$A_t(u) := \tau_t^0\Big(\mathbf{R}\big(v_t, \tau_0^t(u)\big)v_t\Big), \qquad \forall u \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M)).$$

Since it holds

$$\|A_t\|_{\mathrm{op}} = \|\mathbf{R}(v_t, \cdot)v_t\|_{\mathrm{op}} \le \mathcal{C}S^2 + 3L^2,$$

by standard arguments there is a unique $C^{1,1}$ solution $(\overline{J}_t)$ of the above system. It remains to prove that the corresponding $C^{1,1}$ solution $(\mathbf{J}_t)$ is C^∞. This is a consequence of the smoothness of $\mathbf{R}(v_t, \cdot)v_t$. Indeed, since $(\mathbf{J}_t)$ is $C^{1,1}$, from the equality

$$\frac{\boldsymbol{D}^2}{dt^2}\mathbf{J}_t = -\mathbf{R}(v_t, \mathbf{J}_t)v_t,$$

we get that the second covariant derivative of $(\mathbf{J}_t)$ is $C^{1,1}$. By proposition 5.31 and remark 6.9, it follows that $(\mathbf{J}_t)$ is $C^{3,1}$. But then, again from the Jacobi equation, its second covariant derivative is $C^{3,1}$, and so on. The thesis follows. □

8.2 Solutions of the Jacobi equation

Here we prove that the solutions of the Jacobi equation can be explicitly characterized as the differential of the exponential map, as in usual Riemannian manifolds.

In order to prove this, it is better to introduce the following notation. Recall that since our geodesic (μ_t) is the restriction of a geodesic defined in some larger interval, the velocity vector field (v_t) is well defined at $t = 0$; also, for any $t \in [0,1]$ the map $\exp(tv_0)$ is invertible, its inverse being $\exp(-tv_t)$.

Definition 8.3 (The vector field $J_t(u_1, u_2)$) *For any couple of vector fields* $u_1, u_2 \in L^2_{\mu_{t_0}}$ *(thus not necessarily tangent), the vector field* $J_t(u_1, u_2) \in L^2_{\mu_t}$, $t \in [0,1]$ *is defined by*

$$J_t(u_1,u_2)\big(\exp_x(tv_0(x))\big) := \begin{cases} \text{the value at } s = t \text{ of the (usual) Jacobi field } j_s \text{ along the} \\ \text{geodesic } s \mapsto \exp_x(sv_0(x)) \text{ which has the} \\ \text{initial conditions } j_0 = u_1(x), \text{ and } j_0' = u_2(x). \end{cases}$$

The invertibility of $\exp(tv_0)$ ensures that $J_t(u_1,u_2)$ is well defined for μ_t−a.e. x; furthermore, from inequality (1.8b) we have that

$$\|J_t(u_1,u_2)\|^2_{\mu_t} \le 2\|u_1\|^2_{\mu_0}\cosh^2\left(S\sqrt{\mathcal{C}}\right) + 2\frac{\|u_2\|^2_{\mu_0}}{S^2\mathcal{C}}\sinh^2\left(S\sqrt{\mathcal{C}}\right).$$

With respect to the notation of the previous chapter, here we just added the time dependence, indeed it holds:

$$J_1(u_1,u_2) = j_{\mu_0,v_0}(u_1,u_2),$$

By proposition 3.13 and the regularity of the (usual) Jacobi fields on M, we know that $\big(J_t(u_1,u_2)\big)$ is a C^∞ vector field along (μ_t) and that it holds

$$\begin{cases} \dfrac{\boldsymbol{d}^2}{dt^2}J_t + R(v_t,J_t)v_t = 0, \\ J_0 = u_1 \\ \dfrac{\boldsymbol{d}}{dt}J_t|_{t=0} = u_2. \end{cases}$$

where we wrote J_t for $J_t(u_1,u_2)$.

Motivated by remark 7.12 (and the analogous 7.4), we are lead to suppose that the solution of the Jacobi equation with initial conditions (8.2) is given by the vector field $(\mathbf{J}_t)$ defined by

$$\mathbf{J}_t := \mathrm{P}_{\mu_t}\Big(J_t\big(u_0, u_0' - \mathcal{N}_{\mu_0}(u_0,v_0)\big)\Big). \tag{8.3}$$

(well, actually in remark 7.12, read with current notation, the Lipschitz vector field was u_0, while v_0 was arbitrary. However, here we know that v_0 is Lipschitz and that some solution must exists for any u_0, so that by the properties of $\mathcal{N}_\mu$ the 'only possible guess' is given by the above formula also for u_0 non Lipschitz).

This is actually true, as we prove now:

Theorem 8.4 (Jacobi fields) *For any* $u_0, u_0' \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$*, the vector field* $(\mathbf{J}_t)$ *defined by* (8.3) *solves the Jacobi equation* (8.1) *with the initial conditions* (8.2).

Proof Fix $u_0, u_0' \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$ and let

$$J_t := J_t\big(u_0, u_0' - \mathcal{N}_{\mu_0}(u_0,v_0)\big).$$

We know that (J_t) is a C^∞ vector field along (μ_t) and that $\big(\mathrm{P}_{\mu_t}(\cdot)\big)$ is C^∞ as well. Thus $(\mathbf{J}_t)$ is C^∞. Notice that

$$\mathbf{J}_0 = \mathrm{P}_{\mu_0}(J_0) = \mathrm{P}_{\mu_0}(u_0) = u_0,$$

so that the first initial condition is satisfied. Now evaluate the covariant derivative of $(\mathbf{J}_t)$ using theorem 5.14:

$$\frac{\boldsymbol{D}}{dt}\mathbf{J}_t = \mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}J_t - \mathcal{O}^*_{v_t}(J_t)\right).$$

Evaluate this derivative at $t=0$ and recall that since $u_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$ we have $\mathcal{O}^*_{v_t}(u_0)=0$, to get

$$\frac{\boldsymbol{D}}{dt}\mathbf{J}_t|_{t=0} = \mathrm{P}_{\mu_0}\left(u_0' - \mathcal{N}_{\mu_0}(u_0,v_0) - \mathcal{O}^*_{v_t}(J_0)\right) = u_0'.$$

Thus also the second initial condition is fulfilled.

Using again theorem 5.14, corollary 5.23 and recalling that $\frac{\boldsymbol{d}}{dt}v_t \equiv 0$ we can compute the second covariant derivative of $(\mathbf{J}_t)$:

$$\begin{aligned}
\frac{\boldsymbol{D}^2}{dt^2}\mathbf{J}_t &= \frac{\boldsymbol{D}}{dt}\mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}J_t - \mathcal{O}^*_{v_t}(J_t)\right) \\
&= \mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}^2}{dt^2}J_t - \mathcal{O}^*_{v_t}\left(\frac{\boldsymbol{d}}{dt}J_t\right) + R(v,\mathrm{P}^\perp_{\mu_t}(J_t))v_t + \mathcal{O}^*_{v_t}\left(\mathcal{O}^*_{v_t}(J_t)\right) - \mathcal{O}^*_{v_t}\left(\mathcal{O}_{v_t}(\mathrm{P}_{\mu_t}(J_t))\right)\right) \\
&\qquad - \mathrm{P}_{\mu_t}\left(\mathcal{O}^*_{v_t}\left(\frac{\boldsymbol{d}}{dt}J_t - \mathcal{O}^*_{v_t}(J_t)\right)\right) \\
&= -\mathrm{P}_{\mu_t}\Big(R(v_t,J_t)v_t\Big) - 2\mathrm{P}_{\mu_t}\left(\mathcal{O}^*_{v_t}\left(\frac{\boldsymbol{d}}{dt}J_t\right)\right) + \mathrm{P}_{\mu_t}\left(R(v_t,\mathrm{P}^\perp_{\mu_t}(J_t))v_t\right) \\
&\qquad + 2\mathrm{P}_{\mu_t}\left(\mathcal{O}^*_{v_t}\left(\mathcal{O}^*_{v_t}(J_t)\right)\right) - \mathrm{P}_{\mu_t}\Big(\mathcal{O}^*_{v_t}\left(\mathcal{O}_{v_t}(\mathrm{P}_{\mu_t}(J_t))\right)\Big) \\
&= -\mathrm{P}_{\mu_t}\Big(R(v_t,\mathbf{J}_t)v_t\Big) - 2\mathrm{P}_{\mu_t}\left(\mathcal{O}^*_{v_t}\left(\frac{\boldsymbol{d}}{dt}J_t\right)\right) + 2\mathrm{P}_{\mu_t}\left(\mathcal{O}^*_{v_t}\left(\mathcal{O}^*_{v_t}(J_t)\right)\right) - \mathrm{P}_{\mu_t}\Big(\mathcal{O}^*_{v_t}\left(\mathcal{O}_{v_t}(\mathbf{J}_t)\right)\Big).
\end{aligned}$$

Recalling that

$$\mathbf{R}(v_t,\mathbf{J}_t)v_t = \mathrm{P}_{\mu_t}\Big(R(v_t,\mathbf{J}_t)v_t\Big) + 3\mathrm{P}_{\mu_t}\Big(\mathcal{O}^*_{v_t}\left(\mathcal{O}_{v_t}(\mathbf{J}_t)\right)\Big),$$

we just proved that

$$\frac{\boldsymbol{D}^2}{dt^2}\mathbf{J}_t + \mathbf{R}(v_t,\mathbf{J}_t)v_t = -2\mathrm{P}_{\mu_t}\left(\mathcal{O}^*_{v_t}\left(\frac{\boldsymbol{d}}{dt}J_t - \mathcal{O}^*_{v_t}(J_t) - \mathcal{O}_{v_t}(\mathrm{P}_{\mu_t}(J_t))\right)\right)$$

therefore our thesis will be achieved if we show that the right hand side of the above equation is identically 0. To this aim, it is sufficient to show that the vector field (u_t) defined by

$$u_t := \mathrm{P}^\perp_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}J_t - \mathcal{O}^*_{v_t}(J_t) - \mathcal{O}_{v_t}(\mathrm{P}_{\mu_t}(J_t))\right),$$

is identically 0. Notice that since we are along a geodesic in $\mathcal{P}_c(M)$, the gradient of v_t is symmetric: that is, by proposition 5.12, it holds

$$\mathrm{P}^\perp_{\mu_t}(\mathcal{O}^*_{v_t}(J_t)) = \mathcal{O}_{v_t}\left(\mathrm{P}^\perp_{\mu_t}(J_t)\right),$$

and thus we have

$$u_t = \mathrm{P}^\perp_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}J_t - \mathcal{O}_{v_t}(J_t)\right).$$

(Actually, this - important - step of the proof is the only place where we needed proposition 5.12. Thus, it is due to this proof the hard work that we did in section 2.3 to prove that the velocity vector fields of a geodesic are not just Lipschitz, but gradients of $C^{1,\infty}$ functions). The value at $t = 0$ of u_t is:

$$u_0 = \mathrm{P}^{\perp}_{\mu_0}\Big(u'_0 + \mathcal{N}_{\mu_0}(v_0, u_0) - \mathcal{O}_{v_0}(u_0)\Big) = \mathrm{P}^{\perp}_{\mu_0}\Big(\mathcal{N}_{\mu_0}(v_0, u_0) - \mathcal{N}_{\mu_0}(v_0, u_0)\Big) = 0.$$

It is clear that (u_t) is C^∞ along (μ_t). Evaluate its total derivative:

$$\begin{aligned}
\frac{\boldsymbol{d}}{dt}u_t &= \frac{\boldsymbol{d}}{dt}\left(\mathrm{P}^{\perp}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}J_t\right) - \mathcal{O}_{v_t}(J_t)\right)\\
&= \mathrm{P}^{\perp}_{\mu_t}\left(\frac{\boldsymbol{d}^2}{dt^2}J_t\right) + \mathrm{P}_{\mu_t}\left(\mathcal{O}^*_{v_t}\left(\frac{\boldsymbol{d}}{dt}J_t\right)\right) + \mathcal{O}_{v_t}\left(\mathrm{P}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}J_t\right)\right) - \mathcal{O}_{v_t}\left(\frac{\boldsymbol{d}}{dt}J_t\right)\\
&\qquad + \mathrm{P}^{\perp}_{\mu_t}(R(v_t, J_t)v_t) + \mathcal{O}_{v_t}(\mathcal{O}_{v_t}(J_t)) + \mathrm{P}_{\mu_t}\Big(\mathcal{O}^*_{v_t}(\mathcal{O}_{v_t}(J_t))\Big)\\
&= \mathrm{P}_{\mu_t}\left(\mathcal{O}^*_{v_t}\left(\frac{\boldsymbol{d}}{dt}J_t - \mathcal{O}_{v_t}(J_t)\right)\right) - \mathcal{O}_{v_t}\left(\mathrm{P}^{\perp}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}J_t\right) - \mathcal{O}_{v_t}(J_t)\right)\\
&= \mathrm{P}_{\mu_t}\Big(\mathcal{O}^*_{v_t}(u_t)\Big) - \mathcal{O}_{v_t}(u_t).
\end{aligned}$$

Define $f : [0,1] \to [0,1]$ as $f(t) := \|u_t\|_{\mu_t}$. We know that f is absolutely continuous, $f(0) = 0$ and that

$$f'(t) \le \left\|\frac{\boldsymbol{d}}{dt}u_t\right\|_{\mu_t} \le \left\|\mathrm{P}_{\mu_t}\Big(\mathcal{O}^*_{v_t}(u_t)\Big)\right\|_{\mu_t} + \|\mathcal{O}_{v_t}(u_t)\|_{\mu_t} \le 2\mathrm{L}(v_t)\|u_t\|_{\mu_t} = 2\mathrm{L}(v_t)f(t), \quad a.e.\ t \in [0,1].$$

From the Gronwall lemma we deduce that $f(t) \equiv 0$, which means $u_t \equiv 0$. This concludes the proof. □

Remark 8.5 (The case $M = \mathbb{R}^d$) If $M = \mathbb{R}^d$, the norm of the Jacobi field $J_t(u_1, u_2)$ grows linearly in time, regardless of the compactness assumption of the supports of the μ_t's, therefore the vector field $(\mathbf{J}_t)$ is well defined for general $(\mu_t) \subset \mathscr{P}_2(\mathbb{R}^d)$. Also, by remark 5.13, we know that the gradients of the velocity vector field of (μ_t) are symmetric and, by remark 6.5, that the curvature tensor is well defined for any $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, and not just for those with compact support. Thus the previous theorem is true in the Euclidean case even without the compactness assumption. ■

Remark 8.6 (A way to check tangency) In the last part of the proof we used a particular case of the following general statement. Let (μ_t) be a regular curve, (v_t) its velocity vector field and (w_t) an absolutely continuous vector field such that $w_0 \in \mathrm{Tan}_{\mu_0}(\mathscr{P}_2(M))$. Then $w_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$ for every t if and only if

$$\frac{\boldsymbol{d}}{dt}w_t + \mathcal{O}_{v_t}(w_t) \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M)), \qquad a.e.\ t \in [0,1].$$

Indeed the 'only if' is obvious from theorem 5.15. For the 'if' consider the vector field $u_t := \mathrm{P}^{\perp}_{\mu_t}(w_t)$, notice that

$$\begin{aligned}\frac{\boldsymbol{d}}{dt}u_t &= \mathrm{P}^{\perp}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}w_t\right) + \mathrm{P}_{\mu_t}(\mathcal{O}^*_{v_t}(w_t)) + \mathcal{O}_{v_t}(\mathrm{P}_{\mu_t}(w_t))\\ &= \mathrm{P}^{\perp}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}w_t + \mathcal{O}_{v_t}(w_t)\right) + \mathrm{P}_{\mu_t}\Big(\mathcal{O}^*_{v_t}\left(\mathrm{P}^{\perp}_{\mu_t}(w_t)\right)\Big) - \mathcal{O}_{v_t}\left(\mathrm{P}^{\perp}_{\mu_t}(w_t)\right)\\ &= \mathrm{P}_{\mu_t}\Big(\mathcal{O}^*_{v_t}(u_t)\Big) - \mathcal{O}_{v_t}(u_t),\end{aligned}$$

and conclude as before by applying the Gronwall lemma to the function $f(t) := \|u_t\|_{\mu_t}$ ∎

Remark 8.7 (A more geometrical way to conclude) The last part of the proof of theorem 8.4 concerns with the proof that

$$\mathrm{P}^{\perp}_{\mu_t}\left(\frac{\boldsymbol{d}}{dt}J_t - \mathcal{O}_{v_t}(J_t)\right) = 0, \qquad \forall t. \tag{8.4}$$

We proceeded in an analytic way by calling into play the Gronwall lemma, but from a geometrical perspective this is not particularly insightful.

A different approach could be the following (we are going to be sloppy about regularity assumptions - just think that everything is C^∞): the vector field in (8.4) is equal to 0 if and only if the vector field

$$\frac{\boldsymbol{d}}{dt}J_t - \nabla v_t^{\mathrm{t}} \cdot J_t$$

is tangent for any t. This means that we should have something like:

$$\frac{\boldsymbol{d}}{dt}J_t - \nabla v_t^{\mathrm{t}} \cdot J_t = \nabla \phi_t, \tag{8.5}$$

and the geometrical question we could ask is: 'who' is ϕ_t? We are going to give the answer - and run the calculations - in the case $M = \mathbb{R}^d$.

Assume that (J_t) satisfies the initial conditions

$$\begin{aligned}J_0 &= \nabla\psi,\\ \frac{\boldsymbol{d}}{dt}J_t|_{t=0} &= \nabla\varphi + \mathrm{P}^{\perp}_{\mu_0}(\nabla v_0^{\mathrm{t}} \cdot \nabla\psi),\end{aligned}$$

let ϕ be such that

$$\nabla\phi = \nabla\varphi + \mathrm{P}_{\mu_0}(\nabla v_0^{\mathrm{t}} \cdot \nabla\psi),$$

so that the initial conditions may be written as

$$\begin{aligned}J_0 &= \nabla\psi,\\ \frac{\boldsymbol{d}}{dt}J_t|_{t=0} &= \nabla\phi + \nabla v_0^{\mathrm{t}} \cdot \nabla\psi.\end{aligned}$$

We claim that equation (8.5) is satisfied with $\phi_t := \phi \circ \mathbf{T}(t, 0, \cdot)$.

Indeed, observe that on $\mathbb{R}^d$ we have

$$\begin{aligned}
\mathbf{T}(0, t, \cdot) &= Id + tv_0, \\
\mathbf{T}(t, 0, \cdot) &= (Id + tv_0)^{-1} = Id - tv_t,
\end{aligned}$$

and that Jacobi fields on $\mathbb{R}^d$ are linear vector field, so that

$$\begin{aligned}
J_t \circ (Id + tv_0) &= \nabla\psi + t(\nabla\phi + \nabla v_0^{\mathrm{t}} \cdot \nabla\psi), \\
\frac{\boldsymbol{d}}{dt} J_t \circ (Id + tv_0) &= \nabla\phi + \nabla v_0^{\mathrm{t}} \cdot \nabla\psi,
\end{aligned}$$

or, which is the same:

$$\begin{aligned}
J_t &= \big(\nabla\psi + t(\nabla\phi + \nabla v_0^{\mathrm{t}} \cdot \nabla\psi)\big) \circ (Id - tv_t), \\
\frac{\boldsymbol{d}}{dt} J_t &= (\nabla\phi + \nabla v_0^{\mathrm{t}} \cdot \nabla\psi) \circ (Id - tv_t).
\end{aligned}$$

Thus we have

$$\begin{aligned}
\frac{\boldsymbol{d}}{dt} J_t - \nabla v_t^{\mathrm{t}} \cdot J_t &= (\nabla\phi + \nabla v_0^{\mathrm{t}} \cdot \nabla\psi) \circ (Id - tv_t) \\
&\quad - \nabla v_t^{\mathrm{t}} \cdot \Big(\nabla\psi + t(\nabla\phi + \nabla v_0^{\mathrm{t}} \cdot \nabla\psi)\Big) \circ (Id - tv_t) \\
&= \underbrace{\nabla\phi \circ (Id - tv_t) - t\nabla v_t^{\mathrm{t}} \cdot \nabla\phi \circ (Id - tv_t)}_{(\mathbf{A})} \\
&\quad + \Big(\big(\underbrace{\nabla v_0^{\mathrm{t}} - \nabla v_t^{\mathrm{t}} \circ (Id + tv_0) - t\nabla v_t^{\mathrm{t}} \circ (Id + tv_0) \cdot \nabla v_0^{\mathrm{t}}}_{(\mathbf{B})}\big) \cdot \nabla\psi\Big) \circ (Id - tv_t)
\end{aligned}$$

For the term $(\mathbf{A})$ we have

$$\begin{aligned}
\mathbf{A} &= (\mathcal{I} - t\nabla v_t)^{\mathrm{t}} \cdot \nabla\phi \circ (Id - tv_t) \\
&= (\nabla(Id - tv_t))^{\mathrm{t}} \cdot \nabla\phi \circ (Id - tv_t) \\
&= \nabla\big(\phi \circ (Id - tv_t)\big) = \nabla\phi_t.
\end{aligned}$$

Thus to conclude we need only to show that the operator $(\mathbf{B})$ is 0. To this aim, observe that from $v_t \circ (Id + tv_0) = v_0$ we have

$$\begin{aligned}
\nabla v_t \circ (Id + tv_0) &= \nabla v_0 \cdot \big(\nabla(Id + tv_0)\big)^{-1} \\
&= \nabla v_0 \cdot \big(\nabla(Id - tv_t)\big) \circ (Id + tv_0) \\
&= \nabla v_0 - t\nabla v_0 \cdot \nabla v_t \circ (Id + tv_0).
\end{aligned}$$

The claim follows. ■

8.3 Points before the first conjugate

In this short section we want to point out that the existence of Jacobi fields on $\mathscr{P}_2(M)$ allows to give some quantitative estimate on the behavior of the map $\nu \mapsto \{\text{optimal transport map from } \mu \text{ to } \nu\}$, at least in some very special case.

Let (μ_t) as before, $\mu := \mu_0$ and define the time $FirstConj > 0$ as the first positive 0 of the function $f : [0,1] \to \mathbb{R}$ defined by

$$\begin{cases} f(0) &= 0, \\ f'(0) &= 1, \\ f''(t) &= -f(t)\Big(\mathcal{C}S^2 + 3\big(\mathrm{L}(v_t)\big)^2\Big). \end{cases}$$

If f has no zeros in $(0,1]$, then we put $FirstConj := 1$.

Now consider a Jacobi field $(\mathbf{J}_t)$ with initial conditions $\mathbf{J}_0 = 0$, $\frac{\boldsymbol{D}}{dt}\mathbf{J}_t|_{t=0} = u$ for some $u \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$. By standard comparison arguments based on the inequality

$$\left\| \frac{\boldsymbol{D}^2}{dt^2}\mathbf{J}_t \right\|_{\mu_t} = \|\mathbf{R}(v_t, \mathbf{J}_t)v_t\|_{\mu_t} \le \Big(\mathcal{C}S^2 + 3(\mathrm{L}(v_t))^2\Big)\|\mathbf{J}_t\|_{\mu_t}$$

we know that

$$\|\mathbf{J}_t\|_{\mu_t} \ge f(t)\|u\|_\mu, \qquad \forall t \in [0,1].$$

In particular, $\|\mathbf{J}_t\|_{\mu_t} > 0$ for every $t < FirstConj$. Fix $0 < t_0 < FirstConj$, define $\nu := \mu_{t_0}$, $\nu_s := \exp(t_0 v_0 + su)_\# \mu$. Assume also that $\exp(u)$ is an optimal map, so that $\exp(v_0 + su)$ is optimal as well for any $s \ge 0$. Let $\boldsymbol{\gamma}_s \in \mathcal{O}pt(\nu, \nu_s)$. Then by the result 7.8 on the differentiability of the exponential map, we know that $\mathrm{P}_\nu(\mathscr{B}(\frac{1}{s} \cdot \boldsymbol{\gamma}_s))$ weakly converges to $\mathbf{J}_{t_0}$ as $s \to 0$. This implies that

$$\begin{aligned} \varliminf_{s\to 0} \frac{W_2^2(\nu, \nu_s)}{s^2} &= \varliminf_{s\to 0} \int |\mathrm{v}|^2 d\left(\frac{1}{s} \cdot \boldsymbol{\gamma}_s\right) \ge \varliminf_{s\to 0} \left\| \mathscr{B}\Big(\frac{1}{s} \cdot \boldsymbol{\gamma}_s\Big) \right\|_\nu \\ &\ge \varliminf_{s\to 0} \left\| \mathrm{P}_\nu\left(\mathscr{B}\Big(\frac{1}{s} \cdot \boldsymbol{\gamma}_s\Big)\right) \right\|_\nu \ge \|\mathbf{J}_{t_0}\|_\nu^2 \ge \big(f(t_0)\big)^2 \|u\|_\mu^2. \end{aligned} \tag{8.6}$$

This inequality is a kind of reverse of the standard inequality

$$W_2(\nu, \tilde{\nu}) \le \|v - \tilde{v}\|_\mu,$$

valid on Riemannian manifolds with non negative sectional curvature, where $\nu = \exp(v)_\# \mu$ and $\tilde{\nu} = \exp(\tilde{v})_\# \mu$. Indeed, let $v_s := v + su$, so that $\exp(v_s)$ is an optimal transport map from μ to ν_s: we can rewrite (8.6) as

$$\varlimsup_{s\to 0} \frac{\|v_s - v\|_\mu}{W_2(\nu, \nu_s)} = \varlimsup_{s\to 0} \frac{s\|u\|_\mu}{W_2(\nu, \nu_s)} \le \frac{1}{f(t_0)}, \tag{8.7}$$

which is a sort of Lipschitz continuity in time of the optimal transport maps.

Let us underline that this is a very special situation: in order for the above estimate to hold, the measures μ, ν have to lie in the internal part of a geodesic and need to be sufficiently near w.r.t. to length of the geodesic itself. In general, as shown in [11], if (ν_s) is an absolutely continuous curve, the maximum regularity we can expect from $s \mapsto \{\text{optimal transport map from } \mu \text{ to } \nu_s\}$ is $\frac{1}{2}$-Hölder continuity.

9 Appendix

9.1 Density of regular curves

Here we prove that regular curves are dense in the set of transport couples. The notion of convergence related to this density is the following.

Definition 9.1 (L^2-convergence of transport couples) *Let (μ_t^n, v_t^n), $n \in \mathbb{N}$, and (μ_t, v_t) be transport couples satisfying*

$$\int_0^1 \|v_t\|_{\mu_t}^2 dt < \infty, \tag{9.1}$$

and similarly for (μ_t^n, v_t^n). We say that the sequence (μ_t^n, v_t^n) L^2-converges to (μ_t, v_t) if:

i) $W_2(\mu_t^n, \mu_t) \to 0$ as $n \to \infty$ uniformly on $t \in [0, 1]$,

ii) v_t^n strongly converges to v_t as $n \to \infty$ for a.e. $t \in [0, 1]$ in the sense of definition 1.8,

iii) $\overline{\lim}_{n\to\infty} \int_0^1 \|v_t^n\|_{\mu_t^n}^2 dt \leq \int_0^1 \|v_t\|_{\mu_t}^2 dt.$

This definition differs a bit from the analogous 2.5, indeed here we are requiring a stronger integrability condition on the norm of the vector fields, and, with the third condition, we are requiring that this kind of L^2 norm converges. We prefer to work with this new definition, as it fits more naturally into the theory. Once our density result will be proven w.r.t. L^2 convergence of transport couples, it is only a matter of standard reparametrization arguments to show that there is density also w.r.t. the convergence 2.5. We won't stress this point further.

The main result of this section is:

Theorem 9.2 (Density of regular curves) *Let (μ_t, v_t) be a transport couple such that* (9.1) *holds. Then there exists a sequence of regular curves (μ_t^n) such that:*

i) the sequence of transport couples (μ_t^n, v_t^n) L^2-converges to (μ_t, v_t), where (v_t^n) is the velocity vector field of (μ_t^n),

ii) for any n, t, μ_t^n is absolutely continuous w.r.t. the volume measure with and $\mathrm{supp}(\mu_t^n) \subset K^n$ for some compact set K^n independent on t. Also, the map $(x, t) \mapsto \frac{d\mu_t^n}{d\mathrm{vol}}(x)$ is C^∞, *iii)* $v_t^n = \nabla \varphi_t^n$ *with* $\varphi_t^n \in C_c^\infty(M)$ *for every* n, t*. Also,* $\mathrm{supp}(\varphi_t^n) \subset K^n$ *and the map* $(x,t) \mapsto \varphi_t^n(x)$ *is* C^∞*.*

Remark 9.3 There is a great freedom in the choice of the support of the measures μ_t^n: it may be chosen compact (as in the statement of the theorem), or equal to the whole M. Also, if the given transport couple (μ_t, v_t) is made of measures concentrated on the same compact K, the measures μ_t^n may be chosen with support in some compact K' independent on n, t. ■

Remark 9.4 Observe that we can approximate any transport couple, and not only those for which the vector field (v_t) is tangent. That is, we can produce convergence to v_t with gradients of smooth functions, regardless of the tangency of the v_t's themselves. In view of the discussion made in section 4.4 this is not that surprising. ■

This result is slightly different from the one obtained in [1, 10] for the case $M = \mathbb{R}^d$, where it was proven that the measures μ_t^n can be chosen such that $\overline{\lim}_n \left\| \frac{d\mu_t^n}{d\mathcal{L}^d} \right\|_\infty \leq \left\| \frac{d\mu_t}{d\mathcal{L}^d} \right\|_\infty$ for a.e. t, where $\| \frac{d\mu}{d\mathcal{L}^d} \|_\infty$ is the essential supremum of the density of the absolutely continuous measure μ, but giving no informations on the smoothness of the density nor on the supports. This proof for the case $M = \mathbb{R}^d$, found by Ambrosio, strongly relies on the geometrical properties of $\mathbb{R}^d$, and it is unclear - to the author - whether an analogous result holds for general manifolds or not.

Our strategy for the proof of the above theorem is the following:

- we recall that for any transport couple there exists a (non unique in general) associated measure on the space $\mathscr{P}(\mathrm{AC}([0,1], \mathrm{M}))$ of absolutely continuous curves in M,
- we prove that convergence of transport couples is implied by convergence of the associated measures in $\mathscr{P}_2(\mathrm{AC}_2([0,1], M))$,
- using such implication, we first approximate the given transport couple with transport couples for which the underlying measures have finite support: for these kind of transport couples it will be easy to check regularity
- now that the problem is reduced to approximating a transport couple whose measures have finite support, we can 'put our hands in' and, by a standard smoothening argument, get the conclusion.

Unfortunately, the development of this plan is a bit lengthy in terms of vocabulary that needs to be introduced.

We start recalling the following important result which we state without proof. It was proved for the case $M = \mathbb{R}^d$ in [2]: the generalization to the case of general Riemannian manifolds follows by Nash's embedding theorem and presents no difficulties, we won't give the details. We denote by $\mathrm{e_t}$ the evaluation map defined, for any curve $\gamma : [0,1] \to M$, as $\mathrm{e_t}(\gamma) := \gamma(t)$.

Theorem 9.5 (Probabilistic interpretation of continuity equation) *Let* (μ_t, v_t) *be a transport couple. Then there exists a Borel probability measure* $\boldsymbol{\mu}$ *on* $\mathrm{AC}([0,1], \mathrm{M})$ *such that:*

i) $\boldsymbol{\mu}$ *is concentrated on the set of curves* γ *which are solution of the ODE* $\gamma'(t) = v_t(\gamma(t))$ *for a.e.* $t \in [0,1]$*,*

ii) $\mu_t = \mathrm{e}_{\mathrm{t}\#}\boldsymbol{\mu}$*.*

Remark 9.6 In general the measure $\boldsymbol{\mu}$ is not unique. Consider for instance the curve $[-1,1] \ni t \mapsto \mu_t := \frac{1}{2}(\delta_t + \delta_{-t}) \in \mathscr{P}_2(\mathbb{R})$, and let (v_t) its velocity vector field. Then a possible choice of $\boldsymbol{\mu}$ is the measure $\frac{1}{2}(\delta_{\gamma(\cdot)} + \delta_{\tilde{\gamma}(\cdot)})$, where $\gamma, \tilde{\gamma} : [-1,1] \to \mathbb{R}$ are defined by

$$\gamma(t) := t,$$
$$\tilde{\gamma}(t) := -t.$$

A different $\boldsymbol{\mu}$ associated to (μ_t, v_t) is given by the same formula above with:

$$\gamma(t) := \begin{cases} t & \text{if } t \leq 0, \\ -t & \text{if } t \geq 0, \end{cases} \qquad \tilde{\gamma}(t) := \begin{cases} -t & \text{if } t \leq 0, \\ t & \text{if } t \geq 0, \end{cases}.$$

If the curve (μ_t) is regular and (v_t) is its velocity vector field, then the uniquen ■

We will denote by $\mathrm{AC}_2([0,1],M)$ the set of absolutely continuous curves γ from $[0,1]$ to M satisfying $\int_0^1 |\gamma'|^2(t)dt < \infty$ endowed with the distance

$$\overline{d}(\gamma_1(\cdot), \gamma_2(\cdot)) := \sqrt{\left(\sup_{t\in[0,1]} \mathrm{d}(\gamma_1(t), \gamma_2(t))\right)^2 + \int_0^1 \mathrm{D}^2\Big(\gamma_1'(t), \gamma_2'(t)\Big)dt}.$$

It is easy to check that $(\mathrm{AC}_2([0,1],M), \overline{d})$ is complete and separable. Also, if $\boldsymbol{\mu}$ is a measure associated to a transport couple (μ_t, v_t) satisfying (9.1), the identity

$$\int\int_0^1 |\gamma'|^2(t)dt\boldsymbol{\mu}(\gamma) = \int_0^1 \|v_t\|^2_{\mu_t}dt < \infty,$$

gives that $\boldsymbol{\mu}$ is concentrated on $\mathrm{AC}_2([0,1],M)$.

Now recall that the Wasserstein distance is well defined on all complete and separable metric spaces (X,d). Choose $\overline{x} \in X$ and define

$$\mathscr{P}_2(X) := \left\{\mu \in \mathscr{P}(X) \ : \ \int d^2(x,\overline{x})d\mu(x) < \infty\right\},$$

it is clear that this set does not depend on the chosen point $\overline{x}$. The set $\mathscr{P}_2(X)$ is endowed with the distance

$$W_2(\mu,\nu) := \sqrt{\inf \int d^2(x,y)d\boldsymbol{\gamma}(x,y)},$$

where the infimum is taken over all $\boldsymbol{\gamma}$ such that $\pi^1_\#\boldsymbol{\gamma} = \mu$ and $\pi^2_\#\boldsymbol{\gamma} = \nu$, where π^1 and π^2 are the projections onto the first and second coordinate respectively. The following proposition is well known, we skip the proof.

Proposition 9.7 *Let (X, d) be a complete separable metric space and $\mu \in \mathscr{P}_2(X)$. Then there exists a sequence $(\mu_n) \subset \mathscr{P}_2(X)$ such that $W_2(\mu_n, \mu) \to 0$ as $n \to \infty$ and the support of μ_n is made of exactly n points. Also, the following statements are equivalent:*

- $W_2(\mu_n, \mu) \to 0$,

- $\int \varphi d\mu_n \to \int \varphi d\mu$ *for every real valued, bounded and continuous function* φ *and* $\lim_{n\to\infty} W_2(\mu_n, \nu) = W_2(\mu, \nu)$, *for some* $\nu \in \mathscr{P}(X)$.

- $\int \varphi d\mu_n \to \int \varphi d\mu$ *for every real valued, continuous function* φ *with linear growth, i.e. satisfying*
$$|\varphi(x)| \leq \frac{C}{1 + d(x, x_0)}, \qquad \forall x \in X$$
for some $x_0 \in X$, *and* $\lim_{n\to\infty} W_2(\mu_n, \nu) = W_2(\mu, \nu)$, *for some* $\nu \in \mathscr{P}(X)$.

- $\int \varphi d\mu_n \to \int \varphi d\mu$ *for every real valued, continuous function* φ *with quadratic growth, i.e. satisfying*
$$|\varphi(x)| \leq \frac{C}{1 + d^2(x, x_0)}, \qquad \forall x \in X$$
for some $x_0 \in X$,

In particular, the space $\mathscr{P}_2(\mathrm{AC}_2([0,1], M))$ is well defined. To characterize the measures in such space, we define the *cost* $C(\boldsymbol{\mu})$ of a measure $\boldsymbol{\mu} \in \mathscr{P}(\mathrm{AC}([0,1], \mathrm{M}))$ by:
$$C(\boldsymbol{\mu}) := \int \int_0^1 |\gamma'|^2(t) dt \, d\boldsymbol{\mu}(\gamma).$$
It holds the following result.

Proposition 9.8 *Let $\boldsymbol{\mu} \in \mathscr{P}(\mathrm{AC}([0,1], \mathrm{M}))$. Then $\boldsymbol{\mu} \in \mathscr{P}_2(\mathrm{AC}_2([0,1], M))$ if and only if $C(\boldsymbol{\mu}) < \infty$ and $\mathrm{e}_{0\#}\boldsymbol{\mu} \in \mathscr{P}_2(M)$. In particular, any measure $\boldsymbol{\mu}$ associated via theorem 9.5 to a transport couple (μ_t, v_t) satisfying (9.1) belongs to $\mathscr{P}_2(\mathrm{AC}_2([0,1], M))$.*

Proof We start with the 'only if' part of the statement. We know that
$$\int \left(\left(\sup_{t\in[0,1]} \mathrm{d}(\gamma(t), \overline{x}) \right)^2 + \int_0^1 \mathrm{D}^2(\gamma'(t), 0_{\overline{x}}) dt \right) d\boldsymbol{\mu}(\gamma) < \infty,$$
where $\overline{x}$ is some chosen point of the manifold and $0_{\overline{x}} \in T_{\overline{x}}M$ is the 0 vector, and we want to prove that
$$\int \int_0^1 |\gamma'|^2(t) dt d\boldsymbol{\mu}(\gamma) < \infty,$$
$$\int \mathrm{d}^2(\gamma(0), \overline{x}) d\boldsymbol{\mu}(\gamma) < \infty.$$

To prove this it is enough to observe that from (1.6) it follows the inequality

$$\mathrm{D}(\gamma'(t), 0_{\overline{x}}) \geq |\gamma'(t)|.$$

For the converse implication, start that recalling that (1.7) tells that

$$\mathrm{D}^2(\gamma'(t), 0_{\overline{x}}) \leq \mathrm{d}^2(\gamma(t), \overline{x}) + |\gamma'(t)|^2.$$

To conclude, observe that it holds

$$\mathrm{d}^2(\gamma(t), \overline{x}) \leq \left(\mathrm{d}(\gamma(0), \overline{x}) + \int_0^t |\gamma'|(t)dt\right)^2 \leq 2\mathrm{d}^2(\gamma(0), \overline{x}) + 2\int_0^1 |\gamma'|^2(t)dt$$

□

Now that the set $\mathscr{P}_2(\mathrm{AC}_2([0,1], M))$ is identified, we pass to the description of the relation between convergence of transport couples and of the corresponding measures in $\mathscr{P}_2(\mathrm{AC}_2([0,1], M))$. To do so, start observing that (TM, D) is a complete separable metric space, and thus we may consider the space $(\mathscr{P}_2(TM), W_2)$, where here W_2 is the Wasserstein distance associated to D. There is a simple relation between convergence of maps in the sense of 1.8 and convergence in $\mathscr{P}_2(TM)$:

Proposition 9.9 *Let $n \mapsto \mu_n \in \mathscr{P}_2(M)$ a sequence W_2-converging to some μ, $u_n \in L^2_{\mu_n}$ for any $n \in \mathbb{N}$ and $u \in L^2_\mu$. Then u_n strongly converges to u in the sense of 1.8 if and only if the sequence of plans $\boldsymbol{\gamma}_n := (Id, u_n)_\# \mu_n \in \mathscr{P}_2(TM)$ converges to $\boldsymbol{\gamma} := (Id, u)_\# \mu$ in $\mathscr{P}_2(TM)$.*

Proof Start assuming that $W_2(\boldsymbol{\gamma}_n, \boldsymbol{\gamma}) \to 0$. We will use proposition 9.7. Fix $\xi \in \mathcal{V}(M)$ and consider the test function $TM \ni (x, \mathrm{v}) \mapsto \langle \xi(x), \mathrm{v} \rangle \in \mathbb{R}$ (observe that it has linear growth) to get

$$\lim_{n\to\infty} \int \langle \xi(x), u_n(x) \rangle \, d\mu_n(x) = \int \langle \xi(x), u_n(x) \rangle \, d\mu(x).$$

Now consider the measure $\boldsymbol{\sigma} \in \mathscr{P}_2(TM)$ defined by $\boldsymbol{\sigma} := (Id, 0)_\# \mu$. Observe that $W_2^2(\boldsymbol{\gamma}, \boldsymbol{\sigma}) = \|v\|^2_\mu$. Similarly, define $\boldsymbol{\sigma}_n := (Id, 0)_\# \mu_n$. Obviously $W_2(\boldsymbol{\sigma}_n, \boldsymbol{\sigma}) \to 0$ as $n \to \infty$, therefore we obtain

$$\|v\|_\mu = W_2(\boldsymbol{\sigma}, \boldsymbol{\gamma}) = \lim_{n\to\infty} W_2(\boldsymbol{\sigma}, \boldsymbol{\gamma}_n) = \lim_{n\to\infty} W_2(\boldsymbol{\sigma}_n, \boldsymbol{\gamma}_n) = \lim_{n\to\infty} \|v_n\|_{\mu_n}.$$

We pass to the converse implication. It is clear that from the uniform bound on the norms of v_n it follows the tightness of the family $\{\boldsymbol{\gamma}_n\}$. Thus, up to pass to a subsequence, not relabeled, we may assume that $n \mapsto \boldsymbol{\gamma}_n$ converges to some $\tilde{\boldsymbol{\gamma}}$ in duality with continuous and bounded functions. The functional $\boldsymbol{\sigma} \mapsto \int |\mathrm{v}|^2 d\boldsymbol{\sigma}(x, \mathrm{v})$ is lower semicontinuous w.r.t. convergence in duality with continuous and bounded functions, so we get

$$\int |\mathrm{v}|^2 d\tilde{\boldsymbol{\gamma}}(x, \mathrm{v}) \leq \varliminf_{n\to\infty} \int |\mathrm{v}|^2 d\tilde{\boldsymbol{\gamma}}_n(x, \mathrm{v}) = \varliminf_{n\to\infty} \|v_n\|^2_{\mu_n}.$$

By the continuity of the projection on M we have $\pi^M_\# \tilde{\boldsymbol{\gamma}} = \mu$. Define the vector field $\tilde{v}$ as

$$\tilde{v}(x) := \int \mathrm{v} d\tilde{\boldsymbol{\gamma}}(x, \mathrm{v}),$$

from the strict convexity of the L^2 norm it follows that

$$\|\tilde{v}\|^2_\mu \leq \int |\mathrm{v}|^2 d\tilde{\boldsymbol{\gamma}}(x, \mathrm{v}),$$

and that equality holds if and only if $\tilde{\boldsymbol{\gamma}} = (Id, \tilde{v})_\# \mu$. We claim that $\tilde{v} = v$. Indeed for any $\xi \in \mathcal{V}(M)$ we have

$$\lim_{n\to\infty} \langle v_n, \xi \rangle_{\mu_n} = \langle v, \xi \rangle_\mu ,$$

$$\lim_{n\to\infty} \langle v_n, \xi \rangle_{\mu_n} = \lim_{n\to\infty} \int \langle \mathrm{v}, \xi(x) \rangle \, d\boldsymbol{\gamma}_n(x, \mathrm{v}) = \int \langle \mathrm{v}, \xi(x) \rangle \, d\tilde{\boldsymbol{\gamma}}(x, \mathrm{v}) = \langle \tilde{v}, \xi \rangle_\mu .$$

From the fact that $\|v_n\|_{\mu_n} \to \|v\|_\mu$ we deduce that in the chain of inequalities

$$\|v\|^2_\mu \leq \int |\mathrm{v}|^2 d\tilde{\gamma}(x, \mathrm{v}) \leq \varliminf_{n\to\infty} | \int \mathrm{v}|^2 d\tilde{\gamma}_n(x, \mathrm{v}) = \varliminf_{n\to\infty} \|v_n\|^2_{\mu_n} = \|v\|^2_\mu,$$

all inequalities are equality. In particular, we obtain that

$$\tilde{\gamma} = (Id, v)_\# \mu.$$

Thus we proved that $\boldsymbol{\gamma}_n$ converges to $\boldsymbol{\gamma}$ in duality with continuous functions with compact support. The fact that $\lim_{n\to\infty} W_2(\boldsymbol{\gamma}_n, (Id, 0)_\# \mu) = W_2(\boldsymbol{\gamma}, (Id, 0)_\# \mu)$ follows by arguments similar to those used above. □

Now observe that to each measure $\boldsymbol{\mu}$ in $\mathscr{P}_2(\mathrm{AC}([0,1], \mathrm{M}))$ we can associate the plans $\boldsymbol{\gamma}^{\boldsymbol{\mu}}_t \in \mathscr{P}_2(TM)$ defined by $\boldsymbol{\gamma}^{\boldsymbol{\mu}}_t := (\mathrm{e_t}, \mathrm{d_t})_\# \boldsymbol{\mu}$, where $(\mathrm{e_t}, \mathrm{d_t})(\gamma) := (\gamma(t), \gamma'(t))$. It is clear that if the measure $\boldsymbol{\mu}$ is associated to a certain transport couple (μ_t, v_t) via theorem 9.5, the plans $\boldsymbol{\gamma}^{\boldsymbol{\mu}}_t$ are given by $\boldsymbol{\gamma}^{\boldsymbol{\mu}}_t = (Id, v_t)_\# \mu_t$ for a.e. $t \in [0, 1]$.

The desired relation between convergence of transport couples and of associated measures in $\mathscr{P}_2(\mathrm{AC}_2([0,1], M))$ is based on the following proposition:

Proposition 9.10 *Let* $\boldsymbol{\mu}^n, \boldsymbol{\mu} \in \mathscr{P}_2(\mathrm{AC}_2([0,1], M))$, $n \in \mathbb{N}$ *and assume that* $\lim_{n\to\infty} W_2(\boldsymbol{\mu}, \boldsymbol{\mu}^n) = 0$. *Then it holds*

$$\lim_{n\to\infty} C(\boldsymbol{\mu}^n) = C(\boldsymbol{\mu}),$$

$$\lim_{n\to\infty} \int_0^1 W^2_2(\boldsymbol{\gamma}^{\boldsymbol{\mu}^n}_t, \boldsymbol{\gamma}^{\boldsymbol{\mu}}_t) dt = 0.$$

Proof To prove the first statement, observe that the map $\gamma \mapsto c(\boldsymbol{\gamma}) := \int_0^1 |\gamma'(t)|^2 dt$ is continuous from $(\mathrm{AC}_2([0,1], M), \overline{d})$ to $\mathbb{R}$ with quadratic growth. Therefore the map $\boldsymbol{\mu} \mapsto \int c(\gamma) d\boldsymbol{\mu}(\gamma) = C(\boldsymbol{\mu})$ from $\mathscr{P}_2(\mathrm{AC}_2([0,1], M))$ to $\mathbb{R}$ is continuous by proposition 9.7.

For the second, let $\boldsymbol{\alpha}^n$ be an optimal plan from $\boldsymbol{\mu}$ to $\boldsymbol{\mu}^n$, so that it holds

$$W_2^2(\boldsymbol{\mu}, \boldsymbol{\mu}^n) = \int \left(\Big(\sup_{t\in[0,1]} \{ \mathrm{d}(\gamma_1(t), \gamma_2(t)) \} \Big)^2 + \int_0^1 \mathrm{D}^2(\gamma_1'(t), \gamma_2'(t)) dt \right) d\boldsymbol{\alpha}^n(\gamma_1, \gamma_2).$$

For any $t \in [0,1]$, define the map $I_t : [\mathrm{AC}_2([0,1], M)]^2 \to [TM]^2$ by $I_t(\gamma_1, \gamma_2) := \big((\gamma_1(t), \gamma_1'(t)), (\gamma_2(t), \gamma_2'(t))\big)$ and the plan $\boldsymbol{\alpha}_t^n := (I_t)_\# \boldsymbol{\alpha}^n$. For any $t \in [0,1]$, the plan $\boldsymbol{\alpha}_t^n$ is admissible for the couple $(\boldsymbol{\gamma}_t^{\boldsymbol{\mu}}, \boldsymbol{\gamma}_t^{\boldsymbol{\mu}^n})$, thus the conclusion follows from

$$\begin{aligned}
W_2^2(\boldsymbol{\mu}, \boldsymbol{\mu}^n) &= \int \left(\Big(\sup_{t\in[0,1]} \{ \mathrm{d}(\gamma_1(t), \gamma_2(t)) \} \Big)^2 + \int_0^1 \mathrm{D}^2(\gamma_1'(t), \gamma_2'(t)) dt \right) d\boldsymbol{\alpha}^n(\gamma_1, \gamma_2) \\
&\geq \int \int_0^1 \mathrm{D}^2(\gamma_1'(t), \gamma_2'(t)) dt \, d\boldsymbol{\alpha}^n(\gamma_1, \gamma_2) \\
&= \int_0^1 \int \mathrm{D}^2(\mathrm{v}_1, \mathrm{v}_2) d\boldsymbol{\alpha}_t^n\big((x_1, \mathrm{v}_1), (x_2, \mathrm{v}_2)\big) \\
&\geq \int_0^1 W_2^2(\boldsymbol{\gamma}_t^{\boldsymbol{\mu}}, \boldsymbol{\gamma}_t^{\boldsymbol{\mu}^n}) dt.
\end{aligned}$$

□

Remark 9.11 The fact that in general there is not a unique $\boldsymbol{\mu}$ associated to a given transport couple, shows that the converse implication in this proposition is not true. ■

Corollary 9.12 *Let $(\mu_t^n, v_t^n), (\mu_t, v_t)$ be transport couples and $\boldsymbol{\mu}^n, \boldsymbol{\mu}$ be measures associated to them via theorem 9.5. Assume that $W_2(\boldsymbol{\mu}^n, \boldsymbol{\mu}) \to 0$ as $n \to \infty$. Then there is a sequence $k \mapsto n_k$ such that $k \mapsto (\mu_t^{n_k}, v_t^{n_k})$ L^2converges to (μ_t, v_t) as $k \to \infty$.*

Proof It is a direct consequence of the previous proposition, of the characterization of convergence of maps given in proposition 9.9 and of the fact that a sequence of L^2 functions converging to 0 admits a subsequence converging almost everywhere to 0. □

We are now ready to prove our main result

Proof of theorem 9.2 Let $\boldsymbol{\mu}$ be a measure on $\mathrm{AC}_2([0,1], M)$ associated to (μ_t, v_t) via theorem 9.5. By proposition 9.7 we know that there exists a sequence of measures $n \mapsto \boldsymbol{\nu}^n$ such that $W_2(\boldsymbol{\nu}^n, \boldsymbol{\mu}) \to 0$ as $n \to \infty$ and such that the support of $\boldsymbol{\nu}^n$ is made of exactly n distinct points (i.e. curves). For every $n \in \mathbb{N}$, let $\gamma_i^n : [0,1] \to M$, $i = 1, \ldots, n$ be the curves in the support of $\boldsymbol{\nu}^n$. Up to changing a bit these curves, we may assume that they are C^∞ and that

$$\gamma_i^n(t) \neq \gamma_j^n(t), \qquad \forall i \neq j, \forall t \in [0,1].$$

Working in chart, it is not hard to see that for every $n \in \mathbb{N}$ there exists a C^∞ map $(x,t) \mapsto \varphi_t^n(x)$ satisfying $(\nabla \varphi_t^n)(\gamma_i^n(t)) = (\gamma_i^n)'(t)$ for any $i = 1, \ldots, n$, and $\varphi_t^n(x) = 0$ if $x \notin \tilde{K}^n$, where $\tilde{K}^n \subset M$ is a compact set whose interior contains the range of all the γ_i^n's. By smoothness and compactness it holds $L^n := \sup_{t\in[0,1]} \mathrm{L}(\nabla\varphi_t^n) < \infty$ for any $n \in \mathbb{N}$. Apply part (ii) of theorem 2.6 with $v_t^n := \nabla\varphi_t^n$ to get the existence of C^∞ maps $\mathbf{T}^n(t,s,\cdot)$ from M into itself satisfying equations (2.2). From the bound (2.4a) we get $\sup_{t\in[0,1]} \mathrm{Lip}(\mathbf{T}^n(0,t,\cdot)) \le e^{L^n} < \infty$.

Let $I^n : M \mapsto \mathrm{AC}_2([0,1],M)$ be the map which associate to each $x \in M$ the curve $t \mapsto I^n(x)(t) := \mathbf{T}^n(0,t,x)$. From lemma 9.14 below, we have the estimate:

$$\begin{aligned}
\overline{d}^2(I^n(x),I^n(y)) &= \sup_{t\in[0,1]} \mathrm{d}^2\Big(I^n(x)(t), I^n(y)(t)\Big) + \int_0^1 \mathrm{D}^2\Big(\big(I^n(x)\big)'(t), \big(I^n(y)\big)'(t)\Big)dt,\\
&\le \sup_{t\in[0,1]} \mathrm{d}^2\Big(I^n(x)(t), I^n(y)(t)\Big) + \sup_{t\in[0,1]} \mathrm{d}^2\Big(I^n(x)(t), I^n(y)(t)\Big)\left(1+\sup_{t\in[0,1]}(\mathrm{L}(\varphi_t^n))^2\right)\\
&\le A^n \mathrm{d}^2(x,y),
\end{aligned} \tag{9.2}$$

where $A^n := e^{2L^n}\Big(2+(L^n)^2\Big)$.

Now choose $\mu_0^n \in \mathscr{P}_2(M)$ absolutely continuous w.r.t. the volume measure, with smooth density and compact support, satisfying $W_2(\mu_0^n, \mathrm{e}_{0\#}\boldsymbol{\nu}^n) < (nA^n)^{-1}$. Define $\boldsymbol{\mu}^n := I^n_{\#}\mu_0^n$ and $\mu_t^n := \mathrm{e_t}\#\boldsymbol{\mu}^n = \mathbf{T}(0,t,\cdot)_{\#}\mu_0^n$. Given that the maps $\mathbf{T}^n(0,t,\cdot)$ are C^∞ with C^∞ inverse we have that μ_t^n is absolutely continuous with smooth density and compact support for any n,t Also, for any n the map $(x,t) \mapsto \frac{d\mu_t^n}{d\mathrm{vol}}(x)$ is C^∞ and the support of μ_t^n is contained in the compact set K^n defined by

$$K^n := \bigcup_{t\in[0,1]} \mathbf{T}^n(0,t,\mathrm{supp}(\mu_0^n)).$$

Furthermore, it is clear that $t \mapsto \mu_t^n \in \mathscr{P}_2(M)$ is absolutely continuous and satisfies the continuity equation

$$\frac{d}{dt}\mu_t^n + \nabla\cdot(\mu_t^n \nabla\varphi_t^n) = 0,$$

so that from the uniform bound on the Lipschitz constant of $\nabla\varphi_t^n$ it follows that (μ_t^n) is regular. It is also clear that $\boldsymbol{\mu}^n$ is the unique measure in $\mathscr{P}_2(\mathrm{AC}_2([0,1],M))$ associated to the transport couple $(\mu_t^n, \nabla\varphi_t^n)$.

Define $J^n : M^2 \to [\mathrm{AC}_2([0,1],M)]^2$ as $J^n(x,y) := (I^n(x), I^n(y))$ and choose an optimal plan $\boldsymbol{\gamma}^n$ from μ_0^n to $\mathrm{e}_{0\#}\boldsymbol{\nu}^n$. By construction, the plan $J^n_{\#}\boldsymbol{\gamma}^n$ is admissible for the couple $(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n)$. Integrating inequality (9.2) over $J^n_{\#}\boldsymbol{\gamma}^n$ we get

$$\begin{aligned}
W_2^2(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n) &\le \int \overline{d}(\gamma_1,\gamma_2)\, d\big(J^n_{\#}\boldsymbol{\gamma}^n\big)(\gamma_1,\gamma_2)\\
&= \int \overline{d}^2(I^n(x), I^n(y))d\boldsymbol{\gamma}^n(x,y)\\
&\le W_2^2(\mu_0^n, \mathrm{e}_{0\#}\boldsymbol{\nu}^n)A_n\\
&\le \frac{1}{n}.
\end{aligned}$$

Thus it holds

$$W_2(\boldsymbol{\mu}^n, \boldsymbol{\mu}) \leq W_2(\boldsymbol{\mu}^n, \boldsymbol{\nu}^n) + W_2(\boldsymbol{\nu}_n, \boldsymbol{\mu}) \to 0,$$

as $n \to \infty$. By corollary 9.12 we get the thesis. □

Remark 9.13 If we want to produce an approximating sequence for which the measures μ_t^n have positive density everywhere, it is enough to choose the measures μ_0^n absolutely continuous, and with density smooth and everywhere positive. Then the proof goes on analogously.

Similarly, if our limit transport couple is made of measures whose supports are all contained in the same compact set $K \subset M$, we can produce an approximating sequence for which all the measures μ_t^n are concentrated on some compact set K' in the following way. Find a connected open set $\Omega \supset K$ with compact closure, and apply proposition 9.7 to find the approximating sequence $\boldsymbol{\nu}^n$ such that $\text{supp}(\text{e}_{\text{t}\#}\boldsymbol{\nu}^n) \subset \Omega$ for every n, t. Then choose the measures μ_0^n with support contained in Ω and proceed as before. ■

Lemma 9.14 *Let $\xi \in \mathcal{V}(M)$ and $x, y \in M$. Then it holds*

$$\text{D}^2(\xi(x), \xi(y)) \leq \big(1 + (\text{L}(\xi))^2\big)\text{d}^2(x, y). \tag{9.3}$$

Proof Estimating $\text{D}(\xi(x), \xi(y))$ from above using a geodetic $\gamma : [0, 1] \to M$ between x and y in equation (1.6), we get

$$\text{D}^2(\xi(x), \xi(y)) \leq \text{d}^2(x, y) + |T_0^1(\xi(x)) - \xi(y)|^2, \tag{9.4}$$

where T_s^t is the parallel transport map from $\gamma(s)$ to $\gamma(t)$ along γ. Now observe that

$$\begin{aligned}\frac{d}{dt}\Big|T_t^1\Big(\xi\big(\gamma(t)\big)\Big) - \xi(y)\Big|^2 &= -2\Big\langle T_t^1\Big(\xi\big(\gamma(t)\big)\Big) - \xi(y), \nabla_{\gamma'(t)}\xi\big(\gamma(t)\big)\Big\rangle \\ &= -2\Big\langle T_t^1\Big(\xi\big(\gamma(t)\big)\Big) - \xi(y), \nabla\xi\big(\gamma(t)\big) \cdot \gamma'(t)\Big\rangle \\ &\leq 2\Big|T_t^1\Big(\xi\big(\gamma(t)\big)\Big) - \xi(y)\Big|\text{L}(\xi)|\gamma'(t)|,\end{aligned}$$

from which it follows

$$\frac{d}{dt}\Big|T_t^1\Big(\xi\big(\gamma(t)\big)\Big) - \xi(y)\Big| \leq |\text{L}(\xi)|\gamma'(t)|.$$

Integrating this inequality over $[0, 1]$ we get

$$\big|T_0^1\big(\xi(x)\big) - \xi(y)\big| \leq \text{L}(\xi)\int_0^1 |\gamma'(t)|dt = \text{L}(\xi)\text{d}(x, y).$$

Plugging this into (9.4) we get the result. □

9.2 C^1 curves

In this work, it often happened to consider absolutely continuous curves (μ_t) having a velocity vector field, generally defined only for a.e. t, which admits a representative (v_t) continuous and tangent on the whole $[0,1]$. Here we want to show that this notion, a priori purely dependent on the topology of convergence of maps, is actually, not surprisingly, linked to the analytical and geometrical properties of (μ_t).

Definition 9.15 (C^1 curves) *Let $(\mu_t) \subset \mathscr{P}_2(M)$ be an absolutely continuous curve. We say that (μ_t) is C^1 if there exists a choice of the velocity vector field (v_t) for* every $t \in [0,1]$ *which is continuous w.r.t. strong convergence of maps and tangent for every $t \in [0,1]$.*

Lemma 9.16 *Let $\mu \in \mathcal{P}_c(M)$, $v \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$ and define $\mu_t := \mathbf{exp}_\mu(tv) = (\exp(tv))_\#\mu$. Then it holds*

$$\lim_{t \downarrow 0} \frac{W_2(\mu_t, \mu)}{t} = \|v\|_\mu.$$

If the manifold M has non negative sectional curvature, then the result is true for any $\mu \in \mathscr{P}_2(M)$.

Proof Since $(Id, \exp(tv))_\#$ is an admissible plan for the couple μ, μ_t, we know that it holds

$$\frac{W_2(\mu_t, \mu)}{t} \le t\|v\|_\mu, \qquad \forall t \in [0,1].$$

Suppose $v = \nabla\varphi$, where $\varphi \in C_c^\infty(M)$. In this case the thesis follows by proposition 6.1.

To pass to the limit in $\nabla\varphi$, observe that given $v, w \in L^2_\mu$ such that $\sup|v(x)|, \sup|w(x)| \le S$ (here the supremum is an essential supremum w.r.t. μ), equation (1.10) gives

$$W_2\Big(\mathbf{exp}_\mu(v), \mathbf{exp}_\mu(w)\Big) \le \|v - w\|_\mu \frac{\sinh\Big(S\sqrt{\mathcal{C}(\mathrm{supp}(\mu))}\Big)}{S\sqrt{\mathcal{C}(\mathrm{supp}(\mu))}}$$

Now, for every $v \in \mathrm{Tan}_\mu(\mathscr{P}_2(M))$, define the truncation $T_N(v) \in L^2_\mu$ as

$$T_N(v)(x) := \begin{cases} v(x) & \text{if } |v(x)| \le N, \\ \dfrac{Nv(x)}{|v(x)|} & \text{if } |v(x)| \ge N. \end{cases}$$

It clearly holds

$$W_2\Big(\mathbf{exp}_\mu(tv), \mathbf{exp}_\mu(tT_N(v))\Big) \le t\|v - T_N(v)\|_\mu, \qquad \forall t \in [0,1].$$

We have

$$
\begin{aligned}
W_2\Big(\mathbf{exp}_\mu(tv),&\mathbf{exp}_\mu(t\nabla\varphi)\Big)\\
&\le W_2\Big(\mathbf{exp}_\mu(tv),\mathbf{exp}_\mu(tT_N(v))\Big)+W_2\Big(\mathbf{exp}_\mu(tT_N(v)),\mathbf{exp}_\mu(tT_N(\nabla\varphi))\Big)\\
&\quad +W_2\Big(\mathbf{exp}_\mu(tT_N(\nabla\varphi)),\mathbf{exp}_\mu(t\nabla\varphi)\Big),\\
&\le t\|v-T_N(v)\|_\mu+t\|\nabla\varphi-T_N(\nabla\varphi)\|_\mu\\
&\quad +t\|T_N(v)-T_N(\nabla\varphi)\|\frac{\sinh(N\sqrt{\mathcal{C}(\mathrm{supp}(\mu))})}{N\sqrt{\mathcal{C}(\mathrm{supp}(\mu))}}.
\end{aligned}
$$

And therefore

$$
\begin{aligned}
\frac{W_2(\mu_t,\mu)}{t}&\ge\frac{W_2(\mathbf{exp}_\mu(t\nabla\varphi),\mu)}{t}-\frac{W_2\big(\mu_t,\mathbf{exp}_\mu(t\nabla\varphi)\big)}{t}\\
&\ge\frac{W_2(\mathbf{exp}_\mu(t\nabla\varphi),\mu)}{t}-\|v-T_N(v)\|_\mu-\|\nabla\varphi-T_N(\nabla\varphi)\|_\mu\\
&\quad -\|T_N(v)-T_N(\nabla\varphi)\|\frac{\sinh(N\sqrt{\mathcal{C}(\mathrm{supp}(\mu))})}{N\sqrt{\mathcal{C}(\mathrm{supp}(\mu))}}.
\end{aligned}
$$

Letting first $t\to 0$, then $\nabla\varphi\to v$ in $\mathrm{Tan}_\mu(\mathscr{P}_2(M))$ and then $N\to\infty$ we conclude.

The conclusion for the case of non negative sectional curvature follows by equation (1.9), which shows that in this case $W_2(\mathbf{exp}_\mu(v),\mathbf{exp}_\mu(w))\le\|v-w\|_\mu$. The thesis follows as above.

□

Lemma 9.17 *Let $\mu_n,\mu\in\mathscr{P}_2(M)$ be such that $W_2(\mu_n,\mu)\to 0$ as $n\to\infty$ and $u_n\in L^2_{\mu_n}$, $u\in L^2_\mu$. Then u_n strongly converges to u if and only if for any choice of plans $\boldsymbol{\gamma}_n\in\mathcal{Adm}(\mu,\mu_n)$ (not necessarily optimal) satisfying $\lim_n\int \mathrm{d}^2(x,y)d\boldsymbol{\gamma}_n=0$ it holds*

$$
\lim_{n\to\infty}\int \mathrm{D}^2(u(x),u_n(y))d\boldsymbol{\gamma}_n(x,y)=0. \tag{9.5}
$$

Proof Assume that u_n strongly converges to u. Fix $\xi\in\mathcal{V}(M)$ and observe that it holds

$$
\lim_{n\to\infty}\|u_n-\xi\|^2_{\mu_n}=\|u\|^2_\mu+\|\xi\|^2_\mu-2\lim_{n\to\infty}\langle u_n,\xi\rangle_{\mu_n}=\|u-\xi\|^2_\mu.
$$

By lemma 9.14 we get that

$$
\int \mathrm{D}^2(\xi(x),\xi(y))d\boldsymbol{\gamma}_n(x,y)\le\big(1+(\mathrm{L}(\xi))^2\big)\int \mathrm{d}^2(x,y)d\boldsymbol{\gamma}_n(x,y),
$$

therefore we obtain

$$
\begin{aligned}
\int \mathrm{D}^2\big(u(x),&u_n(y)\big)d\boldsymbol{\gamma}_n(x,y)\\
&\le 3\int \mathrm{D}^2\big(u(x),\xi(x)\big)d\mu(x)+3\int \mathrm{D}^2\big(\xi(x),\xi(y)\big)d\boldsymbol{\gamma}_n(x,y)+3\int \mathrm{D}^2\big(\xi(y),u_n(y)\big)d\mu_n(y)\\
&\le 3\|u-\xi\|_\mu+3\big(1+(\mathrm{L}(\xi))^2\big)\int \mathrm{d}^2(x,y)d\boldsymbol{\gamma}_n(x,y)+3\|u_n-\xi\|_{\mu_n}.
\end{aligned}
$$

The conclusion follows by letting first $n \to \infty$ and then $\xi \to u$ in L^2_μ.

Conversely, assume that (9.5) holds for some sequence of plans $\boldsymbol{\gamma}_n$ satisfying $\int \mathrm{d}^2(x,y)d\boldsymbol{\gamma}_n(x,y) \to 0$. From the inequality $\mathrm{D}((x,u),(y,v)) \geq \big||u|-|v|\big|$ we get

$$\Big|\|u\|_\mu - \|u_n\|_{\mu_n}\Big| \leq \int \mathrm{D}^2(u(x),u_n(y))d\boldsymbol{\gamma}_n(x,y),$$

which gives the convergence of norms.

Now for every $(x,y) \in \operatorname{supp}(\boldsymbol{\gamma}_n)$ choose the optimal curve $(\gamma_n)_x^y$ in the definition of $\mathrm{D}(u(x),u(y))$ and denote by $(T_n)_x^y(\mathrm{u})$ be the optimal transport map of $\mathrm{u} \in T_xM$ to T_yM along $(\gamma_n)_x^y$. Arguing as in the proof of lemma 9.14, it is immediate to check that

$$\big|\xi(y) - (T_n)_x^y(\xi(x))\big| \leq \mathrm{L}(\xi)\int_0^1 \big|\big((\gamma_n)_x^y\big)'\big|(t)dt, \qquad \forall \xi \in \mathcal{V}(M)$$

so that

$$\begin{aligned}\int \big|\xi(y) - (T_n)_x^y(\xi(x))\big|^2 d\boldsymbol{\gamma}_n(x,y) &\leq (\mathrm{L}(\xi))^2 \int \left(\int_0^1 \big|\big((\gamma_n)_x^y\big)'\big|(t)dt\right)^2 d\boldsymbol{\gamma}_n(x,y) \\ &\leq (\mathrm{L}(\xi))^2 \int \mathrm{D}^2\big(u(x),u_n(y)\big)d\boldsymbol{\gamma}_n(x,y) \to 0,\end{aligned} \tag{9.6}$$

for any $\xi \in \mathcal{V}(M)$.

Now fix $\xi \in \mathcal{V}(M)$ and observe that

$$\begin{aligned}\langle u,\xi\rangle_\mu - \langle u_n,\xi\rangle_{\mu_n} &= \int \langle u(x),\xi(x)\rangle + \langle u_n(y),\xi(y)\rangle\, d\boldsymbol{\gamma}_n(x,y) \\ &= \int \big\langle u(x),\xi(x) - (T_n)_y^x\big(\xi(y)\big)\big\rangle + \big\langle (T_n)_x^y\big(u(x)\big) - u_n(y),\xi(y)\big\rangle\, d\boldsymbol{\gamma}_n(x,y).\end{aligned}$$

The conclusion follows from equations (9.6), (9.5) and inequality

$$\Big|(T_n)_x^y\big(u(x)\big) - u_n(y)\Big|^2 \leq \mathrm{D}^2(u(x),u_n(y)).$$

□

Now we turn to the main result regarding C^1 curves. For simplicity of exposition, we will state and prove this result in the case $M = \mathbb{R}^d$: it is not difficult to check, using lemma 9.16 above and adapting the formalism, that the same holds for measures on manifolds.

Theorem 9.18 (Behavior of C^1 curves) *Let (μ_t) be a C^1 curve and (v_t) its velocity vector field. Then:*

i) The continuity equation holds 'for every time', that is:

$$\frac{d}{dt}\int \varphi d\mu_t\Big|_{t=t_0} = \langle v_{t_0},\nabla\varphi\rangle_{\mu_{t_0}}, \qquad \forall t_0 \in [0,1], \varphi \in C_c^\infty(\mathbb{R}^d). \tag{9.7}$$

ii) For any $t \in [0,1]$ it holds

$$\lim_{h\to 0} \frac{W_2(\mu_{t+h}, (Id + hv_t)_\# \mu_t)}{h} = 0.$$

iii) The metric derivative $|\dot{\mu}_t|$ of (μ_t) exists for every time and is continuous.

iv) For any $t, s \in [0,1]$ and any choice of optimal plans $\boldsymbol{\gamma}_t^s \in Opt(\mu_t, \mu_s)$, the rescaled plans

$$\boldsymbol{\sigma}_t^s := \left(\pi^1, \frac{\pi^2 - \pi^1}{s-t}\right)_\# \boldsymbol{\gamma}_t^s,$$

converge to $(Id, v_t)_\# \mu_t$ in $\mathscr{P}_2(TM)$ as $s \to t$.

Proof We start with (i). Observe that from the validity of the continuity equation it follows immediately that $t \mapsto \int \varphi d\mu_t$ is absolutely continuous for every $\varphi \in C_c^\infty(\mathbb{R}^d)$ and that equation

$$\frac{d}{dt}\int \varphi d\mu_t = \langle v_t, \nabla\varphi\rangle_{\mu_t}$$

holds for a.e. t. The hypothesis of continuity of (v_t) gives that the right hand side of the above equation is continuous, thus the function $t \mapsto \int \varphi d\mu_t$ is C^1 and equality holds for every t.

Now we turn to (ii). Let $\boldsymbol{\eta} \in \mathscr{P}(\mathrm{AC}([0,1], \mathbb{R}^{\mathrm{d}}))$ be a measure associated to (μ_t) via theorem 9.5. Observe that

$$\begin{aligned}
\int |\gamma(t) - \gamma(s)|^2 d\boldsymbol{\eta}(\gamma) &= \int \left| \int_t^s \gamma'(r) dr \right|^2 d\boldsymbol{\eta}(\gamma) = \int \left| \int_t^s v_r(\gamma(r)) dr \right|^2 d\boldsymbol{\eta}(\gamma) \\
&\leq |s-t| \int \int_t^s |v_r(\gamma(r))|^2 \, dr \, d\boldsymbol{\eta}(\gamma) \\
&= |s-t| \int_t^s \|v_r\|_{\mu_r}^2 dr \to 0 \qquad\qquad \text{as } s \to t.
\end{aligned}$$

Thus defining $\boldsymbol{\mu}_t^s := (\mathrm{e_t}, \mathrm{e_s})_\# \boldsymbol{\eta}$ we have $\|x - y\|_{\boldsymbol{\mu}_t^s} \to 0$ as $s \to t$ for any $t \in [0,1]$. By the characterization of convergence of maps given in lemma 9.17 above, and by the hypothesis of continuity of (v_t) we get

$$\int |v_t(\gamma(t)) - v_s(\gamma(s))|^2 d\boldsymbol{\eta}(\gamma) = \|v_t(x) - v_s(y)\|_{\boldsymbol{\mu}_t^s}^2 \to 0, \qquad \text{as } s \to t. \tag{9.8}$$

Now define the function $f : [0,1]^2 \to \mathbb{R}$ as

$$f(t,s) := \|\mathrm{e_t} + (s-t)v_t(\mathrm{e_t}) - \mathrm{e_s}\|_{\boldsymbol{\eta}} = \sqrt{\int \Big|\gamma(t) + (s-t)v_t(\gamma(t)) - \gamma(s)\Big|^2 d\boldsymbol{\eta}(\gamma)}.$$

It is immediate to verify that f is absolutely continuous as a function of s (that is, for any 'frozen' t). For the derivative of f^2 it holds

$$\begin{aligned}\frac{1}{2}\frac{d}{ds}f^2(t,s) &= \int \big\langle \gamma(t)+(s-t)v_t(\gamma(t))-\gamma(s), v_t(\gamma(t))-\gamma'(s)\big\rangle \, d\boldsymbol{\eta}(\gamma)\\ &= \int \langle \gamma(t)+(s-t)v_t(\gamma(t))-\gamma(s), v_t(\gamma(t))-v_s(\gamma(s))\rangle \, d\boldsymbol{\eta}(\gamma)\\ &\leq f(t,s)\sqrt{\int |v_t(\gamma(t))-v_s(\gamma(s))|^2 d\boldsymbol{\eta}(\gamma)}.\end{aligned}$$

Thus from equation (9.8) we get

$$\lim_{s\to t}\frac{f(t,s)}{|s-t|}=0.$$

Observe that the plan $(\mathrm{e_t}+(s-t)v_t(\mathrm{e_t}),\mathrm{e_s})_\#\boldsymbol{\eta}$ is admissible for the couple $\big((Id+(s-t)v_t)_\#\mu_t,\mu_s\big)$, thus the previous equation yields

$$\lim_{s\to t}\frac{W_2\Big((Id+(s-t)v_t)_\#\mu_t,\mu_s\Big)}{|s-t|}\leq \lim_{s\to t}\frac{f(t,s)}{|s-t|}=0.$$

Part (iii) follows immediately from (ii) and lemma 9.16: indeed for any $t\in[0,1]$ it holds

$$\lim_{s\to t}\frac{W_2(\mu_t,\mu_s)}{|s-t|}=\lim_{s\to t}\frac{W_2(\mu_t,(Id+(s-t)v_t)_\#\mu_t)}{|s-t|}=\|v_t\|_{\mu_t}. \tag{9.9}$$

To prove part (iv), start proving that $\|y\|_{\boldsymbol{\sigma}_t^s}\to\|v_t\|_{\mu_t}$ as $s\to t$ and that $\int\langle\nabla\varphi(x),y\rangle\,d\boldsymbol{\sigma}_t^s\to\langle\nabla\varphi,v_t\rangle_{\mu_t}$ as $s\to t$ for any $\varphi\in C_c^\infty(\mathbb{R}^d)$. From the equality

$$\|y\|_{\boldsymbol{\sigma}_t^s}=\frac{\|y-x\|_{\boldsymbol{\gamma}_t^s}}{|s-t|}=\frac{W_2(\mu_t,\mu_s)}{|s-t|}$$

and equation (9.9) we get the convergence of norms. Now fix $\varphi\in C_c^\infty(\mathbb{R}^d)$ and observe that

$$\begin{aligned}\langle v_t,\nabla\varphi\rangle_{\mu_t} &=\frac{d}{dt}\int\varphi d\mu_t=\lim_{s\to t}\frac{\int\varphi d\mu_s-\int\varphi d\mu_t}{s-t}=\lim_{s\to t}\frac{\int\varphi(y)-\varphi(x)d\boldsymbol{\gamma}_t^s(x,y)}{s-t}\\ &=\lim_{s\to t}\int\left\langle\nabla\varphi(x),\frac{y-x}{s-t}\right\rangle d\boldsymbol{\gamma}_t^s(x,y)+R(t,s,\varphi)=\lim_{s\to t}\int\langle\nabla\varphi(x),y\rangle\,d\boldsymbol{\sigma}_t^s,\end{aligned}$$

where the last equality follows from the fact that the reminder term $R(t,s,\varphi)$ is bounded by $|R(t,s,\varphi)|\leq\mathrm{Lip}(\nabla\varphi)\frac{W_2^2(\mu_t,\mu_s)}{2(s-t)}$ and thus converges to 0 as $s\to t$.

Now let $v_t^s(x):=\int y d(\boldsymbol{\sigma}_t^s)_x(y)\in L^2_{\mu_t}$, where $(\boldsymbol{\sigma}_t^s)_x$ is the disintegration of $\boldsymbol{\sigma}_t^s$ w.r.t. the projection onto the first coordinate. Observe that we have

$$\lim_{s\to t}\langle\nabla\varphi,v_t^s\rangle_{\mu_t}=\lim_{s\to t}\int\langle\nabla\varphi(x),y\rangle\,d\boldsymbol{\sigma}_t^s=\langle\nabla\varphi,v_t\rangle_{\mu_t},\qquad\forall\varphi\in C_c^\infty(\mathbb{R}^d),$$

which tells that $\mathrm{P}_{\mu_t}(v_t^s)$ weakly converges to v_t in $L^2_{\mu_t}$ (since $v_t \in \mathrm{Tan}_{\mu_t}(\mathscr{P}_2(M))$). From

$$\|v_t\|_{\mu_t} \leq \varliminf_{s\to t} \|\mathrm{P}_{\mu_t}(v_t^s)\|_{\mu_t} \leq \varlimsup_{s\to t} \|\mathrm{P}_{\mu_t}(v_t^s)\|_{\mu_t} \leq \varlimsup_{s\to t} \|v_t^s\|_{\mu_t} \leq \varlimsup_{s\to t} \|y\|_{\boldsymbol{\sigma}_t^s} = \|v_t\|_{\mu_t},$$

we deduce that $\lim_{s\to t} \|v_t^t\|_{\mu_t} = \|v_t\|_{\mu_t}$ and therefore

$$\varlimsup_{s\to t} \|\mathrm{P}^\perp_{\mu_t}(v_t^s)\|^2_{\mu_t} \leq \varlimsup_{s\to t} \|v_t^s\|^2_{\mu_t} \varliminf_{s\to t} \|\mathrm{P}_{\mu_t}(v_t^s)\|^2_{\mu_t} = 0,$$

which implies $v_t^s \to v_t$ in $L^2_{\mu_t}$ as $s \to t$. It is easy to check that such L^2-convergence implies that $(Id, v_t^s)_\# \mu_t$ converges to $(Id, v_t)_\# \mu_t$ in $\mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$, therefore to conclude it is enough to check that $W_2(\boldsymbol{\sigma}_t^s, (Id, v_t^s)_\# \mu_t)$ goes to 0 as $s \to t$. The plan $(\pi^1, \pi^2, \pi^1, v_t^s \circ \pi^1)_\# \boldsymbol{\sigma}_t^s$ is admissible for the couple $(\boldsymbol{\sigma}_t^s, (Id, v_t^s)_\# \mu_t)$, therefore we have

$$W_2(\boldsymbol{\sigma}_t^s, (Id, v_t^s)_\# \mu_t) \leq \|y - v_t^s(x)\|_{\boldsymbol{\sigma}_t^s} \leq \Big| \|y\|_{\sigma_t^s} - \|v_t^s\|_{\mu_t} \Big| \to 0,$$

as $s \to t$. □

We conclude with some examples which help understanding the behavior of C^1 curves.

Example 9.19 (Being tangent everytime matters) In the definition of C^1 curves, it is clear that requirement of continuity in time of the velocity vector field plays a key role. It may be less clear that the requirement for this vector field to be tangent for everytime is both non trivial and necessary. Here we are going to show why. Note that the example is the same with which we discussed the difficulties in finding the pointwise derivative of the exponential map.

Let $Q = [0,1] \times [0,1]$ be the unit square in $\mathbb{R}^2$ and T_i, $i = 1,2,3,4$ be the four triangles in which Q is divided by its diagonal. Define the vector field $v : Q \to \mathbb{R}^2$ as

$$\begin{aligned} v_{|T_1} &:= (1,1), \\ v_{|T_2} &:= (-1,1), \\ v_{|T_3} &:= (-1,-1), \\ v_{|T_4} &:= (1,-1). \end{aligned}$$

Now define $\mu := \mathcal{L}^2_{|Q}$ and observe that $v \in L^2_\mu$ and that for $t > 0$, $\mu_t := (Id + tv)_\# \mu$ is made of 4 connected components, each one the translation of one of the T_i. Also, it is easy to check that $v \notin \mathrm{Tan}_\mu(\mathscr{P}_2(M))$.

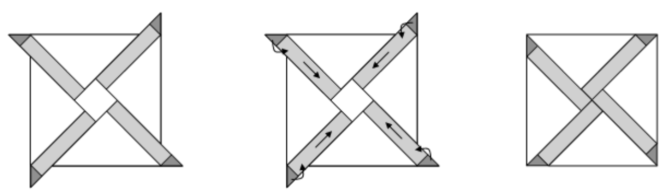
The function $v_t := v \circ (Id + tv)^{-1}$ is constant on each of those components and therefore clearly tangent. It follows that for positive t we have $|\dot{\mu}_t| = \|v_t\|_{\mu_t} = \|v\|_\mu = \sqrt{2}$.

Thus we have an absolutely continuous curve (μ_t) and a velocity vector field (v_t) which is continuous in time and tangent for every $t > 0$ but not for $t = 0$. We want to show that the metric derivative is not continuous in 0, and that the rescaled optimal plans do not converge to v_0. Both of this claims will be achieved if we show that $|\dot{\mu}_0| \leq 1$. To see this consider the function g on Q defined as

$$\begin{aligned} g_{|T_1} &:= (0,1), \\ g_{|T_2} &:= (-1,0), \\ g_{|T_3} &:= (0,-1), \\ g_{|T_4} &:= (1,0), \end{aligned}$$

and let $\nu_t := (Id + tg)_{\#}\mu$.

Then clearly $W(\mu_t, \nu_t) \leq \|f - g\|_\mu = 1$, therefore the conclusion will follow if we show that $W(\nu_t, \mu) = o(t)$. This can be proved by finding a family of transport maps whose cost goes to 0 faster than t: see the picture below (a mass of order t is moved of a distance of order t, giving a cost of order $t^{3/2}$).



■

Example 9.20 (Discontinuity of the projection operator) A simple variation of the previous example shows that the projection operator may very well be discontinuous along a C^1 curve. Indeed, consider the curve $(\tilde{\mu}_t)$ defined by

$$\tilde{\mu}_t := \mu_{t^2},$$

with (μ_t) as above. Let $(\tilde{v}_t)$ be its velocity vector field and observe that $\tilde{v}_t = 2tv_{t^2}$ and thus is continuous and tangent on the whole $[0,1]$.

Now, the continuity of the projection operator along a curve is not affected by the parametrization of the curve itself, thus to conclude is enough to observe that the projection operator is discontinuous along (μ_t) (the vector field (v_t) is continuous along (μ_t), but the projected one $(\mathrm{P}_{\mu_t}(v_t))$ is not).

This example confirms that the continuity of the velocity vector field has little to do with the continuity of the projection operator: what matters the most, is the Lipschitz property. ■

Example 9.21 (Supporting curves are not C^1) Let (μ_t) be a C^1 curve, (v_t) its velocity vector field and $\boldsymbol{\eta} \in \mathscr{P}(\Gamma([0,1], M))$ a measure associated to (μ_t, v_t) via theorem 9.5. We know that for almost every γ in the support of $\boldsymbol{\eta}$ it holds

$$\gamma'(t) = v_t(\gamma(t)), \qquad a.e.\ t \in [0,1].$$

Thus, given the continuity in time of the v_t's, a natural question is whether almost every γ in the support of $\boldsymbol{\eta}$ is C^1 and whether the above equality holds for every time, at least for some choice of $\boldsymbol{\eta}$.

The answer to both questions is no, as shown by the following example. Let $M = \mathbb{R}$ and for every $t \in [0,1]$ define

$$\mu_t := \frac{1}{2}\mathcal{L}^1|_{[0,2t]} + \mathcal{L}^1|_{[2t,1+t]}.$$

It is immediate to verify that (μ_t) is absolutely continuous and that its velocity vector field (v_t) is given by

$$v_t(x) = \begin{cases} 0 & \text{if } x \in [0,2t], \\ 1 & \text{if } x \in [2t, 1+t], \end{cases}$$

which is tangent and continuous (in time) in the whole $[0,1]$. Thus (μ_t) is C^1. Our claim follows by noticing that the absolutely continuous solution γ of

$$\begin{cases} \gamma(0) = x_0 \in [0,1] \\ \gamma'(t) = v_t(\gamma(t)), \end{cases}$$

is given by

$$\gamma(t) = \begin{cases} x_0 + t & \text{if } t \in [0,x_0], \\ 2x_0 & \text{if } t \in [x_0,1], \end{cases}$$

and thus is not C^1 for any $x_0 \in (0,1)$. ■

9.3 A weak notion of absolute continuity of vector fields

We want to conclude this work by mentioning the possibility of defining a notion of absolute continuity of vector fields along a general absolutely continuous curve, thus generalizing the definition given for regular curves. Not surprisingly, this notion is given in terms of weak derivation. The approach we propose is not conclusive, and further studies are needed to fully understand the subject.

Before beginning, let us point out the following fact: we know that 'something goes wrong' when trying to deal with general absolutely continuous curves. For instance, the parallel transport doesn't necessarily exist. However, the kind of problems which come out, are generally related to the lack of regularity of time variation of the tangent spaces along a curve. What we try do here, has some hope of making sense because we completely avoid the study of tangent vector fields, and just focus on general ones.

Our starting point is the following proposition, which is a direct consequence of the density of vector fields in $\mathcal{V}(M \times [0,1])$ in the class of absolutely continuous vector fields along a given regular curve (proposition 3.20). We will denote by $\mathcal{V}(M \times (0,1))$ the set of vector fields smooth in time and space and compactly supported in $M \times (0,1)$.

Proposition 9.22 *Let (μ_t) be a regular curve and (u_t) an L^2 vector field defined along it. Then (u_t) admits an absolutely continuous representative (that is, is absolutely continuous up to redefining it for a negligible set of times) if and only if there exists an L^1 vector field (w_t) such that*

$$\int_0^1 \langle u_t, \partial_t \xi_t + \nabla\xi \cdot v_t \rangle_{\mu_t} dt = -\int_0^1 \langle w_t, \xi_t \rangle_{\mu_t} dt, \qquad \forall (\xi_t) \in \mathcal{V}(M \times (0,1)), \tag{9.10}$$

in this case $w_t = \frac{\boldsymbol{d}}{dt} u_t$.

Proof The 'only if' part is obvious by the Leibniz rule (3.2) and the formula for the total derivative of (ξ_t) (3.3). So we turn to the 'if' part.

Up to restricting to an interval of the kind $[\varepsilon, 1-\varepsilon]$ and reparametrizing, we may assume that our class of test vector fields is the whole $\mathcal{V}(M \times [0,1])$.

Thus assume that equation (9.10) holds for every $(\xi_t) \in \mathcal{V}(M \times [0,1])$. Let $\{e_j\}_{j\in\mathbb{N}}$ be an orthonormal base of $L^2_{\mu_0}$, fix $j \in \mathbb{N}$ and consider the absolutely continuous vector field $(\tau_0^t(e_j))$. Arguing as in proposition 3.18 and using the fact that $\int_0^1 \|v_t\|^2_{\mu_t} dt < \infty$, it is not hard to find a sequence $n \mapsto (\xi_t^n) \in \mathcal{V}(M \times [0,1])$ of vector fields uniformly converging to $(\tau_0^t(e_j))$ such that the maps $t \mapsto \|\frac{\boldsymbol{d}}{dt}\xi_t^n\|_{\mu_t}$ converge to 0 in $L^2(0,1)$ as $n \to \infty$

The uniform convergence of (ξ_t^n) to $(\tau_0^t(e_j))$ implies that for every $t \in [0,1]$ the number $\langle u_t, \xi_t^n \rangle_{\mu_t}$ converges to $\langle u_t, \tau_0^t(e_j) \rangle_{\mu_t}$ as $n \to \infty$, hence from

$$\begin{aligned} \left| \lim_{n\to\infty} \int_t^s \left\langle u_r, \frac{\boldsymbol{d}}{dr}\xi_r^n \right\rangle_{\mu_r} dr \right| &\le \overline{\lim_{n\to\infty}} \int_t^s \left| \left\langle u_r, \frac{\boldsymbol{d}}{dr}\xi_r^n \right\rangle_{\mu_r} \right| dr \\ &\le \overline{\lim_{n\to\infty}} \int_t^s \|u_r\|_{\mu_r} \left\| \frac{\boldsymbol{d}}{dr}\xi_r^n \right\|_{\mu_r} dr \\ &\le \left(\int_0^1 \|u_r\|^2_{\mu_r} dr \right) \overline{\lim_{n\to\infty}} \int_t^s \left\| \frac{\boldsymbol{d}}{dr}\xi_r^n \right\|^2_{\mu_r} dr = 0, \end{aligned}$$

we get:

$$\begin{aligned} \langle u_s, \tau_0^s(e_j) \rangle_{\mu_s} - \langle u_t, \tau_0^t(e_j) \rangle_{\mu_t} &= \lim_{n\to\infty} \langle u_s, \xi_s^n \rangle_{\mu_s} - \langle u_t, \xi_t^n \rangle_{\mu_t} \\ &= \lim_{n\to\infty} \int_t^s \left\langle u_r, \frac{\boldsymbol{d}}{dr}\xi_r^n \right\rangle_{\mu_r} + \langle w_r, \xi_r^n \rangle_{\mu_r} dr = \int_t^s \langle w_r, \tau_0^r(e_j) \rangle_{\mu_r} dr. \end{aligned} \tag{9.11}$$

Equation (9.11) shows that $t \mapsto \langle u_t, \tau_0^t(e_j) \rangle_{\mu_t}$ is absolutely continuous and that its derivative is given, for a.e. t, by $\langle w_t, \tau_0^t(e_j) \rangle_{\mu_t}$.

Now consider the curves $\overline{u}_t := \tau_t^0(u_t) \in L^2_{\mu_0}$ and $\overline{w}_t := \tau_t^0(w_t) \in L^2_{\mu_0}$. Since $\langle \overline{u}_t, e_j \rangle_{\mu_0} =$

$\left\langle u_t, \tau_0^t(e_j)\right\rangle_{\mu_t}$, we have that $t \mapsto \langle \overline{u}_t, e_j\rangle_{\mu_0}$ is absolutely continuous. Also, we know that

$$\left|\langle \overline{u}_s - \overline{u}_t, e_j\rangle_{\mu_0}\right|^2 = \left|\langle u_s, \tau_0^s(e_j)\rangle_{\mu_s} - \left\langle u_t, \tau_0^t(e_j)\right\rangle_{\mu_t}\right|^2 = \left|\int_t^s \langle w_r, \tau_0^r(e_j)\rangle_{\mu_r} \, dr\right|^2$$
$$= \left|\int_t^s \langle \overline{w}_r, e_j\rangle_{\mu_0} \, dr\right|^2 = \left\langle \left|\int_t^s \overline{w}_r dr, e_j\right\rangle_{\mu_0}\right|^2.$$

Adding up over $j \in \mathbb{N}$ we finally obtain that

$$\|\overline{u}_s - \overline{u}_t\|_{\mu_0}^2 = \sum_{j\in\mathbb{N}} \left|\langle \overline{u}_s - \overline{u}_t, e_j\rangle_{\mu_0}\right|^2 = \sum_{j\in\mathbb{N}} \left|\left\langle \int_t^s \overline{w}_r dr, e_j\right\rangle_{\mu_0}\right|^2 = \left|\int_t^s \overline{w}_r \, dr\right|^2,$$

which shows that $t \mapsto \overline{u}_t$ is absolutely continuous and that its derivative is $\overline{w}_t$. By definition, this is the same as to say that (u_t) is absolutely continuous and that its derivative is (w_t). □

What is important in this proposition is that it characterizes the absolute continuity without the need of the flow maps, the translation maps and the Lipschitz property of the velocity vector field: all the ingredients needed make sense along any absolutely continuous curve. Thus we give the following general definition:

Definition 9.23 (Absolutely continuous vector fields along non-regular curves) *Let (μ_t) be an absolutely continuous curve and (u_t) an L^1 vector field defined along it. We say that (u_t) is absolutely continuous if there exists an L^1 vector field (w_t) such that*

$$\int_0^1 \langle u_t, \partial_t \xi_t + \nabla \xi_t \cdot v_t\rangle_{\mu_t} \, dt = -\int_0^1 \langle w_t, \xi_t\rangle_{\mu_t} \, dt, \qquad \forall (\xi_t) \in \mathcal{V}(M \times (0,1)). \tag{9.12}$$

In this case we say that the vector field (w_t) (which is clearly uniquely identified by the above equation) is the derivative $(\frac{\boldsymbol{d}}{dt} u_t)$ of (u_t).

As example of application of this definition, we show that for any geodesic in $\mathscr{P}_2(M)$, the velocity vector field is absolutely continuous and satisfies $\frac{\boldsymbol{d}}{dt} v_t = 0$. This, regardless of the regularity of the geodesic. For simplicity of exposition, we will state and prove the result for the case $M = \mathbb{R}^d$, but it is immediate to verify that it holds in generic Riemannian manifolds (well, in the case $M = \mathbb{R}^d$ we know that any restriction of geodesic is regular, so the statement in this form does not add anything, however, the proof we propose generalizes to manifolds - we run the calculations in the Euclidean case to focus on the most important aspects of the proof).

Let us mention that after finishing the work on this paper, we discovered that in the recent paper [9] Gangbo, Nguyen and Tudorascu run similar computation to discover the equation satisfied by the minimizers of certain Lagrangian, more general than the one examined here (see theorem 3.9 of [9]).

Proposition 9.24 (The simplest Lagrangian) *Let $(\mu_t) \subset \mathscr{P}_2(\mathbb{R}^d)$ be a geodesic. Then its velocity vector field (v_t) is absolutely continuous in $(0,1)$ in the sense of the above definition and*

$$\frac{\boldsymbol{d}}{dt} v_t = 0, \qquad a.e.\ t \in (0,1).$$

Proof Up to splitting the problem in the analysis over the three intervals $[0, \frac{1}{2}]$, $[\frac{1}{3}, \frac{2}{3}]$, $[\frac{1}{2}, 1]$, we may assume that (μ_t) is the unique geodesic from μ_0 to μ_1.

Let $\mathcal{A}$ be the set of absolutely continuous curves $(\tilde{\mu}_t)$ on $[0,1]$ with values in $\mathscr{P}_2(M)$ such that $\tilde{\mu}_0 = \mu_0$ and $\tilde{\mu}_1 = \mu_1$. Define the functional $\mathcal{L} : \mathcal{A} \to \mathbb{R}$ as

$$\mathcal{L}(\tilde{\mu}_t) := \frac{1}{2} \int_0^1 \|\tilde{v}_t\|^2_{\tilde{\mu}_t} dt,$$

where $\tilde{v}_t$ is the velocity vector field of $\tilde{\mu}_t$.

From the fact that

$$\frac{1}{2} \int_0^1 \|v_t\|^2_{\mu_t} dt \geq \frac{1}{2} \int_0^1 \|v_t\|_{\mu_t} dt \geq \frac{1}{2} W_2(\mu_0, \mu_1),$$

we get that the range of any minimizer has to be a minimal geodesic, and, from the first inequality, that this geodesic has to have constant speed. Thus (μ_t) is the unique global minimizer of $\mathcal{L}$ in the class $\mathcal{A}$.

Now choose $(x,t) \mapsto \xi_t(x) \in C^\infty_c(\mathbb{R}^d \times (0,1), \mathbb{R}^d)$, fix $\varepsilon \in \mathbb{R}$ and define:

$$\mu^\varepsilon_t := \mathbf{exp}_{\mu_t}(\varepsilon \xi_t) = \big(Id + \varepsilon \xi_t\big)_\# \mu_t, \qquad \forall t \in [0,1].$$

Using the plans $(\pi^1 + \varepsilon \xi_t \circ \pi^1, \pi^2 + \varepsilon \xi_s \circ \pi^2)_\# \boldsymbol{\gamma}^s_t$, where $\boldsymbol{\gamma}^s_t \in \mathcal{O}pt(\mu_t, \mu_s)$, it is possible to check that (μ^ε_t) is absolutely continuous - we skip the details. We want to identify its velocity vector field. To this aim, start observing that for $|\varepsilon|$ sufficiently small, the map $x \mapsto x + \varepsilon \xi_t(x)$ is invertible for any $t \in [0,1]$. Now choose $\varphi \in C^\infty_c(\mathbb{R}^d)$ and calculate:

$$\begin{aligned} \frac{d}{dt} \int \varphi d\mu^\varepsilon_t &= \frac{d}{dt} \int \varphi \circ (Id + \varepsilon \xi_t) d\mu_t \\ &= \int \langle \nabla \varphi \circ (Id + \varepsilon \xi_t), \varepsilon \partial \xi_t \rangle \, d\mu_t + \int \big\langle \nabla \big(\varphi \circ (Id + \varepsilon \xi_t)\big), v_t \big\rangle \, d\mu_t, \\ &= \varepsilon \int \big\langle \nabla \varphi, \partial \xi_t \circ (Id + \varepsilon \xi_t)^{-1} \big\rangle \, d\mu^\varepsilon_t + \int \langle (\nabla \varphi) \circ (Id + \varepsilon \xi_t), (Id + \varepsilon \nabla \xi_t) \cdot v_t \rangle \, d\mu_t, \\ &= \big\langle \nabla \varphi, v_t \circ (Id + \varepsilon \xi_t)^{-1} \big\rangle_{\mu^\varepsilon_t} + \varepsilon \big\langle \nabla \varphi, \big(\partial_t \xi_t + \nabla \xi_t \cdot v_t\big) \circ (Id + \varepsilon \xi_t)^{-1} \big\rangle_{\mu^\varepsilon_t} . \end{aligned} \tag{9.13}$$

Thus, defing the vector field (v^ε_t) along (μ^ε_t) bying the vector field (v^ε_t) along (μ^ε_t) by

$$v^\varepsilon_t := \Big(v_t + \varepsilon \big(\partial_t \xi_t + \nabla \xi_t \cdot v_t\big)\Big) \circ (Id + \varepsilon \xi_t)^{-1},$$

we just proved that the velocity vector field of (μ^ε_t) is given by $(\mathrm{P}_{\mu^\varepsilon_t}(v^\varepsilon_t))$. Observe that

$$\begin{aligned} \mathcal{L}(\mu^\varepsilon_t) &= \frac{1}{2} \int_0^1 \big\|\mathrm{P}_{\mu^\varepsilon_t}(v^\varepsilon_t)\big\|^2_{\mu \varepsilon_t} dt \leq \frac{1}{2} \int_0^1 \|v^\varepsilon_t\|^2_{\mu \varepsilon_t} dt = \frac{1}{2} \int_0^1 \big\|v_t + \varepsilon \big(\partial_t \xi_t + \nabla \xi_t \cdot v_t\big)\big\|^2_{\mu_t} dt \\ &= \mathcal{L}(\mu_t) + \varepsilon \int_0^1 \langle v_t, \partial_t \xi_t + \nabla \xi_t \cdot v_t \rangle_{\mu_t} dt + \frac{\varepsilon^2}{2} \int_0^1 \|\partial_t \xi_t + \nabla \xi_t \cdot v_t\|^2_{\mu_t} dt, \end{aligned}$$

so that from the minimality of $\mathcal{L}(\mu_t)$, we get

$$0 \leq \lim_{\varepsilon \downarrow 0} \frac{\mathcal{L}(\mu_t^\varepsilon) - \mathcal{L}(\mu_t)}{\varepsilon} \leq \int_0^1 \langle v_t, \partial_t \xi_t + \nabla \xi_t \cdot v_t \rangle_{\mu_t} dt.$$

Letting $\varepsilon \uparrow 0$ we obtain the other inequality, so that we proved:

$$\int_0^1 \langle v_t, \partial_t \xi_t + \nabla \xi_t \cdot v_t \rangle_{\mu_t} dt = 0,$$

which is the thesis. □

Starting from this result, one would like to try to build a general theory of Lagrangian in $(\mathscr{P}_2(M), W_2)$, however, it is important to underline that the definition of weak absolute continuity is still unfit for such a purpose.

For instance, a natural question is whether the inverse implication in the above theorem is true or not, that is: is that true that if the velocity vector field (v_t) of a curve (μ_t) is absolutely continuous in the sense of definition 9.23 and satisfies

$$\frac{\boldsymbol{d}}{dt} v_t = 0,$$

then (μ_t) is a minimal geodesic? The answer is no, as shown in the following example:

Example 9.25 (Other critical points) Let $[-1, 1] \ni t \mapsto \mu_t \in \mathscr{P}_2(\mathbb{R})$ be given by

$$\mu_t := \frac{1}{2}(\delta_{-t} + \delta_t),$$

and let (v_t) its velocity vector field, given, for $t \neq 0$, by

$$v_t(x) := \begin{cases} 1 & \text{if } x = -t, \\ -1 & \text{if } x = t. \end{cases}$$

An explicit calculation shows that (v_t) *is* absolutely continuous in the sense of definition 9.23 and that its derivative is 0. ■

Thus, minimal geodesics are not the only critical points of the Lagrangian $\mathcal{L}$. This is not that surprising, as the same happens in Riemannian manifolds, where it is sufficient for a curve to be locally minimizing geodesics. The curve considered in the example is, heuristically said, 'almost locally minimizing', indeed: its restriction to the intervals $[-1, 0]$ and $[0, 1]$ is a minimal geodesic, and around 0 the mass continue moving straight. Still, observe that such a curve is *not* minimizing in any interval of the form $[-\varepsilon, \varepsilon]$.

Also, there is only one definition of v_0 such that the resulting vector field is continuous in time w.r.t. weak convergence of maps: namely, $v_0 = 0$ (this is a general fact: if a vector field is absolutely continuous in the sense of 9.23, then up to change it on a negligible set of times we can have a vector field continuous w.r.t. *weak* convergence. This is an immediate consequence of the definition). However with such a choice (v_t) is *not* continuous w.r.t. strong convergence of maps.

The situation becomes much worse when considering the following example:

Example 9.26 (Lack of uniqueness of the integral) Let (μ_t), (v_t) as in the previous example. Define the vector field (u_t) along (μ_t) as

$$u_t := \begin{cases} v_t & \text{if } t \le 0, \\ 0 & \text{if } t \ge 0, \end{cases}$$

so that $u_t = v_t$ for half of the interval of definition of the curve. Again, by direct calculation it is easy to check that (u_t) is absolutely continuous in the sense of definition 9.23 and its derivative is 0. ∎

Our impression, is that the irregularities just described cannot be avoided without the introduction of plans into the theory. The point is that the 'true' value of v_0 is the plan $\boldsymbol{\gamma}_0$ given by

$$\boldsymbol{\gamma}_0 := \frac{(Id, -1)_\# \mu_0 + (Id, 1)_\# \mu_0}{2},$$

and with this choice, the family of plans $t \mapsto \boldsymbol{\gamma}_t$, where $\boldsymbol{\gamma}_t := (Id, v_t)_\# \mu_t$ for $t \neq 0$, is absolutely continuous in $(\mathscr{P}_2(TM), W_2)$, where the distance underlying the Wasserstein distance is the Sasaki metric D, like in in section 9.1.

This avoids situations like the one of the second example, indeed observe that defining $\tilde{\boldsymbol{\gamma}}_t := (Id, u_t)_\# \mu_t$ for $t \neq 0$, there is no choice of $\tilde{\boldsymbol{\gamma}}_0$ which lets the family be absolutely continuous on the whole $[-1, 1]$.

References

[1] L. Ambrosio, N. Gigli, *Construction of parallel transport in the Wasserstein space*, Methods and Applications of Analysis, 15 (2008), vol.1, pp. 1–30.

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

[3] 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.

[4] X. Cabré, *Nondivergent elliptic equations on manifolds with nonegative curvature*, Comm. on Pure and Applied Math., 50 (1998), 623–665.

[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] A. Fathi, A. Figalli, *Optimal transportation on non-compact manifolds*, accepter paper Israel J. Math.

[9] W. Gangbo, T. Nguyen, A. Tudorascu, *Hamilton-Jacobi equations in the Wasserstein space*, Methods and Applications of Analysis, 15 (2008), vol.2, pp. 155–184.

[10] N. Gigli, *On the geometry of the space of measures in* $\mathbb{R}^d$ *endowed with the quadratic optimal transportation distance*, PhD thesis, Scuola Normale Superiore, Pisa, 2008. Available at: http://cvgmt.sns.it

[11] N. Gigli, *on Hölder continuity in time of the optimal transport map towards measures along a curve*, submitted paper, available at: http://cvgmt.sns.it

[12] N. Gigli, *On the inverse implication of Brenier-McCann theorems and the structure of* $(\mathscr{P}_2(M), W_2)$, in preparation

[13] 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).

[14] J. Lott, *Some geometric calculations in the wasserstein space*, (2007).

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

[16] Ma, X.-N., Trudinger, N. S., and Wang, X.-J., *Regularity of potential functions of the optimal transportation problem.* Arch. Ration. Mech. Anal. 177, 2 (2005), pp .151–183.

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

[18] McCann, R. J., *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal. 11, 3 (2001), 589–608.

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

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

[21] 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.

[22] K.-T. Sturm, M.K. von Renesse, *Entropic Measure and Wasserstein Diffusion*, submitted paper.

[23] K.-T. Sturm *Entropic measure on multidimensional spaces*, preprint.

[24] K.-T. Sturm, *On the geometry of metric measure spaces. I-II*, Acta Math. 196, 1 (2006), 65–177.

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

[26] L. Zajíček, *On the differentiability of convex functions in finite and infinite dimensional spaces*, Czechoslovak Math. J. 29 (104) (1979), pp. 340–348.