On the differential structure of metric measure spaces and applications

Nicola Gigli *

May 21, 2013

Abstract

The main goals of this paper are:

i) To develop an abstract differential calculus on metric measure spaces by investigating the duality relations between differentials and gradients of Sobolev functions. This will be achieved without calling into play any sort of analysis in charts, our assumptions being: the metric space is complete and separable and the measure is Radon and non-negative.

ii) To employ these notions of calculus to provide, via integration by parts, a general definition of distributional Laplacian, thus giving a meaning to an expression like $\Delta g = \mu$, where g is a function and μ is a measure.

iii) To show that on spaces with Ricci curvature bounded from below and dimension bounded from above, the Laplacian of the distance function is always a measure and that this measure has the standard sharp comparison properties. This result requires an additional assumption on the space, which reduces to strict convexity of the norm in the case of smooth Finsler structures and is always satisfied on spaces with linear Laplacian, a situation which is analyzed in detail.

Contents

*Université de Nice. email: nicola.gigli@unice.fr

1

1   Introduction

The original motivation behind the project which lead to this paper, was to understand if and up to what extent the standard Laplacian comparison estimates for the distance function - valid on smooth Riemannian manifolds with Ricci curvature bounded from below - hold on a non-smooth setting. Recall that if M is a Riemannian manifold with Ricci curvature not less than, say, 0, then for every $x_0 \in M$ the function $g(x) := \mathsf{d}(x, x_0)$ satisfies

$$\Delta g \leq \frac{N-1}{g}, \qquad \text{on } M \setminus \{x_0\}, \tag{1.1}$$

where N is (a bound from above on) the dimension of M. Similar bounds from above hold if the Ricci is bounded from below by $K \in \mathbb{R}$. Even in a smooth context, the distance function is not smooth, in particular is not C^2, therefore the inequality (1.1) must be properly interpreted. There are various ways of doing this, a possibility is to think (1.1) as an inequality between distributions, so that its meaning is: the distributional Laplacian of g is a measure whose singular part w.r.t. the volume measure is non-positive on $M \setminus \{x_0\}$ and such that the density of the absolutely continuous part is bounded from above by $\frac{N-1}{g}$.

    In the seminal papers [30] and [46], [47] Lott-Villani on one side and Sturm on the other proposed a general definition of metric measure space with Ricci curvature bounded from below by $K \in \mathbb{R}$ and dimension bounded from above by $N \in (1, \infty]$ (for finite N's, in [30] only the case $K = 0$ was considered), in short, these are called $\mathsf{CD}(K, N)$ spaces. The study of the properties of these spaces is still a very active research area. Broadly speaking, a central question is 'which of the properties valid on a manifold with Ricci curvature bounded below and dimension bounded above are true in the non-smooth context?'.

    As said, this paper addresses the problem of the Laplacian comparison for the distance function. Given that we expect such Laplacian to be a measure, part of the work that will be carried out here is to give a meaning, in the setting of metric measure spaces, to expressions like $\Delta g = \mu$, where g is a function and μ a measure. This is a non-trivial task because the standard analytic tools available in this setting are only sufficient to speak about 'the norm of the distributional gradient' of a Sobolev function (sometimes called minimal generalized upper gradient [13], or minimal weak upper gradient [25], [44]), which is not enough to give a meaning to an integration by parts formula. We refer to [24] for an overview on the subject and detailed bibliography (see also [7] for recent progresses). In terms of defining the differential of

a function, in the seminal paper [13], Cheeger, assuming a doubling condition on the measure and the validity of a local weak Poincaré inequality, proved the existence of an abstract notion of differential for which the Leibniz rule holds (for a comparison of his approach with ours, see Remark 3.21). A different approach has been proposed by Weaver [49] using the notion of L^∞-module. Yet, these approaches are not sufficient to define integration by parts because the duality relations between differentials and gradients has been not investigated. This is the focus of the current work.

The distinction between differentials and gradients is particularly important in this framework, because $\mathsf{CD}(K, N)$ spaces include Finsler structures (as proved by Ohta in [32] by generalizing the analogous result valid for flat normed spaces proved by Cordero-Erasquin, Sturm and Villani - see the last theorem in [48]), so we must be ready to deal with situations where tangent and cotangent vectors - whatever they are - can't be identified. Recall also that on a Finsler setting the Laplacian is *not* a linear operator (see e.g. [45]). In particular, since the construction we propose reduces to the standard one on Finsler spaces, the Laplacian we define is different from the linear, chart dependent Laplacian introduced in [13].

Shortly said, what we will do is to define the $\mathfrak{m}$-a.e. value of $Df(\nabla g)$, i.e. the differential of f applied to the gradient of g, for Sobolev functions f, g. Then the definition of distributional Laplacian $\Delta g = \mu$ will be given by

$$\int f \, \mathrm{d}\mu = - \int Df(\nabla g) \, \mathrm{d}\mathfrak{m},$$

for any f in an appropriate class of Lipschitz test functions. In order to illustrate the construction we will use, we open a parenthesis and show how it works in a flat normed situation. This will serve as guideline - intended for non-specialists - for the rest of the introduction and provide a quick and simple reference for the procedures that will be used in the paper. Also, it will hopefully make clear why the distributional Laplacian, in general, is not only not linear but also multivalued.

1.1 The simple case of normed spaces

Consider the space $(\mathbb{R}^d, \| \cdot \|, \mathcal{L}^d)$, where $\| \cdot \|$ is any norm, possibly not coming from a scalar product. Assume for the moment that $\| \cdot \|$ is smooth and strictly convex. Given a smooth function $f : \mathbb{R}^d \to \mathbb{R}$, its differential Df is intrinsically well defined as a cotangent vector field, so that we may think at it as a map from $\mathbb{R}^d$ to its dual $(\mathbb{R}^d)^*$. To define the gradient ∇f, which is a tangent vector field, we need some duality map from the cotangent to the tangent space. This is produced by using the norm $\| \cdot \|$: the differential of the map $\frac{\| \cdot \|^2}{2} : \mathbb{R}^d \to \mathbb{R}$ at some point $v \in \mathbb{R}^d$ is a linear map from $T_v \mathbb{R}^d \sim \mathbb{R}^d$ to $\mathbb{R}$, i.e. a cotangent vector. In other words, the differential of $\frac{\| \cdot \|^2}{2}$ is a map $\mathrm{Dual} : \mathbb{R}^d \to (\mathbb{R}^d)^*$. If $\| \cdot \|$ is strictly convex, as we are assuming for the moment, Dual is injective and can therefore be inverted to get a map $\mathrm{Dual}^{-1} : (\mathbb{R}^d)^* \to \mathbb{R}^d$. Then the gradient $\nabla g : \mathbb{R}^d \to \mathbb{R}^d$ of the smooth function g is defined by

$$\nabla g := \mathrm{Dual}^{-1}(Dg).$$

The divergence operator $\nabla \cdot$, which takes a smooth tangent vector field ξ and returns a function, is defined as (the opposite of) the adjoint of the differential operator, i.e.:

$$\int f \, \nabla \cdot \xi \, \mathrm{d}\mathcal{L}^d := - \int Df(\xi) \, \mathrm{d}\mathcal{L}^d, \qquad \forall f \in C_c^\infty(\mathbb{R}^d).$$

3

In particular, in order to define the divergence, we need to use the reference measure.

Finally, the Laplacian of the smooth function g is defined as

$$\Delta g := \nabla \cdot (\nabla g) = \nabla \cdot (\mathrm{Dual}^{-1}(Dg)),$$

(in particular, in order to define Δ we had to use both the norm and the reference measure. This is an heuristic explanation of why the correct abstract setting where a Laplacian can be defined is that of metric measure spaces). Notice that the differential and the divergence operators are always linear functionals, but the duality map Dual^{-1} is linear if and only if the norm comes from a scalar product. Hence, in the general normed situation, the Laplacian is non-linear.

Now suppose that we want to give a meaning to a formula like $\Delta g = \mu$, where μ is some measure on $\mathbb{R}^d$, i.e., suppose that we want to provide a notion of distributional Laplacian. An effect of the non-linearity of Δ, is that there is no hope to 'drop all the derivatives on the test function', that is, we can't hope that a formula like

$$\int f \, \mathrm{d}\Delta g := \int \Delta^{\mathrm{tr}} f \, g \, \mathrm{d}\mathcal{L}^d, \qquad \forall f \in C_c^\infty(\mathbb{R}^d),$$

holds, whatever the operator Δ^{tr} is. Simply because the right hand side would be linear in g, while the left one is not.

Hence, the best we can do is to drop only one derivative on the test function and define

$$\int f \, \mathrm{d}\Delta g := - \int Df(\nabla g) \, \mathrm{d}\mathcal{L}^d, \qquad \forall f \in C_c^\infty(\mathbb{R}^d). \tag{1.2}$$

Observe that an effect of this approach, is that in order to define the distributional Laplacian, we need to first impose some Sobolev regularity on g in order to be sure that ∇g is well defined. This is of course different from the standard approach in the Euclidean setting, where distributional derivatives make sense as soon as the function is in L_{loc}^1. Yet, a posteriori this is not a big issue, see Remark 4.6 for comments in this direction.

In proposing definition (1.2), which is the standard one in Finsler geometries, we worked under the assumption that $\| \cdot \|$ was smooth and strictly convex, so that the gradient of a Sobolev function was a well defined vector field. Actually, we didn't really use the smoothness assumption: if this fails but the norm is still strictly convex, then the duality map Dual is possibly multivalued, but still injective, so that its inverse Dual^{-1} is still uniquely defined. A new behavior comes if we drop the assumption of strict convexity: in this case the map Dual is not injective anymore, which means that its inverse Dual^{-1} is multivalued. Hence the gradient ∇g of a function is typically a multivalued vector field. A simple example which shows this behavior is given by the space $(\mathbb{R}^2, \| \cdot \|_\infty)$ and the linear map $g(x_1, x_2) := x_1$: all the vectors of the kind $(1, x_2)$, with $x_2 \in [-1, 1]$, have the right to be called gradient of g.

Thus in general, the object $Df(\nabla g)$ is not single valued, not even for smooth f, g, so that the best we can do is to define the two single valued maps $D^+ f(\nabla g)$, $D^- f(\nabla g)$ by

$$D^+ f(\nabla g)(x) := \max_{v \in \nabla g(x)} Df(v), \qquad D^- f(\nabla g)(x) := \min_{v \in \nabla g(x)} Df(v), \tag{1.3}$$

and then the distributional Laplacian of g by the chain of inequalities

$$- \int D^+ f(\nabla g) \, \mathrm{d}\mathcal{L}^d \le \int f \, \mathrm{d}\Delta g \le - \int D^- f(\nabla g) \, \mathrm{d}\mathcal{L}^d, \qquad \forall f \in C_c^\infty(\mathbb{R}^d). \tag{1.4}$$

A clear drawback of this approach is that in general the Laplacian can be multivalued (see Proposition 4.9 for explicit examples). Yet, even for this potentially multivalued notion, a reasonable calculus can be developed, see Section 4.2.

Both for theoretical and practical reasons, it is good to have an alternative description of the objects $D^{\pm}f(\nabla g)$. Here it comes a key point: it holds

$$D^+ f(\nabla g) = \inf_{\varepsilon > 0} \frac{\|D(g + \varepsilon f)\|_*^2 - \|Dg\|_*^2}{2\varepsilon}, \qquad (1.5)$$

$$D^- f(\nabla g) = \sup_{\varepsilon < 0} \frac{\|D(g + \varepsilon f)\|_*^2 - \|Dg\|_*^2}{2\varepsilon}. \qquad (1.6)$$

To see why, observe that for any $g \in C_c^\infty(\mathbb{R}^d)$, any $x \in \mathbb{R}^d$ and any tangent vector $v \in T_x\mathbb{R}^d \sim \mathbb{R}^d$ it holds

$$Dg(v) \leq \frac{\|Dg(x)\|_*^2}{2} + \frac{\|v\|^2}{2},$$

and that $v \in \nabla g(x)$ if and only if

$$Dg(v) \geq \frac{\|Dg(x)\|_*^2}{2} + \frac{\|v\|^2}{2}.$$

Write the first inequality with g replaced by $g + \varepsilon f$, subtract the second and divide by $\varepsilon > 0$ (resp. $\varepsilon < 0$) to get that $\leq$ holds in (1.5) (resp. $\geq$ in (1.6)). The fact that equality holds can be deduced with a compactness argument from the convexity of $\varepsilon \mapsto \|D(g + \varepsilon f)\|_*^2(x)$, by considering $v_\varepsilon \in \nabla(g + \varepsilon f)$ and letting $\varepsilon \downarrow 0$ (resp. $\varepsilon \uparrow 0$).

1.2 The general situation

Let us now analyze how to deal with the case of metric measure spaces $(X, \mathsf{d}, \mathfrak{m})$. Recall that our assumptions are: (X, d) is complete and separable and $\mathfrak{m}$ is a Radon non-negative measure. There is a standard notion concerning the meaning of '$f : X \to \mathbb{R}$ has a p-integrable p-weak upper gradient', $p \in (1, \infty)$, one of the various equivalent ways (see [5] and Appendix B) of formulating it is: there exists $G \in L^p(X, \mathfrak{m})$ such that

$$\int |f(\gamma_1) - f(\gamma_0)| \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \iint_0^1 G(\gamma_t)|\dot{\gamma}_t| \, \mathrm{d}t \, \mathrm{d}\boldsymbol{\pi}(\gamma), \qquad (1.7)$$

for any $\boldsymbol{\pi} \in \mathscr{P}(C([0,1], X))$ such that:

- $\iint_0^1 |\dot{\gamma}_t|^q \, \mathrm{d}t \, \mathrm{d}\boldsymbol{\pi}(\gamma) < \infty$, where q is the conjugate exponent of p,

- $(\mathrm{e}_t)_\sharp \boldsymbol{\pi} \leq C\mathfrak{m}$ for any $t \in [0,1]$ and some $C \in \mathbb{R}$, where $\mathrm{e}_t : C([0,1], X) \to X$ is the evaluation map defined by $\mathrm{e}_t(\gamma) := \gamma_t$ for any $\gamma \in C([0,1], X)$.

The class of such f's will be called p-Sobolev class and denoted by $\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$. Notice that we are not asking for any integrability of the functions themselves. For $f \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ there exists a minimal $G \geq 0$ satisfying (1.7), which we will denote by $|Df|_w$ ($|Df|_w$ is typically called minimal p-weak/generalized upper gradient, but from the perspective we are trying to propose, this object is closer to a differential than to a gradient, given that it acts in duality

5

with derivative of curves, whence the notation used). For the potential dependence of $|Df|_w$ on p, see Remark 2.5.

Chapter 3 The real valued map $|Df|_w$ is what plays the role of the norm of the distributional differential for Sobolev functions $f \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$. Hence by analogy with (1.5), (1.6), for $f, g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ we define $\mathfrak{m}$-a.e. the maps $D^\pm f(\nabla g) : X \to \mathbb{R}$ by:

$$
\begin{aligned}
D^+ f(\nabla g) &:= \inf_{\varepsilon > 0} \frac{|D(g + \varepsilon f)|_w^2 - |Dg|_w^2}{2\varepsilon} = \inf_{\varepsilon > 0} \frac{|D(g + \varepsilon f)|_w^p - |Dg|_w^p}{p\varepsilon |Dg|_w^{p-2}}, \\
D^- f(\nabla g) &:= \sup_{\varepsilon < 0} \frac{|D(g + \varepsilon f)|_w^2 - |Dg|_w^2}{2\varepsilon} = \sup_{\varepsilon < 0} \frac{|D(g + \varepsilon f)|_w^p - |Dg|_w^p}{p\varepsilon |Dg|_w^{p-2}}.
\end{aligned}
\tag{1.8}
$$

Since in general $D^+ f(\nabla g) \neq D^- f(\nabla g)$, for given g the maps $f \mapsto D^\pm f(\nabla g)$ are typically not linear, yet some structural properties are retained even in this generality, because $f \mapsto D^+ f(\nabla g)$ is pointwise convex, and $f \mapsto D^- f(\nabla g)$ pointwise concave. For given f, the maps $g \mapsto D^\pm f(\nabla g)$ have some very general semicontinuity properties, see Proposition 3.2. We will call spaces such that $D^+ f(\nabla g) = D^- f(\nabla g)$ $\mathfrak{m}$-a.e. for any $f, g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$, q-infinitesimally strictly convex spaces: in normed spaces, this notion reduces to the strict convexity of the norm. On such spaces, the common value of $D^+ f(\nabla g)$ and $D^- f(\nabla g)$ will be denoted by $Df(\nabla g)$.

From the definition (1.8) it is not hard to prove that $D^\pm f(\nabla g)$ satisfies natural chain rules. The Leibniz rule could then be proved as a consequence of them, see Remark 3.20. Yet we will proceed in a different way, as this will give us the opportunity to introduce important concepts which play a key role in the proof of the Laplacian comparison: we will define the notion of gradient vector field. The idea is the following. On general metric measure spaces we certainly do not have tangent vectors. But we have curves. And we can think at a tangent vector as the derivation at time 0 along a sufficiently smooth curve. Even on $\mathbb{R}^d$, in order to get this interpretation it is not sufficient to deal with Lipschitz curves, because nothing ensures that their derivatives at 0 exist. In order to somehow enforce this derivative to exist in a non-smooth setting, we will take advantage of ideas coming from the gradient flow theory: notice that from (1.7) it is not difficult to see that for any $g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and $\boldsymbol{\pi}$ as in (1.7) it holds

$$
\varlimsup_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \frac{1}{p} \int |Dg|_w^p(\gamma_0) \, \mathrm{d}\boldsymbol{\pi}(\gamma) + \frac{1}{q} \varlimsup_{t \downarrow 0} \frac{1}{t} \iint_0^t |\dot{\gamma}_t|^q \, \mathrm{d}t \mathrm{d}\boldsymbol{\pi}(\gamma).
\tag{1.9}
$$

Therefore if for some $g, \boldsymbol{\pi}$ it happens that

$$
\varliminf_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \geq \frac{1}{p} \int |Dg|_w^p(\gamma_0) \, \mathrm{d}\boldsymbol{\pi}(\gamma) + \frac{1}{q} \varliminf_{t \downarrow 0} \frac{1}{t} \iint_0^t |\dot{\gamma}_t|^q \, \mathrm{d}t \mathrm{d}\boldsymbol{\pi}(\gamma),
\tag{1.10}
$$

it means that the curves where $\boldsymbol{\pi}$ is concentrated 'indicate the direction of maximal increase of g at time $t = 0$'. In a smooth setting, this would mean $\gamma_0' = \nabla g(\gamma_0)$ for $\boldsymbol{\pi}$-a.e. γ. We then use (1.10) as definition and say that if it holds, then $\boldsymbol{\pi}$ is a plan which q-represents ∇g. Arguing exactly as at the end of the Section (1.1), we get the following crucial result: for $f, g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and $\boldsymbol{\pi}$ which q-represents ∇g it holds

$$
\begin{aligned}
\int D^+ f(\nabla g) |Dg|_w^{p-2} \, \mathrm{d}\boldsymbol{\pi}(\gamma_0) &\geq \varlimsup_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma), \\
\int D^- f(\nabla g) |Dg|_w^{p-2} \, \mathrm{d}\boldsymbol{\pi}(\gamma_0) &\leq \varliminf_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma).
\end{aligned}
\tag{1.11}
$$

6

Notice that if the space is q-infinitesimally strictly convex, then (1.11) tells that the limit $\lim_{t\downarrow 0}\int\frac{f(\gamma_t)-f(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}(\gamma)$ exists and in particular the map

$$\mathrm{S}^p(X,\mathsf{d},\mathfrak{m})\ni f \qquad \mapsto \qquad \lim_{t\downarrow 0}\int\frac{f(\gamma_t)-f(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}(\gamma),$$

is well defined and linear, so that we can really think at plans $\boldsymbol{\pi}$ representing gradients as differentiation operators (see also Appendix A).

The non-trivial task here is to prove that plans q-representing gradients exist. This is achieved by calling into play optimal transport as key tool: we will combine gradient flow theory, a powerful duality principle due to Kuwada [26] and a general superposition principle due to Lisini [28], see the proof of Lemma 3.13. Such existence result as well as part of the computations done in this chapter were already present, although in an embryonal form and for some special class of spaces, in [6].

Coming back to the Leibniz rule, it will be easy to show - essentially as a consequence of its validity on the real line - that for $\boldsymbol{\pi}$ which q-represents ∇g, $g\in\mathrm{S}^p(X,\mathsf{d},\mathfrak{m})$ and non-negative $f_1,f_2\in\mathrm{S}^p(X,\mathsf{d},\mathfrak{m})\cap L^\infty(X,\mathfrak{m})$ it holds

$$\overline{\lim_{t\downarrow 0}}\int\frac{f_1(\gamma_t)f_2(\gamma_t)-f_1(\gamma_0)f_2(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}(\gamma)\leq\overline{\lim_{t\downarrow 0}}\int f_1(\gamma_0)\frac{f_2(\gamma_t)-f_2(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}(\gamma)$$
$$+\overline{\lim_{t\downarrow 0}}\int f_2(\gamma_0)\frac{f_1(\gamma_t)-f_1(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}(\gamma).$$

From this inequality and a little bit of work, taking advantage of (1.11) it will then be possible to show that the Leibniz rule holds in the following general form:

$$D^+(f_1f_2)(\nabla g)\leq f_1 D^{s_1}f_2(\nabla g)+f_2 D^{s_2}f_1(\nabla g),$$
$$D^-(f_1f_2)(\nabla g)\geq f_1 D^{-s_1}f_2(\nabla g)+f_2 D^{-s_2}f_1(\nabla g),$$

$\mathfrak{m}$-a.e., where $s_i:=\mathrm{sign}f_i$, $i=1,2$.

Chapter 4 Once a rigorous meaning to $D^\pm f(\nabla g)$ is given, the definition of Laplacian can be proposed along the same lines of (1.4). We say that $g:X\to\mathbb{R}$ is in the domain of the Laplacian, and write $g\in D(\boldsymbol{\Delta})$, provided it belongs to $\mathrm{S}^p(X,\mathsf{d},\mathfrak{m})$ for some $p\in(1,\infty)$ and there exists a Radon measure μ such that

$$-\int D^+f(\nabla g)\,\mathrm{d}\mathfrak{m}\leq\int f\,\mathrm{d}\mu\leq-\int D^-f(\nabla g)\,\mathrm{d}\mathfrak{m},\qquad(1.12)$$

holds for any $f:X\to\mathbb{R}$ Lipschitz, in $L^1(X,|\mu|)$ and with support bounded and of finite $\mathfrak{m}$-measure - see Definition 4.4. In this case we write $\mu\in\boldsymbol{\Delta}g$ (we use the bold character to emphasize that we will always think of the Laplacian as a measure). The calculus tools developed in Chapter 3 allow to prove that also for this general potentially non-linear and multivalued notion of Laplacian, a reasonable calculus can be developed. For instance, the chain rule

$$\boldsymbol{\Delta}(\varphi\circ g)\supset\varphi'\circ g\,\boldsymbol{\Delta}g+\varphi''\circ g|Dg|_w^2\,\mathfrak{m},$$

holds under very general assumptions on g,φ (see Propositions 4.11 and 4.28).

A particularly important class of spaces is that of 'infinitesimally Hilbertian spaces', i.e. those such that $W^{1,2}(X,\mathsf{d},\mathfrak{m})$ is an Hilbert space. In these spaces, the Laplacian is unique and

linearly depends on g, at least for $g \in \mathrm{S}^2(X, \mathsf{d}, \mathfrak{m})$. The key consequence of this assumption is that, as already noticed in [6], not only the space is 2 infinitesimally strictly convex, so that $Df(\nabla g)$ is well defined for $f, g \in \mathrm{S}^2(X, \mathsf{d}, \mathfrak{m})$, but also

$$Df(\nabla g) = Dg(\nabla f), \qquad \mathfrak{m} - a.e. \qquad \forall f, g \in \mathrm{S}^2(X, \mathsf{d}, \mathfrak{m}).$$

This identity is the non-smooth analogous of the fact that on Riemannian framework we can identify differentials and gradients via Riesz theorem. This common value will be denoted by $\nabla f \cdot \nabla g$ to emphasize its symmetry. An important consequence of such symmetry is that the Leibniz rule for the Laplacian

$$\boldsymbol{\Delta}(g_1 g_2) = g_1 \boldsymbol{\Delta} g_2 + g_2 \boldsymbol{\Delta} g_1 + 2 \nabla g_1 \cdot \nabla g_2 \mathfrak{m},$$

holds, see Theorem 4.29.

On a smooth setting, a completely different approach to a definition of distributional Laplacian is to look at the short time behavior of the heat flow. It is therefore natural to question whether such approach is doable on abstract spaces and how it compares with the one we just proposed. Certainly, in order to do this, one has to deal with spaces where a 'well-behaved' heat flow exists. We will analyze this approach, and show that under pretty general assumptions it produces the same notion as the one coming from integration by parts, on spaces with Riemannian Ricci curvature bounded from below. This class of spaces, introduced in [6], is the class of infinitesimally Hilbertian $\mathsf{CD}(K, \infty)$ spaces, see Definition 4.30.

Chapter 5 Our task is now to prove that the sharp Laplacian comparison estimates for the distance function are true in abstract $\mathsf{CD}(K, N)$ spaces. The problem here is that the typical proof of these comparison results relies either on Jacobi fields calculus, or on the Bochner inequality. None of the two tools are available, at least as of today, in the non-smooth setting.

The only information that we have is the one coming from the $\mathsf{CD}(K, N)$ condition. In order to illustrate how to get the Laplacian comparison out of this, let us assume for a moment that we are dealing with a smooth Finsler manifold $(F, \mathsf{d}, \mathfrak{m})$ that satisfies the $\mathsf{CD}(0, N)$ condition and such that the norms in the tangent spaces are all strictly convex. The $\mathsf{CD}(0, N)$ condition tells, in this case, that for any $\mu_0, \mu_1 \in \mathscr{P}(F)$ with bounded support the unique (constant speed and minimizing) geodesic (μ_t) on the Wasserstein space $(\mathscr{P}_2(F), W_2)$ connecting μ_0 to μ_1 satisfies

$$-\int \rho_t^{1-\frac{1}{N}} \, \mathrm{d}\mathfrak{m} \leq -(1-t) \int \rho_0^{1-\frac{1}{N}} \, \mathrm{d}\mathfrak{m} - t \int \rho_1^{1-\frac{1}{N}} \, \mathrm{d}\mathfrak{m}, \qquad \forall t \in [0, 1], \qquad (1.13)$$

where ρ_t is the density of the absolutely continuous part of μ_t w.r.t. $\mathfrak{m}$. Fix $x_0 \in F$, let $\varphi(x) := \frac{\mathsf{d}^2(x, x_0)}{2}$ and let ρ_0 be an arbitrary smooth probability density with compact support. Put $\mu := \rho_0 \mathfrak{m}$ and notice that the only geodesic connecting $\mu_0 := \mu$ to $\mu_1 := \delta_{x_0}$ is given by $\mu_t := (\exp(-t \nabla \varphi))_\sharp \mu$. Hence, letting ρ_t be the density of μ_t, by explicit computation we get

$$\lim_{t \downarrow 0} \frac{-\int \rho_t^{1-\frac{1}{N}} \, \mathrm{d}\mathfrak{m} + \int \rho_0^{1-\frac{1}{N}} \, \mathrm{d}\mathfrak{m}}{t} = -\frac{1}{N} \int D\big(\rho^{1-\frac{1}{N}}\big)(\nabla \varphi) \, \mathrm{d}\mathfrak{m}. \qquad (1.14)$$

8

On the other hand, (1.13) gives

$$\varlimsup_{t\downarrow 0} \frac{-\int \rho_t^{1-\frac{1}{N}}\,\mathrm{d}\mathfrak{m} + \int \rho_0^{1-\frac{1}{N}}\,\mathrm{d}\mathfrak{m}}{t} \leq \int \rho_0^{1-\frac{1}{N}}\,\mathrm{d}\mathfrak{m}, \qquad (1.15)$$

so that it holds

$$\int \rho_0^{1-\frac{1}{N}}\,\mathrm{d}\mathfrak{m} \geq -\frac{1}{N}\int D\big(\rho_0^{1-\frac{1}{N}}\big)(\nabla\varphi)\,\mathrm{d}\mathfrak{m}.$$

Given that ρ_0 is arbitrary, non-negative and was chosen independently on φ, we get $\Delta\varphi \leq N$, which - up to a chain rule - is (1.1). For bounds on the Ricci curvature different than 0, (1.13) is modified into a distorted geodesic convexity of $\rho \mapsto -\int \rho^{1-\frac{1}{N}}\,\mathrm{d}\mathfrak{m}$, and the very same computations give the general sharp Laplacian comparison estimate (we also remark that on a smooth setting, taking two derivatives of the internal energy instead than one leads, via the $\mathsf{CD}(K,N)$ condition, to the Bochner inequality - see [20]).

The difficulty in proceeding this way in a non-smooth setting lies in proving that (1.14) holds, at least with $\geq$ replacing $=$. Notice indeed that we don't have any sort of change of variable formula. The idea to overcome this issue, which was also used in [6], is to first notice - as in the proof of Proposition 3.36 of [30] - that the convexity of $z \mapsto u_N(z) := -z^{1-\frac{1}{N}}$ yields

$$\frac{-\int \rho_t^{1-\frac{1}{N}}\,\mathrm{d}\mathfrak{m} + \int \rho_0^{1-\frac{1}{N}}\,\mathrm{d}\mathfrak{m}}{t} \geq \frac{1}{t}\int u_N'(\rho_0)(\rho_t - \rho_0)\,\mathrm{d}\mathfrak{m} = \frac{1}{t}\int u_N'(\rho_0)\circ\mathrm{e}_t - u_N'(\rho_0)\circ\mathrm{e}_0\,\mathrm{d}\boldsymbol{\pi},$$

where $\boldsymbol{\pi} \in \mathscr{P}(C([0,1],X))$ is such that $(\mathrm{e}_t)_\sharp\boldsymbol{\pi} = \mu_t$ for any $t \in [0,1]$ and $\iint_0^1 |\dot\gamma_t|^2\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi}(\gamma) = W_2^2(\mu_0,\mu_1)$. Then, taking advantage of Cheeger's results in [13], one can see that $\boldsymbol{\pi}$ 2-represents $\nabla(-\varphi)$ (see Corollary 5.8 for the rigorous statement), hence the limit as $t \downarrow 0$ of the last term in the previous expression can be estimated using (1.11) to get

$$\varliminf_{t\downarrow 0} \frac{1}{t}\int u_N'(\rho_0)\circ\mathrm{e}_t - u_N'(\rho_0)\circ\mathrm{e}_0\,\mathrm{d}\boldsymbol{\pi} \geq \int D^-(u_N'(\rho))(\nabla(-\varphi))\rho_0\,\mathrm{d}\mathfrak{m}$$

$$= -\frac{1}{N}\int D^+\big(\rho_0^{1-\frac{1}{N}}\big)(\nabla\varphi)\,\mathrm{d}\mathfrak{m}.$$

Using (1.15), which remains valid in the abstract setting, we then get

$$-\int D^+\big(\rho_0^{1-\frac{1}{N}}\big)(\nabla\varphi)\,\mathrm{d}\mathfrak{m} \leq N\int \rho_0^{1-\frac{1}{N}}\,\mathrm{d}\mathfrak{m}, \qquad (1.16)$$

this inequality being true for any Lipschitz probability density ρ_0 with compact support (see the proof of Theorem 5.14 for the rigorous justification). Notice that from (1.16) and the definition (1.12) we could deduce that $\boldsymbol{\Delta}\varphi \leq N\mathfrak{m}$ if there was $D^-\big(\rho_0^{1-\frac{1}{N}}\big)$ in place of $D^+\big(\rho_0^{1-\frac{1}{N}}\big)$ in the left hand side. This means that at least for spaces which are infinitesimally strictly convex for some $q \in (1,\infty)$, where $D^+\big(\rho_0^{1-\frac{1}{N}}\big)(\nabla\varphi) = D^-\big(\rho_0^{1-\frac{1}{N}}\big)(\nabla\varphi)$, we can conclude that the Laplacian comparison holds, in line with what is known in the smooth Finsler setting (see [34]).

9

2 Preliminaries

2.1 Metric spaces and Wasserstein distance

All the metric spaces (X, d) considered will be complete and separable. The open ball of center $x \in X$ and radius $r > 0$ is denoted by $B_r(x)$. $C([0,1], X)$ is the space of continuous curves from $[0,1]$ to X equipped with the sup norm, which is complete and separable. For $t \in [0,1]$, the evaluation map $\mathsf{e}_t : C([0,1], X) \to X$ is given by

$$\mathsf{e}_t(\gamma) := \gamma_t, \qquad \forall \gamma \in C([0,1], X).$$

For $t, s \in [0,1]$, the restriction operator $\mathrm{restr}_t^s : C([0,1], X) \to C([0,1], X)$ is defined by

$$(\mathrm{restr}_t^s(\gamma))_r := \gamma_{t+r(s-t)}, \qquad \forall \gamma \in C([0,1], X). \tag{2.1}$$

So that for $t < s$, restr_t^s restricts the curve γ to the interval $[t, s]$ and then 'stretches' it on the whole $[0,1]$, for $t > s$ also a change in the orientation occurs.

A curve $\gamma \in C([0,1], X)$ is said absolutely continuous provided there exists $f \in L^1([0,1])$ such that

$$\mathsf{d}(\gamma_t, \gamma_s) \leq \int_t^s f(r) \, \mathrm{d}r, \qquad \forall t, s \in [0,1], \ t < s. \tag{2.2}$$

The set of absolutely continuous curves from $[0,1]$ to X will be denoted by $AC([0,1], X)$. More generally, if the function f in (2.2) belongs to $L^q([0,1])$, $q \in [1, \infty]$, γ is said q-absolutely continuous, and $AC^q([0,1], X)$ is the corresponding set of q-absolutely continuous curves. It turns out (see for instance Theorem 1.1.2 of [3]) that for $\gamma \in AC^q([0,1], X)$ the limit

$$\lim_{h \to 0} \frac{\mathsf{d}(\gamma_{t+h}, \gamma_t)}{|h|},$$

exists for a.e. $t \in [0,1]$, and defines an L^q function. Such function is called metric speed or metric derivative, is denoted by $|\dot{\gamma}_t|$ and is the minimal (in the a.e. sense) L^q function which can be chosen as f in the right hand side of (2.2).

(X, d) is said to be a length space provided for any $x, y \in X$ it holds

$$\mathsf{d}(x, y) = \inf \int_0^1 |\dot{\gamma}_t| \, \mathrm{d}s,$$

the infimum being taken among all $\gamma \in AC([0,1], X)$ such that $\gamma_0 = x$, $\gamma_1 = y$. If the infimum is always a minimum, then the space is said geodesic, and a geodesic from x to y is any minimizer which is parametrized by constant speed (they are sometime referred as constant speed and minimizing geodesics). Equivalently, γ is a geodesic from x to y provided

$$\mathsf{d}(\gamma_s, \gamma_t) = |s - t|\mathsf{d}(\gamma_0, \gamma_1), \qquad \forall t, s \in [0,1], \qquad \gamma_0 = x, \qquad \gamma_1 = y.$$

The space of all geodesics on X will be denoted by $\mathrm{Geo}(X)$. It is a closed subset of $C([0,1], X)$.

Given Borel functions $f : X \to \mathbb{R}$, $G : X \to [0, \infty]$ we say that G is an upper gradient of f provided

$$|f(\gamma_1) - f(\gamma_0)| \leq \int_0^1 G(\gamma_t)|\dot{\gamma}_t| \, \mathrm{d}t, \qquad \forall \gamma \in AC([0,1], X).$$

For $f : X \to \mathbb{R}$ the local Lipschitz constant $\mathrm{lip}(f) : X \to [0, \infty]$ is defined by

$$\mathrm{lip}(f)(x) := \overline{\lim_{y \to x}} \frac{|f(y) - f(x)|}{\mathsf{d}(y, x)}, \quad \text{if } x \text{ is not isolated, } 0 \text{ otherwise.} \tag{2.3}$$

The one sided analogous $\mathrm{lip}^+(f), \mathrm{lip}^-(f)$, defined by

$$\mathrm{lip}^+(f)(x) := \overline{\lim_{y \to x}} \frac{(f(y) - f(x))^+}{\mathsf{d}(y, x)}, \qquad \mathrm{lip}^-(f)(x) := \overline{\lim_{y \to x}} \frac{(f(y) - f(x))^-}{\mathsf{d}(y, x)}, \tag{2.4}$$

if x is not isolated and 0 otherwise, are called ascending and descending slope respectively, where $z^+ := \max\{z, 0\}$, $z^- := \max\{-z, 0\}$. It is easy to check that if f is locally Lipschitz, then $\mathrm{lip}^\pm(f), \mathrm{lip}(f)$ are all upper gradients of f.

The space of all Lipschitz real valued functions on X will be denoted by $\mathrm{LIP}(X)$.

Given a Borel measure σ on X, by $\mathrm{supp}(\sigma)$ we intend the smallest closed set where σ is concentrated.

The set of Borel probability measures on X will be denoted by $\mathscr{P}(X)$. For $q \in [1, \infty)$, $\mathscr{P}_q(X) \subset \mathscr{P}(X)$ is the set of measures with finite q-moment, i.e. $\mu \in \mathscr{P}_q(X)$ if $\mu \in \mathscr{P}(X)$ and $\int \mathsf{d}^q(x, x_0) \, \mathrm{d}\mu(x) < \infty$ for some (and thus every) $x_0 \in X$. $\mathscr{P}_q(X)$ is endowed with the q-Wasserstein distance W_q, defined by

$$W_q^q(\mu, \nu) := \inf_{\boldsymbol{\gamma}} \int \mathsf{d}^q(x, y) \, \mathrm{d}\boldsymbol{\gamma}(x, y),$$

where the inf is taken among all transport plans $\boldsymbol{\gamma}$, i.e. among all $\boldsymbol{\gamma} \in \mathscr{P}(X \times X)$ such that $\pi^1_\sharp \boldsymbol{\gamma} = \mu$, $\pi^2_\sharp \boldsymbol{\gamma} = \nu$.

We recall the following superposition result, proved by Lisini in [28] as a generalization of an analogous result proved in [3] in the Euclidean context.

Theorem 2.1 (Lisini) *Let (X, d) be a complete and separable metric space, $q \in (1, \infty)$, and $[0, 1] \ni t \mapsto \mu_t \in \mathscr{P}_q(X)$ a q-absolutely continuous curve w.r.t. W_q. Then there exists a measure $\boldsymbol{\pi} \in \mathscr{P}(C([0, 1], X))$ concentrated on $AC^q([0, 1], X)$ such that*

$$(\mathrm{e}_t)_\sharp \boldsymbol{\pi} = \mu_t, \qquad \forall t \in [0, 1],$$
$$\int |\dot{\gamma}_t|^q \, \mathrm{d}\boldsymbol{\pi}(\gamma) = |\dot{\mu}_t|^q, \qquad a.e. \ t. \tag{2.5}$$

We remark that for a plan $\boldsymbol{\pi} \in \mathscr{P}(C([0, 1], X))$ such that $\iint_0^1 |\dot{\gamma}_t|^q \, \mathrm{d}t \, \mathrm{d}\boldsymbol{\pi}(\gamma) < \infty$ and satisfying $(\mathrm{e}_t)_\sharp \boldsymbol{\pi} = \mu_t$ for any $t \in [0, 1]$, it always holds $\int |\dot{\gamma}_t|^q \, \mathrm{d}\boldsymbol{\pi}(\gamma) \geq |\dot{\mu}_t|^q$ for a.e. $t \in [0, 1]$, so that Lisini's result is of variational nature, as it selects plans with minimal energy.

Although we will have to work with general q-Wasserstein distances, we will reserve to the special case $q = 2$ the wording of 'c-concavity', 'Kantorovich potential' and 'optimal geodesic plan'. Thus the c-transform $\varphi^c : X \to \mathbb{R} \cup \{-\infty\}$ of a map $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ is defined by

$$\varphi^c(y) = \inf_{x \in X} \frac{\mathsf{d}^2(x, y)}{2} - \varphi(x).$$

A function $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ is called c-concave provided it is not identically $-\infty$ and it holds $\varphi = \psi^c$ for some ψ. Notice that for a generic function $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ it always holds $\varphi^{ccc} = \varphi^c$ and

$$\varphi^{cc} \geq \varphi, \tag{2.6}$$

thus φ is c-concave if and only if $\varphi^{cc} \leq \varphi$. The c-superdifferential $\partial^c \varphi$ of the c-concave map φ is the subset of X^2 defined by

$$(x,y) \in \partial^c \varphi \qquad \Leftrightarrow \qquad \varphi(x) + \varphi^c(y) = \frac{\mathsf{d}^2(x,y)}{2}.$$

The set $\partial^c \varphi(x) \subset X$ is the set of those y's such that $(x,y) \in \partial^c \varphi$.

Given $\mu, \nu \in \mathscr{P}_2(X)$, a Kantorovich potential from μ to ν is any c-concave map φ such that

$$\frac{1}{2} W_2^2(\mu, \nu) = \int \varphi \, \mathrm{d}\mu + \int \varphi^c \, \mathrm{d}\nu.$$

The set $\mathrm{OptGeo}(\mu, \nu) \subset \mathscr{P}(\mathrm{Geo}(X))$ is the set of those $\boldsymbol{\pi}$'s such that

$$W_2\big((\mathrm{e}_s)_\sharp \boldsymbol{\pi}, (\mathrm{e}_t)_\sharp \boldsymbol{\pi}\big) = |s-t| W_2\big((\mathrm{e}_1)_\sharp \boldsymbol{\pi}, (\mathrm{e}_0)_\sharp \boldsymbol{\pi}\big), \qquad \forall t,s \in [0,1], \qquad (\mathrm{e}_0)_\sharp \boldsymbol{\pi} = \mu, \quad (\mathrm{e}_1)_\sharp \boldsymbol{\pi} = \nu.$$

Recall that $(\mu_t) \subset \mathscr{P}_2(X)$ is a constant speed geodesic if and only if for some $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu_0, \mu_1)$ it holds

$$(\mathrm{e}_t)_\sharp \boldsymbol{\pi} = \mu_t, \qquad \forall t \in [0,1].$$

We conclude with the following simple result, whose proof is extracted from [18]. We remind that (X, d) is said proper, provided bounded closed sets are compact.

Proposition 2.2 *Let (X, d) be a proper geodesic space and $\varphi : X \to \mathbb{R}$ a locally Lipschitz c-concave function. Then for every $x \in X$ the set $\partial^c \varphi(x)$ is non-empty, and $\cup_{x \in K} \partial^c \varphi(x)$ is compact for any compact set $K \subset X$.*

proof Fix $x \in X$ and notice that since φ is c-concave, there exists a sequence $(y^n) \subset X$ such that

$$\varphi(x) = \lim_{n \to \infty} \frac{\mathsf{d}^2(x, y^n)}{2} - \varphi^c(y^n),$$
$$\varphi(z) \leq \frac{\mathsf{d}^2(z, y^n)}{2} - \varphi^c(y^n), \qquad \forall n \in \mathbb{N}, \ z \in X.$$

Now let $\gamma^n : [0, \mathsf{d}(x, y^n)] \to X$ be a unit speed geodesic connecting x to y^n, choose $z = \gamma_1^n$ in the inequality above and subtract $\varphi(x)$ to get

$$\varliminf_{n \to \infty} \varphi(\gamma_1^n) - \varphi(x) \leq \varliminf_{n \to \infty} \frac{\mathsf{d}^2(\gamma_1^n, y^n)}{2} - \frac{\mathsf{d}^2(x, y^n)}{2} = \varliminf_{n \to \infty} -\mathsf{d}(x, y^n) + \frac{1}{2}.$$

The set $\{\gamma_1^n\}_n$ is bounded by construction, hence, since φ is Lipschitz on bounded sets, the left hand side of the inequality is finite, which forces $\overline{\lim}_{n \to \infty} \mathsf{d}(x, y_n) < \infty$. Thus, since the space is proper, the sequence (y^n) converges, up to pass to subsequences, to some $y \in X$, which can be easily seen to belong to $\partial^c \varphi(x)$.

This argument also gives the quantitative estimate $\mathsf{d}(x, y) \leq \mathrm{Lip}(\varphi|_{B_1(x)}) + \frac{1}{2}$ for $y \in \partial^c \varphi(x)$. Therefore for a compact $K \subset X$, the set $\cup_{x \in K} \partial^c \varphi(x)$ is bounded. Since it can be easily checked to be closed, the proof is finished. $\qquad \square$

2.2 Metric measure spaces

The main object of investigation in this paper are metric measure spaces $(X, \mathsf{d}, \mathfrak{m})$, which will always be of the following kind

$$(X, \mathsf{d}) \text{ is a complete separable metric space,}$$
$$\mathfrak{m} \text{ is a Radon non-negative measure on } X, \tag{2.7}$$

where by local finiteness we mean that for every $x \in X$ there exists a neighborhood U_x of x such that $\mathfrak{m}(U_x) < \infty$.

From the fact that $\mathfrak{m}$ is locally finite and (X, d) separable, using the Lindelöf property it follows that there exists a Borel probability measure $\tilde{\mathfrak{m}} \in \mathscr{P}(X)$ such that

$$\mathfrak{m} \ll \tilde{\mathfrak{m}} \leq C\mathfrak{m}, \qquad \text{for some constant } C,$$
$$\text{with } \frac{\mathrm{d}\tilde{\mathfrak{m}}}{\mathrm{d}\mathfrak{m}} \text{ locally bounded from below by a positive constant,} \tag{2.8}$$

where again by locally bounded from below we intend that for any $x \in X$ there exists a neighborhood U_x and a constant $c_x > 0$ such that $\mathfrak{m}$-a.e. on U_x it holds $\frac{\mathrm{d}\tilde{\mathfrak{m}}}{\mathrm{d}\mathfrak{m}} \geq c_x$. It is also not difficult to see that $\tilde{\mathfrak{m}}$ can be chosen so that

$$\int \mathsf{d}^q(x, x_0) \, \mathrm{d}\tilde{\mathfrak{m}} < \infty, \qquad \forall q \geq 1, \text{ for some, and thus any, } x_0 \in X. \tag{2.9}$$

We fix once and for all such measure $\tilde{\mathfrak{m}}$: its use will be useful to get some compactness in the proof of Theorem 3.14 (in particular, in Lemma 3.13). Yet, we remark that $\tilde{\mathfrak{m}}$ is not really part of our data, in the sense that none of our results depend on the particular measure $\tilde{\mathfrak{m}}$ satisfying (2.8) and (2.9).

2.3 Sobolev classes

In this section we recall the definition of Sobolev classes $\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$, which are the metric-measure analogous of the space of functions having distributional gradient in L^p when our space is the Euclidean one, regardless of any integrability assumption on the functions themselves. Definitions are borrowed from [4], [5], but the presentation proposed here is slightly different: see Appendix B for the simple proof that the two approaches are actually the same. The basic properties of the Sobolev classes are recalled mainly without proofs, we refer to [5] for a detailed discussion (see also [1] for the case $p = 1$).

Definition 2.3 (q-test plan) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) and $\boldsymbol{\pi} \in \mathscr{P}(C([0, 1], X))$. We say that $\boldsymbol{\pi}$ has bounded compression provided there exists $C > 0$ such that*

$$(\mathrm{e}_t)_\sharp \boldsymbol{\pi} \leq C\mathfrak{m}, \qquad \forall t \in [0, 1].$$

For $q \in (1, \infty)$ we say that $\boldsymbol{\pi}$ is a q-test plan if it has bounded compression, is concentrated on $AC^q([0, 1], X)$ and

$$\iint_0^1 |\dot{\gamma}_t|^q \, \mathrm{d}t \, \mathrm{d}\boldsymbol{\pi}(\gamma) < \infty.$$

13

Definition 2.4 (Sobolev classes) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p \in (1, \infty)$ and q the conjugate exponent. A Borel function $f : X \to \mathbb{R}$ belongs to the Sobolev class $\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ (resp. $\mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$) provided there exists a function $G \in L^p(X, \mathfrak{m})$ (resp. in $L^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$) such that*

$$\int \big| f(\gamma_1) - f(\gamma_0) \big| \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \iint_0^1 G(\gamma_s) |\dot{\gamma}_s| \, \mathrm{d}s \, \mathrm{d}\boldsymbol{\pi}(\gamma), \qquad \forall q\text{-test plan } \boldsymbol{\pi}. \qquad (2.10)$$

In this case, G is called a p-weak upper gradient of f.

Since the class of q-test plans contains the one of q'-test plans for $q \leq q'$, we have that $\mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}) \subset \mathrm{S}^{p'}_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ for $p \geq p'$, and that if $f \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ and G is a p-weak upper gradient, then G is also a p'-weal upper gradient.

Arguing as in Section 4.5 of [5], we get that for $f \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ (resp. $\mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$) there exists a minimal function $G \geq 0$, in the $\mathfrak{m}$-a.e. sense, in $L^p(X, \mathfrak{m})$ (resp. $L^p_{\mathrm{loc}}(X, \mathfrak{m})$) such that (2.10) holds. We will denote such minimal function by $|Df|_w$. In line with the terminology used in this context, we will refer to $|Df|_w$ as the p-minimal upper gradient of f. Yet, given that such quantity is defined in duality with speed of curves, it is closer in spirit to the dual norm of the differential rather than to the norm of the gradient, whence the notation with a 'D' in place of a '∇'.

As for the local Lipschitz constant and the slopes, we switched from the notation $|\nabla f|_w$ used in [4] and [5] to $|Df|_w$ to underline that this is a cotangent notion, a fact which plays a key role in our analysis.

Remark 2.5 Observe that in denoting the minimal p-weak upper 'gradient' by $|Df|_w$, we are losing the reference to the exponent p, which plays a role in the definition as it affects the class of test plans. A more appropriate notation would be $|Df|_{w,p}$. With this notation we have that for $f \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and $p' \leq p$ it holds $|Df|_{w,p'} \leq |Df|_{w,p}$ $\mathfrak{m}$-a.e.. It is a longstanding open problem to understand whether the other inequality holds.

To drop the dependence on p on the notation is certainly risky, but in all our statements we always recall at first to which Sobolev class our functions are belonging, thus indicating which 'p' is chosen in the definition of p-minimal upper gradient and hopefully minimizing the risk of confusion.

We also recall that as a consequence of the results of Cheeger [13], we have that if the measure is doubling and the space supports a weak-local 1-p' Poincaré inequality, then for $f \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$, $p \geq p'$, we have $|Df|_{w,p'} = |Df|_{w,p}$ $\mathfrak{m}$-a.e., thus at least in this case there is no dependence on the Sobolev exponent. ∎

Notice that if $\boldsymbol{\pi}$ is a q-test plan and $\Gamma \subset C([0, 1], X)$ is a Borel set such that $\boldsymbol{\pi}(\Gamma) > 0$, then also the plan $\boldsymbol{\pi}(\Gamma)^{-1}\boldsymbol{\pi}_{|\Gamma}$ is a q-test plan. Also, for any $t, s \in [0, 1]$ the plan $(\mathrm{restr}_t^s)_\sharp \boldsymbol{\pi}$ is a q-test plan as well (recall definition (2.1)). Then a simple localization argument yields that from (2.10) it follows that

$$\forall t < s \in [0, 1] \quad \text{it holds} \qquad |f(\gamma_s) - f(\gamma_t)| \leq \int_t^s |Df|_w(\gamma_r)|\dot{\gamma}_r| \, \mathrm{d}r, \qquad \boldsymbol{\pi} - a.e. \ \gamma, \quad (2.11)$$

whenever $\boldsymbol{\pi}$ is a q-test plan and $f \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$.

It is obvious that $\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and $\mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ are vector spaces and that it holds

$$|D(\alpha f + \beta g)|_w \leq |\alpha||Df|_w + |\beta||Dg|_w, \qquad \mathfrak{m} - a.e., \quad \forall \alpha, \beta \in \mathbb{R}, \qquad (2.12)$$

and that the spaces $S^p(X, \mathsf{d}, \mathfrak{m}) \cap L^\infty(X, \mathfrak{m})$ and $S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}) \cap L^\infty_{\mathrm{loc}}(X, \mathfrak{m})$ are algebras, for which it holds

$$|D(fg)|_w \le |f||Dg|_w + |g||Df|_w, \qquad \mathfrak{m} - a.e.. \tag{2.13}$$

It is also possible to check that the object $|Df|_w$ is local in the sense that

$$\forall f \in S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}) \text{ it holds } |Df|_w = 0, \quad \mathfrak{m} - a.e. \text{ on } f^{-1}(\mathcal{N}), \quad \forall \mathcal{N} \subset \mathbb{R}, \text{ s.t. } \mathcal{L}^1(\mathcal{N}) = 0, \tag{2.14}$$

and

$$|Df|_w = |Dg|_w, \qquad \mathfrak{m} - a.e. \text{ on } \{f = g\}, \qquad \forall f, g \in S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}). \tag{2.15}$$

Also, for $f \in S^p(X, \mathsf{d}, \mathfrak{m})$ (resp. $S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$) and $\varphi : \mathbb{R} \to \mathbb{R}$ Lipschitz, the function $\varphi \circ f$ belongs to $S^p(X, \mathsf{d}, \mathfrak{m})$ (resp. $S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$) as well and it holds

$$|D(\varphi \circ f)|_w = |\varphi' \circ f||Df|_w \qquad \mathfrak{m} - a.e.. \tag{2.16}$$

The object $|Df|_w$ is local also in the following sense (see Theorem 4.19 of [6] for the case $p = 2$ and Section 8.2 of [5] for the general one).

Proposition 2.6 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p \in (1, \infty)$ and $\Omega \subset X$ an open set. Then the following holds.*

i) For $f \in S^p(X, \mathsf{d}, \mathfrak{m})$ (resp. $f \in S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$), the restriction of f to $\overline{\Omega}$ belongs to $f \in S^p(\overline{\Omega}, \mathsf{d}, \mathfrak{m})$ (resp. $f \in S^p_{\mathrm{loc}}(\overline{\Omega}, \mathsf{d}, \mathfrak{m})$) and it holds

$$(|Df|_w)_X = (|Df|_w)_{\overline{\Omega}}, \qquad \mathfrak{m} - a.e. \text{ in } \Omega, \tag{2.17}$$

where by $(|Df|_w)_X$ (resp. $(|Df|_w)_{\overline{\Omega}}$) we are denoting the minimal p-weak upper gradient of f in the space $(X, \mathsf{d}, \mathfrak{m})$ (resp. of the restriction of f to $\overline{\Omega}$ in the space $(\overline{\Omega}, \mathsf{d}, \mathfrak{m}_{|\overline{\Omega}})$).

ii) Viceversa, if $f \in S^p(\overline{\Omega}, \mathsf{d}, \mathfrak{m})$ (resp. $S^p_{\mathrm{loc}}(\overline{\Omega}, \mathsf{d}, \mathfrak{m})$) and $\mathrm{supp}(f) \subset \Omega$ with $\mathsf{d}(\mathrm{supp}(f), X \setminus \Omega) > 0$, then extending f to the whole X by putting $f \equiv 0$ on $X \setminus \Omega$ we have $f \in S^p(X, \mathsf{d}, \mathfrak{m})$ (resp. $f \in S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$) and (2.17) holds.

proof In [6], (*ii*) was stated under the additional assumption that $\mathfrak{m}(\partial\Omega) = 0$ and then the conclusion (2.17) was given $\mathfrak{m}$-a.e. on $\overline{\Omega}$. To check that the statement is true also in the current version, just notice that from the local finiteness of $\mathfrak{m}$ we can easily build an open set $\Omega' \subset \Omega$ such that $\mathrm{supp}(f) \subset \Omega'$, $\mathsf{d}(\Omega', X \setminus \Omega) > 0$ and $\mathfrak{m}(\partial\Omega') = 0$. □

We endow $S^p(X, \mathsf{d}, \mathfrak{m})$ with the seminorm

$$\|f\|_{S^p(X, \mathsf{d}, \mathfrak{m})} := \||Df|_w\|_{L^p(X, \mathfrak{m})}. \tag{2.18}$$

We don't know if (the quotient of) $S^p(X, \mathsf{d}, \mathfrak{m})$ is complete w.r.t. $\|\cdot\|_{S^p}$.

The Sobolev space $W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ is defined as $W^{1,p}(X, \mathsf{d}, \mathfrak{m}) := S^p(X, \mathsf{d}, \mathfrak{m}) \cap L^p(X, \mathfrak{m})$ endowed with the norm

$$\|f\|^p_{W^{1,p}(X, \mathsf{d}, \mathfrak{m})} := \|f\|^p_{L^p(X, \mathfrak{m})} + \|f\|^p_{S^p(X, \mathsf{d}, \mathfrak{m})}.$$

$W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ is always a Banach space (but we remark that in general $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is *not* an Hilbert space). The completeness of $W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ is a consequence of the following result, whose proof follows from the definitions and Mazur's lemma.

Proposition 2.7 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p \in (1, \infty)$, $(f_n) \subset \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and $(G_n) \subset L^p(X, \mathfrak{m})$. Assume that G_n is a p-weak upper gradient of f_n for every $n \in \mathbb{N}$, that $f_n \to f$ pointwise $\mathfrak{m}$-a.e. and $G_n \rightharpoonup G$ in $L^p(X, \mathfrak{m})$ for some functions f, G.*

Then $f \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and G is a p-weak upper gradient of f.

In [4], [5] the following non-trivial result has been proved.

Theorem 2.8 (Density in energy of Lipschitz functions) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p \in (1, \infty)$ and assume that $\mathfrak{m}$ is finite on bounded sets. Then for every $f \in W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ there exists a sequence $(f_n) \subset W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ of Lipschitz functions such that*

$$\lim_{n \to \infty} \|f_n - f\|_{L^p(X, \mathfrak{m})} = 0,$$

$$\lim_{n \to \infty} \|f_n\|_{\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})} = \lim_{n \to \infty} \||Df_n|\|_{L^p(X, \mathfrak{m})} = \lim_{n \to \infty} \||\overline{Df_n}|\|_{L^p(X, \mathfrak{m})} = \|f\|_{\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})},$$

where given $g : X \to \mathbb{R}$, the function $\overline{|Dg|} : X \to [0, \infty]$ is 0 by definition on isolated points and

$$\overline{|Dg|}(x) := \inf_{r>0} \sup_{y_1 \neq y_2 \in B_r(x)} \frac{|g(y_1) - g(y_2)|}{\mathsf{d}(y_1, y_2)}.$$

on the others.

proof The proof given in [5] for $p \in (1, \infty)$ was based also on the assumption that $\mathfrak{m}$ was finite. If $\mathfrak{m}$ is only finite on bounded sets we can argue as follows. We pick a sequence (χ_n) of 1-Lipschitz functions with bounded support and values in $[0, 1]$ such that $\chi_n \equiv 1$ on $B_n(\bar{x})$, for some fixed $\bar{x} \in X$. Then we fix $f \in W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ and notice that by the dominate convergence theorem we have $f\chi_n \to f$ in $L^p(X, \mathfrak{m})$ as $n \to \infty$, by (2.15) that $|D(f\chi_n)|_w = |Df|_w$ $\mathfrak{m}$-a.e. on $B_n(\bar{x})$ and from (2.13) that $|D(f\chi_n)|_w \leq |Df|_w + |f|$ so that $\|f\chi_n\|_{\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})} \to \|f\|_{\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})}$ as $n \to \infty$ ((2.13) was stated for bounded functions, but it is true also with $f \in L^p(X, \mathfrak{m})$ if, as in our case, the other function is both bounded and Lipschitz). Thus $(f\chi_n)$ converges in energy in $W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ to f as $n \to \infty$. Since $f\chi_n$ is in $W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ and has bounded support, taking into account Proposition 2.6, the fact that $\mathfrak{m}$ is finite on bounded sets and the result in [5], a diagonalization argument gives the result. $\square$

In particular, the following holds:

Corollary 2.9 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) be such that $\mathfrak{m}$ is finite on bounded sets and $p \in (1, \infty)$. Assume that $W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ is uniformly convex. Then $\mathrm{LIP}(X) \cap W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ is dense in $W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ w.r.t. the $W^{1,p}$ norm.*

For later use, we collect here some basic duality properties between $\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and the set of q-test plans. It will be useful to introduce the q-energies of a curve $\gamma \in C([0, 1], X)$: for $q \in (1, \infty)$, and $t \in (0, 1]$, the q-energy $E_{q,t} : C([0, 1], X) \to [0, \infty]$ is defined by

$$E_{q,t}(\gamma) := \begin{cases} t \sqrt[q]{\dfrac{1}{t} \int_0^t |\dot{\gamma}_s|^q \, \mathrm{d}s}, & \text{if } \mathrm{restr}_0^t(\gamma) \in AC^q([0, 1], X), \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.19}$$

Proposition 2.10 *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be as in* (2.7)*,* $p \in (1, \infty)$ *and* q *the conjugate exponent,* $f \in S^p(X, \mathsf{d}, \mathfrak{m})$ *and* $\boldsymbol{\pi}$ *a* q-test plan. Then the following are true.

$$\left| \frac{f(\gamma_t) - f(\gamma_0)}{E_{q,t}} \right|^p \leq \frac{1}{t} \int_0^t |Df|_w^p(\gamma_s) \, \mathrm{d}s, \qquad \boldsymbol{\pi} - a.e. \ \gamma, \ \forall t \in [0, 1], \qquad (2.20)$$

$$\overline{\lim_{t \downarrow 0}} \int \left| \frac{f(\gamma_t) - f(\gamma_0)}{E_{q,t}(\gamma)} \right|^p \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \int |Df|_w^p(\gamma_0) \, \mathrm{d}\boldsymbol{\pi}(\gamma), \qquad (2.21)$$

$$\overline{\lim_{t \downarrow 0}} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \frac{1}{p} \int |Df|_w^p(\gamma_0) \, \mathrm{d}\boldsymbol{\pi}(\gamma) + \overline{\lim_{t \downarrow 0}} \frac{1}{qt} \iint_0^t |\dot{\gamma}_s|^q \, \mathrm{d}s \, \mathrm{d}\boldsymbol{\pi}(\gamma). \quad (2.22)$$

proof From (2.11) we obtain

$$\left| f(\gamma_t) - f(\gamma_0) \right| \leq \int_0^t |Df|_w(\gamma_s) |\dot{\gamma}_s| \, \mathrm{d}s \leq \sqrt[p]{\int_0^t |Df|_w^p(\gamma_s) \, \mathrm{d}s} \sqrt[q]{\int_0^t |\dot{\gamma}_s|^q \, \mathrm{d}s},$$

which is a restatement of (2.20). Integrate (2.20) w.r.t. $\boldsymbol{\pi}$ to get

$$\int \left| \frac{f(\gamma_t) - f(\gamma_0)}{E_{q,t}(\gamma)} \right|^p \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \frac{1}{t} \int_0^t \int |Df|_w^p \, \mathrm{d}(\mathrm{e}_s)_\sharp \boldsymbol{\pi} \, \mathrm{d}s = \int |Df|_w^p \left(\frac{1}{t} \int_0^t \rho_s \, \mathrm{d}s \right) \, \mathrm{d}\mathfrak{m},$$

where ρ_s is the density of $(\mathrm{e}_s)_\sharp \boldsymbol{\pi}$ w.r.t. $\mathfrak{m}$. Conclude observing that $|Df|_w^p \in L^1(X, \mathfrak{m})$, that ρ_s weakly converges to ρ_0 as $s \downarrow 0$ in duality with $C_b(X)$ and that $\|\rho_s\|_\infty$ is uniformly bounded to get weak convergence of ρ_s to ρ_0 in duality with $L^1(X, \mathfrak{m})$, thus it holds

$$\lim_{t \downarrow 0} \int |Df|_w^p \left(\frac{1}{t} \int_0^t \rho_s \, \mathrm{d}s \right) \, \mathrm{d}\mathfrak{m} = \int |Df|_w^p \, \mathrm{d}\mathfrak{m},$$

and (2.21) is proved. For (2.22) just notice that for $\boldsymbol{\pi}$-a.e. γ it holds

$$\frac{|f(\gamma_t) - f(\gamma_0)|}{t} = \frac{|f(\gamma_t) - f(\gamma_0)|}{E_{q,t}(\gamma)} \frac{E_{q,t}(\gamma)}{t} \leq \frac{1}{p} \left| \frac{f(\gamma_t) - f(\gamma_0)}{E_{q,t}(\gamma)} \right|^p + \frac{1}{q} \left| \frac{E_{q,t}(\gamma)}{t} \right|^q, \qquad (2.23)$$

then integrate w.r.t. $\boldsymbol{\pi}$ and use (2.21) in passing to the limit as $t \downarrow 0$. $\qquad\square$

Proposition 2.11 *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be as in* (2.7)*,* $p \in (1, \infty)$ *and* q *the conjugate exponent,* $f \in S^p(X, \mathsf{d}, \mathfrak{m})$ *and* $\boldsymbol{\pi}$ *a* q-test plan. Then

$$\lim_{t \downarrow 0} \sqrt[p]{\frac{1}{t} \int_0^t |Df|_w^p \circ \mathrm{e}_s \, \mathrm{d}s} = |Df|_w \circ \mathrm{e}_0, \qquad in \ L^p(\boldsymbol{\pi}). \qquad (2.24)$$

Furthermore, if the family of functions $\frac{E_{q,t}}{t}$ *is dominated in* $L^q(\boldsymbol{\pi})$*, then the family of functions* $\frac{f(\gamma_t) - f(\gamma_0)}{t}$ *is dominated in* $L^1(\boldsymbol{\pi})$*.*

proof To prove (2.24) it is sufficient to show that

$$\lim_{t \downarrow 0} \frac{1}{t} \int_0^t |Df|_w^p \circ \mathrm{e}_s \, \mathrm{d}s = |Df|_w^p \circ \mathrm{e}_0, \qquad in \ L^1(\boldsymbol{\pi}). \qquad (2.25)$$

To this aim, associate to a generic $h \in L^1(X, \mathfrak{m})$ the functions $H_t \in L^1(\pi)$, $t \in [0, 1]$, defined by

$$H_t := \frac{1}{t} \int_0^t h \circ e_s \, ds, \qquad H_0 := h \circ e_0,$$

and notice that if h is Lipschitz, the inequality $|H_t(\gamma) - H_0(\gamma)| \leq \mathrm{Lip}(h)\frac{1}{t}\int_0^t \mathsf{d}(\gamma_s, \gamma_0) \, ds \leq \mathrm{Lip}(h)\int_0^t |\dot\gamma_s| \, ds$ easily yields that $\|H_t - H_0\|_{L^1(\pi)} \to 0$ as $t \downarrow 0$. Now recall that Lipschitz functions are dense in $L^1(X, \mathfrak{m})$ and use the uniform continuity estimates

$$\int \frac{1}{t} \left| \int_0^t h(\gamma_s) \, ds \right| \, d\pi(\gamma) \leq \frac{1}{t} \iint_0^t |h(\gamma_s)| \, ds \, d\pi(\gamma) \leq C \|h\|_{L^1(X, \mathfrak{m})},$$

$$\int |h(\gamma_0)| \, d\pi(\gamma) \leq C \|h\|_{L^1(X, \mathfrak{m})},$$

where $C > 0$ is such that $(e_t)_\sharp \pi \leq C\mathfrak{m}$ for any $t \in [0, 1]$, to conclude that (2.25) is true by choosing $h := |Df|_w^p$.

For the second part of the statement, observe that (2.23) and (2.20) give

$$\frac{|f(\gamma_t) - f(\gamma_0)|}{t} \leq \frac{1}{pt} \int_0^t |Df|_w^p(\gamma_s) \, ds + \frac{1}{q}\left|\frac{E_{q,t}(\gamma)}{t}\right|^q,$$

that (2.25) yields that the family of functions $\frac{1}{pt}\int_0^t |Df|_w^p(\gamma_s) \, ds$ is dominated in $L^1(\pi)$, and use the assumption to conclude. $\qquad \square$

2.4 The Cheeger energy and its gradient flow

In Section 3.2, when discussing the existence of 'plans representing a gradient', we will need to call into play some non-trivial link between gradient flows in L^2 and Wasserstein geometry. In this short section we recall those results which we will need later on.

We start with the following statement concerning the dependence of the Sobolev classes w.r.t. the reference measure, proven in Section 8.2 of [5].

Proposition 2.12 *Let (X, d) be a complete and separable metric space and $\mathfrak{m}, \mathfrak{m}'$ two Radon non-negative measures on it. Assume that $\mathfrak{m} \ll \mathfrak{m}' \ll \mathfrak{m}$ with $\frac{d\mathfrak{m}'}{d\mathfrak{m}}$ locally bounded from above and from below by positive constants. Then for every $p \in (1, \infty)$ the sets $\mathrm{S}_{\mathrm{loc}}^p(X, \mathsf{d}, \mathfrak{m})$ and $\mathrm{S}_{\mathrm{loc}}^p(X, \mathsf{d}, \mathfrak{m}')$ coincide and for $f \in \mathrm{S}_{\mathrm{loc}}^p(X, \mathsf{d}, \mathfrak{m}) = \mathrm{S}_{\mathrm{loc}}^p(X, \mathsf{d}, \mathfrak{m}')$ a function G is a p-weak upper gradient of f in $\mathrm{S}_{\mathrm{loc}}^p(X, \mathsf{d}, \mathfrak{m})$ if and only if it is so on $\mathrm{S}_{\mathrm{loc}}^p(X, \mathsf{d}, \mathfrak{m}')$.*

In summary, the notion of being a Sobolev function is unchanged if we replace the reference measure $\mathfrak{m}$ with an equivalent one $\mathfrak{m}'$ such that $\log(\frac{d\mathfrak{m}'}{d\mathfrak{m}}) \in L_{\mathrm{loc}}^\infty(X, \mathfrak{m})$.

Now let $(X, \mathsf{d}, \mathfrak{m})$ be a space as in (2.7) and $\tilde{\mathfrak{m}} \in \mathscr{P}(X)$ our fixed measure satisfying (2.8) and (2.9). As a consequence of Proposition 2.12 above and the inequality $\tilde{\mathfrak{m}} \leq C\mathfrak{m}$ for some $C > 0$, we also get that $\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m}) \subset \mathrm{S}^p(X, \mathsf{d}, \tilde{\mathfrak{m}})$.

For $p \in (1, \infty)$ define the Cheeger energy functional $\mathrm{Ch}_p : L^2(X, \mathfrak{m}) \to [0, \infty]$ by

$$\mathrm{Ch}_p(f) := \begin{cases} \dfrac{1}{p} \displaystyle\int |Df|_w^p \, d\mathfrak{m}, & \text{if } f \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m}) \cap L^2(X, \mathfrak{m}), \\[2ex] +\infty & \text{otherwise.} \end{cases} \qquad (2.26)$$

And similarly $\widetilde{\mathrm{Ch}}_p : L^2(X, \tilde{\mathfrak{m}}) \to [0, \infty]$ by

$$\widetilde{\mathrm{Ch}}_p(f) := \begin{cases} \dfrac{1}{p} \displaystyle\int |Df|_w^p \, \mathrm{d}\tilde{\mathfrak{m}}, & \text{if } f \in \mathrm{S}^p(X, \mathsf{d}, \tilde{\mathfrak{m}}) \cap L^2(X, \tilde{\mathfrak{m}}), \\[2mm] +\infty & \text{otherwise.} \end{cases} \qquad (2.27)$$

The name 'Cheeger energy', introduced in [4] for the case $p = 2$, has been preferred over 'Dirichlet energy', because in general it is not a Dirichlet form - not even for $p = 2$, and because its first definition, which is equivalent to the one we gave, was based on a relaxation procedure similar to the one done by Cheeger in [13].

The properties of Ch_p, in particular of Ch_2, are linked to some special case of the definition of Laplacian - see Proposition 4.9 -, while those of $\widetilde{\mathrm{Ch}}_p$ will be of key importance in order to get some sort of compactness in the proof of the crucial Lemma 3.13.

Notice that $\widetilde{\mathrm{Ch}}_p$ is convex, lower semicontinuous and with dense domain, thus for any $f \in L^2(X, \tilde{\mathfrak{m}})$ there exists a unique gradient flow of $\widetilde{\mathrm{Ch}}_p$ starting from f.

The following theorem collects the main properties of this gradient flow. In an abstract setting it has been proved firstly in [22] (in Alexandrov spaces and for $p = 2$), generalized in [4] (for a class of spaces covering in particular those as in (2.7), and for $p = 2$) and in its current form in [5]. The difficult part of the statement is the fifth, where the Wasserstein geometry is linked to the dissipation of $\widetilde{\mathrm{Ch}}_p$: the key tool which allows to get the sharp bound on $|\dot\mu_t|$ is Kuwada's lemma, appeared at first in [26].

Theorem 2.13 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) and $\tilde{\mathfrak{m}} \in \mathscr{P}(X)$ as in (2.8), (2.9). Also, let $p \in (1, \infty)$, $f \in L^2(X, \tilde{\mathfrak{m}})$ and $(f_t) \subset L^2(X, \tilde{\mathfrak{m}})$ be the gradient flow of $\widetilde{\mathrm{Ch}}_p$ starting from f. Then the following hold.*

i) Mass preservation *$\int f_t \, \mathrm{d}\tilde{\mathfrak{m}} = \int f \, \mathrm{d}\tilde{\mathfrak{m}}$ for any $t \geq 0$.*

ii) Maximum principle *If $f \leq C$ (resp. $f \geq c$) $\tilde{\mathfrak{m}}$-a.e., then $f_t \leq C$ (resp. $f_t \geq c$) $\tilde{\mathfrak{m}}$-a.e. for any $t \geq 0$.*

iii) Entropy dissipation *If $c \leq f \leq C$ $\tilde{\mathfrak{m}}$-a.e. and $u : [c, C] \to \mathbb{R}$ is a C^2 map, then $t \mapsto \int u(f_t) \, \mathrm{d}\mathfrak{m}$ belongs to $C^1(0, \infty)$ and it holds*

$$\frac{\mathrm{d}}{\mathrm{d}t} \int u(f_t) \, \mathrm{d}\tilde{\mathfrak{m}} = -\int u''(f_t) |Df_t|_w^p \, \mathrm{d}\tilde{\mathfrak{m}}, \qquad \text{for any } t > 0, \qquad (2.28)$$

iv) Dissipation at $t = 0$ *With the same assumptions as in (iii), if furthermore $f \in \mathrm{S}^p(X, \mathsf{d}, \tilde{\mathfrak{m}})$, then (2.28) is true also for $t = 0$ and it holds*

$$\lim_{t \downarrow 0} \int u''(f_t) |Df_t|_w^p \, \mathrm{d}\tilde{\mathfrak{m}} = \int u''(f) |Df|_w^p \, \mathrm{d}\tilde{\mathfrak{m}}. \qquad (2.29)$$

v) Control on the Wasserstein speed *Assume that $c \leq f \leq C$ for some $c, C > 0$ and that $\int f \, \mathrm{d}\tilde{\mathfrak{m}} = 1$. Then the curve $t \mapsto \mu_t := f_t \tilde{\mathfrak{m}}$ (which has values in $\mathscr{P}_q(X)$ thanks to (i), (ii) and (2.9)) is q-absolutely continuous w.r.t. W_q, where $q \in (1, \infty)$ is the conjugate exponent of p, and for its metric speed $|\dot\mu_t|$ it holds*

$$|\dot\mu_t|^q \leq \int \frac{|Df_t|_w^p}{f_t^{q-1}} \, \mathrm{d}\tilde{\mathfrak{m}}, \qquad a.e.\ t. \qquad (2.30)$$

3   Differentials and gradients

3.1   Definition and basic properties

Following the approach described in Section 1.1, our first task toward giving the definition of distributional Laplacian on a metric measure space $(X, \mathsf{d}, \mathfrak{m})$ is to give a meaning to $D^{\pm} f(\nabla g)$ for Sobolev functions f, g. We will imitate the definition given in the right hand sides of (1.5), (1.6).

Start noticing that if $\varphi : \mathbb{R} \to \mathbb{R}^+$ is a convex function, the values of

$$\inf_{\varepsilon > 0} \frac{\varphi(\varepsilon)^p - \varphi(0)^p}{p\varepsilon\varphi(0)^{p-2}}, \qquad \sup_{\varepsilon < 0} \frac{\varphi(\varepsilon)^p - \varphi(0)^p}{p\varepsilon\varphi(0)^{p-2}}$$

are independent on $p \in (1, \infty)$, and equal to $\varphi(0)\varphi'(0^+)$, $\varphi(0)\varphi'(0^-)$ respectively, as soon as $\varphi(0) \neq 0$. The convexity also grants that the inf and sup can be substituted by $\lim_{\varepsilon \downarrow 0}$ and $\lim_{\varepsilon \uparrow 0}$, respectively.

Fix $p \in (1, \infty)$, let $f, g \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ and observe that (2.12) ensures that the map $\varepsilon \mapsto |D(g + \varepsilon f)|_w$ is convex in the sense that

$$|D(g + ((1 - \lambda)\varepsilon_0 + \lambda\varepsilon_1)f)|_w \leq (1 - \lambda)|D(g + \varepsilon_0 f)|_w + \lambda|D(g + \varepsilon_1 f)|_w, \qquad \mathfrak{m} - a.e.,$$

for any $\lambda \in [0, 1]$, $\varepsilon_0, \varepsilon_1 \in \mathbb{R}$.

Definition 3.1 $(D^{\pm} f(\nabla g))$ *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p \in (1, \infty)$ and $f, g \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$. The functions $D^{\pm} f(\nabla g) : X \to \mathbb{R}$ are $\mathfrak{m}$-a.e. well defined by*

$$\begin{aligned}
D^+ f(\nabla g) &:= \inf_{\varepsilon > 0} \frac{|D(g + \varepsilon f)|_w^p - |Dg|_w^p}{p\varepsilon|Dg|_w^{p-2}}, \\
D^- f(\nabla g) &:= \sup_{\varepsilon < 0} \frac{|D(g + \varepsilon f)|_w^p - |Dg|_w^p}{p\varepsilon|Dg|_w^{p-2}},
\end{aligned} \tag{3.1}$$

on $\{x : |Dg|_w(x) \neq 0\}$, and are taken 0 by definition on $\{x : |Dg|_w(x) = 0\}$, where the inf / sup should be intended as essential-inf / sup.

As for the notation $|Df|_w$, we are omitting the explicit dependence on the Sobolev exponent p in writing $D^{\pm} f(\nabla g)$, which affects the definition by the potential dependence on p of the objects $|D(g + \varepsilon f)|_{w,p}$. Yet, the initial discussion and Remark 2.5 grant that at least if the space is doubling and supports a 1-p' Poincaré inequality, then $D^{\pm} f(\nabla g)$ is unambiguously defined for $f, g \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$, where $p \geq p'$.

The initial discussion also ensures that that inf / sup can be replaced with essential limits as $\varepsilon \downarrow 0 / \varepsilon \uparrow 0$ respectively.

Throughout this paper, the expressions $D^{\pm} f(\nabla g)|Dg|_w^{p-2}$ on the set $\{x : |Dg|_w(x) = 0\}$ will always be taken 0 by definition. In this way it always holds

$$\begin{aligned}
D^+ f(\nabla g)|Dg|_w^{p-2} &= \inf_{\varepsilon > 0} \frac{|D(g + \varepsilon f)|_w^p - |Dg|_w^p}{p\varepsilon}, \\
D^- f(\nabla g)|Dg|_w^{p-2} &= \sup_{\varepsilon < 0} \frac{|D(g + \varepsilon f)|_w^p - |Dg|_w^p}{p\varepsilon},
\end{aligned}$$

$\mathfrak{m}$-a.e. on X.

Notice that the inequality $|D(g + \varepsilon f)|_w \leq |Dg|_w + |\varepsilon||Df|_w$ yields

$$|D^{\pm} f(\nabla g)| \leq |Df|_w |Dg|_w, \qquad \mathfrak{m} - a.e., \tag{3.2}$$

and in particular $D^{\pm} f(\nabla g)|Dg|_w^{p-2} \in L^1(X, \mathfrak{m})$ for any $f, g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ (resp. in $L^1_{\mathrm{loc}}(X, \mathfrak{m})$ for $f, g \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$).

The convexity of $f \mapsto |D(g + f)|_w^p$ gives

$$D^- f(\nabla g) \leq D^+ f(\nabla g), \qquad \mathfrak{m} - a.e., \tag{3.3}$$

and

$$
\begin{aligned}
D^+ f(\nabla g)|Dg|_w^{p-2} &= \inf_{\varepsilon > 0} D^- f(\nabla g_\varepsilon)|Dg_\varepsilon|_w^{p-2} = \inf_{\varepsilon > 0} D^+ f(\nabla g_\varepsilon)|Dg_\varepsilon|_w^{p-2}, \qquad \mathfrak{m} - a.e., \\
D^- f(\nabla g)|Dg|_w^{p-2} &= \sup_{\varepsilon < 0} D^+ f(\nabla g_\varepsilon)|Dg_\varepsilon|_w^{p-2} = \sup_{\varepsilon < 0} D^- f(\nabla g_\varepsilon)|Dg_\varepsilon|_w^{p-2}, \qquad \mathfrak{m} - a.e.,
\end{aligned}
\tag{3.4}
$$

where $g_\varepsilon := g + \varepsilon f$.

Also, from the definition it directly follows that

$$D^+(-f)(\nabla g) = -D^- f(\nabla g) = D^+ f(\nabla(-g)), \qquad \mathfrak{m} - a.e., \tag{3.5}$$

and that

$$D^{\pm} g(\nabla g) = |Dg|_w^2, \qquad \mathfrak{m} - a.e.. \tag{3.6}$$

As a consequence of the locality properties (2.14) and (2.15) we also have

$$D^{\pm} f(\nabla g) = 0, \quad \mathfrak{m} - a.e. \quad \text{on } f^{-1}(\mathcal{N}) \cup g^{-1}(\mathcal{N}), \qquad \forall \mathcal{N} \subset \mathbb{R} \quad \text{such that } \mathcal{L}^1(\mathcal{N}) = 0 \tag{3.7}$$

and

$$D^{\pm} f(\nabla g) = D^{\pm} \tilde{f}(\nabla \tilde{g}), \qquad \mathfrak{m} - a.e. \quad \text{on } \{f = \tilde{f}\} \cap \{g = \tilde{g}\}. \tag{3.8}$$

Some further basic properties of $D^{\pm} f(\nabla g)$ are collected in the following proposition.

Proposition 3.2 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p \in (1, \infty)$ and $g \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$. Then*

$$\mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}) \ni f \qquad \mapsto \qquad D^+ f(\nabla g),$$

is positively 1-homogeneous, convex in the $\mathfrak{m}$-a.e. sense, i.e.

$$D^+((1 - \lambda)f_0 + \lambda f_1)(\nabla g) \leq (1 - \lambda)D^+ f_1(\nabla g) + \lambda D^+ f_2(\nabla g), \qquad \mathfrak{m} - a.e.,$$

for any $f_0, f_1 \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ and $\lambda \in [0, 1]$, and 1-Lipschitz in the following sense:

$$\left| D^+ f_1(\nabla g) - D^+ f_2(\nabla g) \right| \leq |D(f_1 - f_2)|_w |Dg|_w, \qquad \mathfrak{m} - a.e. \ \forall f_1, f_2 \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}). \tag{3.9}$$

Similarly,

$$\mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}) \ni f \qquad \mapsto \qquad D^- f(\nabla g)$$

is positively 1-homogeneous, concave and 1-Lipschitz.

Conversely, for any $f \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ it holds

$$
\begin{aligned}
\mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}) \ni g &\mapsto D^+ f(\nabla g) \quad \textit{is positively 1-homogeneous and upper semicontinuous,} \\
\mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}) \ni g &\mapsto D^- f(\nabla g) \quad \textit{is positively 1-homogeneous and lower semicontinuous,}
\end{aligned}
$$

where upper semicontinuity is intended as follows: if $g_n, g \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$, $n \in \mathbb{N}$, and for some Borel set $E \subset X$ it holds $\int_E |Df|^p_w \, \mathrm{d}\mathfrak{m} < \infty$, $\sup_n \int_E |Dg_n|^p_w \, \mathrm{d}\mathfrak{m} < \infty$ and $\int_E |D(g_n - g)|^p_w \, \mathrm{d}\mathfrak{m} \to 0$, then it holds

$$\varlimsup_{n \to \infty} \int_{E'} D^+ f(\nabla g_n) |Dg_n|^{p-2}_w \, \mathrm{d}\mathfrak{m} \leq \int_{E'} D^+ f(\nabla g) |Dg|^{p-2}_w \, \mathrm{d}\mathfrak{m}, \qquad \forall \text{ Borel sets } E' \subset E.$$

Similarly for lower semicontinuity.

proof Positive 1-homogeneity in f, g is obvious.

For convexity (resp. concavity) in f, just notice that from (2.12) we get

$$|Dg + \varepsilon((1-\lambda)f_0 + \lambda f_1)|_w - |Dg|_w \leq (1-\lambda)\Big(|D(g + \varepsilon f_1)|_w - |Dg|_w\Big) + \lambda\Big(|D(g + \varepsilon f_1)|_w - |Dg|_w\Big),$$

so that the conclusion follows dividing by $\varepsilon > 0$ (resp. $\varepsilon < 0$) and letting $\varepsilon \downarrow 0$ (resp. $\varepsilon \uparrow 0$).

The Lipschitz continuity in f is a consequence of

$$\left| \frac{|D(g + \varepsilon f)|_w - |Dg|_w}{\varepsilon} - \frac{|D(g + \varepsilon \tilde{f})|_w - |Dg|_w}{\varepsilon} \right| \leq |D(f - \tilde{f})|_w, \qquad \forall \varepsilon \neq 0.$$

For the semicontinuity in g, let $E \subset X$ be a Borel set as in the assumptions and $V \subset \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ the space of g's such that $\int_E |Dg|^p_w \, \mathrm{d}\mathfrak{m} < \infty$. Endow V with the seminorm $\|g\|_{\mathrm{S}^p, E} := \sqrt[p]{\int_E |Dg|^p_w \, \mathrm{d}\mathfrak{m}}$ and notice that for any $\varepsilon \neq 0$ the real valued map

$$V \ni g \qquad \mapsto \qquad \int_{E'} \frac{|D(g + \varepsilon f)|^p_w - |Dg|^p_w}{p\varepsilon} \, \mathrm{d}\mathfrak{m},$$

is continuous. Hence the map

$$V \ni g \qquad \mapsto \qquad \int_{E'} D^+ f(\nabla g) |Dg|^{p-2}_w \, \mathrm{d}\mathfrak{m} = \inf_{\varepsilon > 0} \int_{E'} \frac{|D(g + \varepsilon f)|^p_w - |Dg|^p_w}{p\varepsilon} \, \mathrm{d}\mathfrak{m},$$

is upper semicontinuous. Similarly for $D^- f(\nabla g)$. $\qquad \square$

We now propose the following definition:

Definition 3.3 (q-infinitesimally strictly convex spaces) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p \in (1, \infty)$ and q the conjugate exponent. We say that $(X, \mathsf{d}, \mathfrak{m})$ is q-infinitesimally strictly convex provided*

$$\int D^+ f(\nabla g) |Dg|^{p-2}_w \, \mathrm{d}\mathfrak{m} = \int D^- f(\nabla g) |Dg|^{p-2}_w \, \mathrm{d}\mathfrak{m}, \qquad \forall f, g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m}). \tag{3.10}$$

In the case of normed spaces analyzed in Section 1.1, q-infinitesimal strict convexity is equivalent to strict convexity of the norm, whatever q is, which in turn is equivalent to the differentiability of the dual norm in the cotangent space. Formula (3.10) describes exactly such differentiability property, whence the terminology used. Notice that q-infinitesimal strict convexity has nothing to do with the strict convexity of $W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ as the former is about strict convexity of the 'tangent space', while the latter is about the one of the 'cotangent space'.

We don't know whether there is any general relation between infinitesimal strict convexity and non-branching (not even under additional assumptions like doubling of the measure, validity of a weak local Poincaré inequality or bounds from below on the Ricci curvature), but observe that the former makes sense also in non-geodesic spaces.

From inequality (3.3) we get that the integral equality (3.10) is equivalent to the pointwise one:

$$D^+ f(\nabla g) = D^- f(\nabla g), \qquad \mathfrak{m} - a.e., \qquad \forall f, g \in S^p(X, \mathsf{d}, \mathfrak{m}). \tag{3.11}$$

Then a simple cut-off argument and the locality (3.8) ensures that (3.11) is true also for $f, g \in S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$. Furthermore, from Remark 2.5 we know that if the measure is doubling and the space supports a p'-weak local Poincaré inequality, then q'-infinitesimal strict convexity implies q-infinitesimal strict convexity for any $q \in (1, q')$, where q' is the conjugate exponent of p'.

On q-infinitesimally strictly convex spaces, we will denote by $Df(\nabla g)$ the common value of $D^+ f(\nabla g) = D^- f(\nabla g)$, for $f, g \in S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$. As a direct consequence of Proposition 3.2 we have the following corollary:

Corollary 3.4 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p, q \in (1, \infty)$ be conjugate exponents. Assume that the space is q-infinitesimally strictly convex. Then:*

- *For any $g \in S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ the map*

$$S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}) \ni f \qquad \mapsto \qquad Df(\nabla g)$$

  is linear $\mathfrak{m}$-a.e., i.e.

$$D(\alpha_1 f_1 + \alpha_2 f_2)(\nabla g) = \alpha_1 Df_1(\nabla g) + \alpha_2 Df_2(\nabla g), \qquad \mathfrak{m} - a.e.$$

  for any $f_1, f_2 \in S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$, $\alpha_1, \alpha_2 \in \mathbb{R}$.

- *For any $f \in S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ the map*

$$S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}) \ni g \qquad \mapsto \qquad Df(\nabla g),$$

  is 1-homogeneous and continuous, the latter meaning that if $g_n, g \in S^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$, $n \in \mathbb{N}$, and for some Borel set $E \subset X$ it holds $\sup_n \int_E |Dg_n|_w^p \, \mathrm{d}\mathfrak{m} < \infty$ and $\int_E |D(g_n - g)|_w^p \, \mathrm{d}\mathfrak{m} \to 0$, then

$$\lim_{n \to \infty} \int_{E'} Df(\nabla g_n)|Dg_n|_w^{p-2} \, \mathrm{d}\mathfrak{m} = \int_{E'} Df(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}\mathfrak{m}, \qquad \forall \text{ Borel sets } E' \subset E.$$

Remark 3.5 (Weak semicontinuity in f) As a direct consequence of Proposition 3.2 we get the following result. For $p \in (1, \infty)$ and $g \in S^p(X, \mathsf{d}, \mathfrak{m})$ the map

$$S^p(X, \mathsf{d}, \mathfrak{m}) \ni f \qquad \mapsto \qquad \int D^+ f(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}\mathfrak{m},$$

is weakly lower semicontinuous in $S^p(X, \mathsf{d}, \mathfrak{m})$, in the sense that if $L(f_n) \to L(f)$ as $n \to \infty$ for any bounded - w.r.t. the seminorm $\| \cdot \|_{S^p}$ - linear functional $L : S^p(X, \mathsf{d}, \mathfrak{m}) \to \mathbb{R}$, then

$$\lim_{n \to \infty} \int D^+ f_n(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}\mathfrak{m} \geq \int D^+ f(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}\mathfrak{m}.$$

(in other words, we are picking the quotient of $S^p(X, \mathsf{d}, \mathfrak{m})$ w.r.t. the subspace of functions f with $\|f\|_{S^p} = 0$, taking its completion and then considering the usual weak convergence in the resulting Banach space. This abstract completion has some right to be called p-cotangent space of $(X, \mathsf{d}, \mathfrak{m})$, see Appendix A for some results in this direction).

To prove this weak lower semicontinuity, just notice that $f \mapsto \int D^+ f(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}\mathfrak{m}$ is convex and continuous.

Similarly, $f \mapsto \int D^- f(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}\mathfrak{m}$ is weakly upper semicontinuous in $S^p(X, \mathsf{d}, \mathfrak{m})$. ∎

3.2 Horizontal and vertical derivatives

In Section 1.1 we proved that on a flat normed space, for any two smooth functions f, g it holds

$$
\begin{aligned}
\inf_{\varepsilon > 0} \frac{\|D(g + \varepsilon f)(x)\|_*^2 - \|Dg(x)\|_*^2}{2\varepsilon} &= \max_{v \in \nabla g(x)} Df(v), \\
\sup_{\varepsilon < 0} \frac{\|D(g + \varepsilon f)(x)\|_*^2 - \|Dg(x)\|_*^2}{2\varepsilon} &= \min_{v \in \nabla g(x)} Df(v).
\end{aligned}
\tag{3.12}
$$

In this section we investigate the validity of this statement in a metric setting. More precisely, in Theorem 3.10 we will prove an analogous of

$$
\inf_{\varepsilon > 0} \frac{\|D(g + \varepsilon f)(x)\|_*^2 - \|Dg(x)\|_*^2}{2\varepsilon} \geq Df(v) \geq \sup_{\varepsilon < 0} \frac{\|D(g + \varepsilon f)(x)\|_*^2 - \|Dg(x)\|_*^2}{2\varepsilon},
\tag{3.13}
$$

for any $v \in \nabla g(x)$, given that this will be sufficient for our purposes. The non-smooth analogous of (3.12) will be proved in Appendix A, see Theorem A.4.

The ideas behind the main results of this part of the paper, namely Theorems 3.10 and 3.14, already appeared in [6] as key tools both to find a formula for the differentiation of the relative entropy along a Wasserstein geodesic, and to develop some notions of calculus in spaces with linear heat flow. In this paper, these ideas are extended and generalized, see in particular the next section for general calculus rules.

The definition we gave of $D^+ f(\nabla g)$ and $D^- f(\nabla g)$ are what play the role of the leftmost and rightmost sides in (3.13). In order to define a metric-measure theoretic analogous of the middle term, we need to introduce the notion of *plan representing the gradient of a function*. We start with the q-norm of a plan. Recall that the q-energies $E_{q,t}$ of a curve were defined in (2.19).

Definition 3.6 (q-norm of a plan) *For $q \in (1, \infty)$ and $\boldsymbol{\pi} \in \mathscr{P}(C([0, 1], X))$ we define the q-norm $\|\boldsymbol{\pi}\|_q \in [0, \infty]$ of $\boldsymbol{\pi}$ by*

$$
\|\boldsymbol{\pi}\|_q^q := \overline{\lim_{t \downarrow 0}} \int \left(\frac{E_{q,t}}{t} \right)^q \mathrm{d}\boldsymbol{\pi}
\tag{3.14}
$$

so that $\|\boldsymbol{\pi}\|_q^q = \overline{\lim}_{t \downarrow 0} \frac{1}{t} \iint_0^t |\dot{\gamma}_s|^q \, \mathrm{d}s \, \mathrm{d}\boldsymbol{\pi}(\gamma)$ if $(\mathrm{restr}_0^T)_\sharp \boldsymbol{\pi}$ is concentrated on absolutely continuous curves for some $T \in (0, 1]$ and $+\infty$ otherwise.

Notice that $\mathscr{P}(C([0, 1], X))$ is not a vector space, so that in calling $\|\cdot\|_q$ a norm we are somehow abusing the notation. Yet, this wording well fits the discussion hereafter, and a rigorous meaning to it can be given, see Appendix A.

If $\boldsymbol{\pi}$ is of bounded compression and $\|\boldsymbol{\pi}\|_q < \infty$, then not necessarily $\boldsymbol{\pi}$ is a q-test plan, because it might be that $\iint_0^1 |\dot\gamma_t|^q \, dt \, d\boldsymbol{\pi}(\gamma) = \infty$. However, for $T > 0$ small enough, $(\mathrm{restr}_0^T)_\sharp \boldsymbol{\pi}$ is a q-test plan. Hence if $p, q \in (1, \infty)$ are conjugate exponents, $\boldsymbol{\pi}$ is of bounded compression with $\|\boldsymbol{\pi}\|_q < \infty$ and $g \in S^p(X, \mathsf{d}, \mathfrak{m})$, (2.22) yields that

$$\overline{\lim_{t \downarrow 0}} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \, d\boldsymbol{\pi} \le \frac{\||Dg|_w\|_{L^p(X,(e_0)_\sharp \boldsymbol{\pi})}^p}{p} + \frac{\|\boldsymbol{\pi}\|_q^q}{q}. \tag{3.15}$$

We then propose the following definition.

Definition 3.7 (Plans representing gradients) *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be as in* (2.7)*,* $p, q \in (1, \infty)$ *conjugate exponents and* $g \in S^p(X, \mathsf{d}, \mathfrak{m})$*. We say that* $\boldsymbol{\pi} \in \mathscr{P}(C([0, 1], X))$ *q-represents* ∇g *if* $\boldsymbol{\pi}$ *is of bounded compression,* $\|\boldsymbol{\pi}\|_q < \infty$ *and it holds*

$$\underline{\lim_{t \downarrow 0}} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \, d\boldsymbol{\pi} \ge \frac{\||Dg|_w\|_{L^p(X,(e_0)_\sharp \boldsymbol{\pi})}^p}{p} + \frac{\|\boldsymbol{\pi}\|_q^q}{q}. \tag{3.16}$$

Remark 3.8 Notice that the definition we just gave is not too far from De Giorgi's original definition of gradient flow in a metric setting (see [3] for a modern approach to the topic). Differences are the fact that here we care only about the behavior of the curves in the support of $\boldsymbol{\pi}$ for t close to 0, and the fact that we are considering a Sobolev notion, rather than a metric one. ∎

Remark 3.9 The key information encoded by a plan $\boldsymbol{\pi}$ q-representing ∇g is contained in the disintegration $\{\boldsymbol{\pi}_x\}_{x \in X}$ of $\boldsymbol{\pi}$ w.r.t. e_0: this provides a Borel map $X \ni x \mapsto \boldsymbol{\pi}_x \in \mathscr{P}(C([0, 1], X))$ which "associates to $(e_0)_\sharp \boldsymbol{\pi}$-a.e. x a weighted set of curves which is a gradient flow at time 0 of g starting from x". Once this disintegration is given, for every measure μ such that $\mu \le C(e_0)_\sharp \boldsymbol{\pi}$ for some $C > 0$, the plan $\boldsymbol{\pi}_\mu := \int \boldsymbol{\pi}_x \, d\mu(x)$ also q-represents ∇g, see also Proposition 3.11 below.

Yet, it is curious to observe that the property of q-representing a gradient can be fully described by looking at the marginals $(e_t)_\sharp \boldsymbol{\pi}$ only. Indeed, assume that $\boldsymbol{\pi}$ q-represents ∇g. Then the curve $t \mapsto \mu_t := (e_t)_\sharp \boldsymbol{\pi}$ is q-absolutely continuous w.r.t. W_q in a neighborhood of 0. Now let $\tilde{\boldsymbol{\pi}} \in \mathscr{P}(C([0, 1], X))$ be any plan associated to (μ_t) via Theorem 2.1. We have

$$\int \frac{g(\gamma_t) - g(\gamma_0)}{t} \, d\boldsymbol{\pi}(\gamma) = \frac{1}{t} \left(\int g \, d\mu_t - \int g \, d\mu_0 \right) = \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \, d\tilde{\boldsymbol{\pi}}(\gamma),$$

$\||Dg|_w\|_{L^p(X,(e_0)_\sharp \boldsymbol{\pi})} = \||Dg|_w\|_{L^p(X,\mu_0)} = \||Dg|_w\|_{L^p(X,(e_0)_\sharp \tilde{\boldsymbol{\pi}})}$ and

$$\overline{\lim_{t \downarrow 0}} \frac{1}{t} \iint_0^t |\dot\gamma_t|^q \, dt \, d\boldsymbol{\pi}(\gamma) \ge \overline{\lim_{t \downarrow 0}} \frac{1}{t} \int_0^t |\dot\mu_t|^q \, dt = \overline{\lim_{t \downarrow 0}} \frac{1}{t} \iint_0^t |\dot\gamma_t|^q \, dt \, d\tilde{\boldsymbol{\pi}}(\gamma), \tag{3.17}$$

which gives that $\tilde{\boldsymbol{\pi}}$ also q-represents ∇g. This argument also shows that the inequality in (3.17) must be an equality, or else for the plan $\tilde{\boldsymbol{\pi}}$ the inequality (3.15) would fail. ∎

The following theorem is the key technical point on which the theory developed in this paper relies. It relates the 'vertical derivatives' $D^\pm f(\nabla g)$ (so called because $D^+ f(\nabla g), D^- f(\nabla g)$ are obtained by a perturbation in the dependent variable), with the 'horizontal derivatives' $\underline{\lim}_{t \downarrow 0} \int \frac{f \circ e_t - f \circ e_0}{t} \, d\boldsymbol{\pi}$, $\overline{\lim}_{t \downarrow 0} \int \frac{f \circ e_t - f \circ e_0}{t} \, d\boldsymbol{\pi}$ (so called because f is perturbed in the dependent variable).

Theorem 3.10 (Horizontal and vertical derivatives) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p, q \in (1, \infty)$ conjugate exponents and $f, g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$. Then for every plan $\boldsymbol{\pi} \in \mathscr{P}(C([0,1], X))$ which q-represents ∇g it holds*

$$\int D^+ f(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}(\mathrm{e}_0)_\sharp \boldsymbol{\pi} \geq \varlimsup_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma)$$

$$\geq \varliminf_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \geq \int D^- f(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}(\mathrm{e}_0)_\sharp \boldsymbol{\pi}.$$

$$(3.18)$$

proof By the inequality (3.15) applied to $g + \varepsilon f$ we have

$$\varlimsup_{t \downarrow 0} \int \frac{(g + \varepsilon f)(\gamma_t) - (g + \varepsilon f)(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \frac{\||D(g + \varepsilon f)|_w\|_{L^p(X, (\mathrm{e}_0)_\sharp \boldsymbol{\pi})}^p}{p} + \frac{\|\boldsymbol{\pi}\|_q^q}{q},$$

and by the assumption on g and $\boldsymbol{\pi}$ we know that

$$\varliminf_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \geq \frac{\||Dg|_w\|_{L^p(X, (\mathrm{e}_0)_\sharp \boldsymbol{\pi})}^p}{p} + \frac{\|\boldsymbol{\pi}\|_q^q}{q}.$$

Subtract the second inequality from the first to get

$$\varlimsup_{t \downarrow 0} \varepsilon \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \int \frac{|D(g + \varepsilon f)|_w^p - |Dg|_w^p}{p} \, \mathrm{d}(\mathrm{e}_0)_\sharp \boldsymbol{\pi}.$$

Divide by $\varepsilon > 0$ (resp. $\varepsilon < 0$) and use the dominate convergence theorem to get the first (resp. the third) inequality in (3.18). $\qquad\square$

The rest of the section is devoted to the study of the properties of plans q-representing gradients. In particular, in Theorem 3.14 we will show that they actually exist for a wide class of initial data (i.e. given value of $(\mathrm{e}_0)_\sharp \boldsymbol{\pi}$).

Proposition 3.11 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p, q \in (1, \infty)$ conjugate exponents, $g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and $\boldsymbol{\pi} \in \mathscr{P}(C([0,1], X))$ a plan of bounded compression.*
Then $\boldsymbol{\pi}$ q-represents ∇g if and only if

$$\lim_{t \downarrow 0} \frac{g \circ \mathrm{e}_t - g \circ \mathrm{e}_0}{E_{q,t}} = \lim_{t \downarrow 0} \left(\frac{E_{q,t}}{t} \right)^{\frac{q}{p}} = |Dg|_w \circ \mathrm{e}_0, \qquad in \qquad L^p(C([0,1], X), \boldsymbol{\pi}). \qquad (3.19)$$

proof The 'if' is obvious, so we turn to the 'only if'.
Thanks to (3.15) and its proof, (3.16) and the equality case in Young's inequality, we have

$$\varlimsup_{t \downarrow 0} \int \frac{g \circ \mathrm{e}_t - g \circ \mathrm{e}_0}{t} \, \mathrm{d}\boldsymbol{\pi} = \varliminf_{t \downarrow 0} \int \frac{g \circ \mathrm{e}_t - g \circ \mathrm{e}_0}{t} \, \mathrm{d}\boldsymbol{\pi} = \|\boldsymbol{\pi}\|_q^q = \||Dg|_w\|_{L^p(X, (\mathrm{e}_0)_\sharp \boldsymbol{\pi})}^p =: L$$

$$(3.20)$$

Define the functions $A_t, B_t, C_t : C([0,1], X) \to \mathbb{R} \cup \{\pm\infty\}$ by

$$A_t := \frac{g \circ \mathrm{e}_t - g \circ \mathrm{e}_0}{E_{q,t}}, \qquad B_t := \frac{E_{q,t}}{t}, \qquad C_t := \sqrt[p]{\frac{1}{t} \int_0^t |Dg|_w^p \circ \mathrm{e}_s \, \mathrm{d}s},$$

and notice that from (2.20) we have

$$|A_t| \leq C_t, \qquad \boldsymbol{\pi} - a.e., \tag{3.21}$$

and that from (2.24) we know that

$$C_t \to |Dg|_w \circ e_0, \qquad \text{in } L^p(\boldsymbol{\pi}). \tag{3.22}$$

From (3.20), (3.21) and (3.22) we deduce

$$
\begin{aligned}
L &= \lim_{t\downarrow 0} \int \frac{g \circ e_t - g \circ e_0}{t} \, d\boldsymbol{\pi} = \lim_{t\downarrow 0} \int A_t B_t \, d\boldsymbol{\pi} \leq \varlimsup_{t\downarrow 0} \int |A_t| B_t \, d\boldsymbol{\pi} \\
&\leq \varlimsup_{t\downarrow 0} \left(\frac{\|A_t\|^p_{L^p(\boldsymbol{\pi})}}{p} + \frac{\|B_t\|^q_{L^q(\boldsymbol{\pi})}}{q} \right) \leq \varlimsup_{t\downarrow 0} \frac{\|A_t\|^p_{L^p(\boldsymbol{\pi})}}{p} + \varlimsup_{t\downarrow 0} \frac{\|B_t\|^q_{L^q(\boldsymbol{\pi})}}{q} \\
&\leq \varlimsup_{t\downarrow 0} \frac{\|C_t\|^p_{L^p(\boldsymbol{\pi})}}{p} + \frac{\|\boldsymbol{\pi}\|^q_q}{q} = L.
\end{aligned}
\tag{3.23}
$$

In particular it holds $\varliminf_{t\downarrow 0} \left(\frac{\|A_t\|^p_{L^p(\boldsymbol{\pi})}}{p} + \frac{\|B_t\|^q_{L^q(\boldsymbol{\pi})}}{q} \right) = \varlimsup_{t\downarrow 0} \frac{\|A_t\|^p_{L^p(\boldsymbol{\pi})}}{p} + \varlimsup_{t\downarrow 0} \frac{\|B_t\|^q_{L^q(\boldsymbol{\pi})}}{q}$, which forces the limits of $\|A_t\|^p_{L^p(\boldsymbol{\pi})}, \|B_t\|^q_{L^q(\boldsymbol{\pi})}$ as $t \downarrow 0$ to exist. Then, since the last inequality of (3.23) is an equality, we also get that $\lim_{t\downarrow 0} \|A_t\|^p_{L^p(\boldsymbol{\pi})} = \lim_{t\downarrow 0} \|C_t\|^p_{L^p(\boldsymbol{\pi})}$, which together with (3.21) and (3.22) gives that $|A_t| \to |Dg|_w \circ e_0$ as $t \downarrow 0$ in $L^p(\boldsymbol{\pi})$. Since also the first inequality in (3.23) is an equality, we deduce also that $A_t \to |Dg|_w \circ e_0$ as $t \downarrow 0$ in $L^p(\boldsymbol{\pi})$.

Using (3.23) again we get that $\lim_{t\downarrow 0} \|B_t\|^q_{L^q(\boldsymbol{\pi})} = L$. Let $B \in L^q(\boldsymbol{\pi})$ be any weak limit of B_t as $t \downarrow 0$ in $L^q(\boldsymbol{\pi})$. From

$$L = \lim_{t\downarrow 0} \int A_t B_t \, d\boldsymbol{\pi} = \int |Dg|_w \circ e_0 \, B \, d\boldsymbol{\pi} \leq \frac{\||Dg|_w \circ e_0\|^p_{L^p(\boldsymbol{\pi})}}{p} + \frac{\|B\|^q_{L^q(\boldsymbol{\pi})}}{q} \leq L, \tag{3.24}$$

we deduce $\|B\|^q_{L^q(\boldsymbol{\pi})} = L$, so that the weak convergence of B_t is actually strong. Finally, since the first inequality in (3.24) is an equality, we deduce $B^q = |Dg|^p_w \circ e_0$ $\boldsymbol{\pi}$-a.e., and the proof is achieved. $\qquad\square$

Direct consequences of characterization (3.19) are the following. For $g \in S^p(X, d, \mathfrak{m})$ it holds:

- if $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2$ are plans which q-represent ∇g then

$$\boldsymbol{\pi} := \lambda \boldsymbol{\pi}_1 + (1 - \lambda) \boldsymbol{\pi}_2 \qquad q - \text{represents } \nabla g, \qquad \text{for any } \lambda \in [0, 1], \tag{3.25}$$

- if $\boldsymbol{\pi}$ is a plan which q-represents ∇g and $F : C([0, 1], X) \to \mathbb{R}$ is a non-negative bounded Borel function such that $\int F \, d\boldsymbol{\pi} = 1$, then

$$\tilde{\boldsymbol{\pi}} := F \boldsymbol{\pi} \qquad q - \text{represents } \nabla g. \tag{3.26}$$

We turn to the existence of plans representing gradients.

Remark 3.12 In connection with Remark 3.8, it is worth stressing the following. The general existence results for gradient flows in metric spaces are based either on some compactness assumption, possibly w.r.t. some weak topology, of the sublevels of the function (see e.g. Chapters 2 and 3 of [3]), or on some structural assumption on the geometry of the space, which, very roughly said, expresses the fact that the metric space 'looks like an Hilbert space on small scales' (see e.g. Chapter 4 of [3] and [42]).

In our current framework, none of the two assumptions is made. Yet, it is still possible to prove that this sort of 'infinitesimal gradient flow of a Sobolev function' exists. The necessary compactness is gained by looking at the evolution of the probability densities in $L^2(X, \tilde{\mathfrak{m}})$, and then coming back to measures on $\mathscr{P}(C([0, 1], X))$ via the use of Theorems 2.13 and 2.1.

∎

Recall that we fixed $\tilde{\mathfrak{m}} \in \mathscr{P}(X)$ satisfying (2.8) and (2.9) and that we defined the Cheeger energy functional $\widetilde{\mathrm{Ch}}_p : L^2(X, \tilde{\mathfrak{m}}) \to [0, \infty]$ in (2.27).

Lemma 3.13 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $\tilde{\mathfrak{m}}$ as in (2.8) and (2.9), $p, q \in (1, \infty)$ conjugate exponents and $g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m}) \cap L^\infty(X, \mathfrak{m})$. Then there exists a plan $\boldsymbol{\pi}$ which q-represents ∇g and such that $c\tilde{\mathfrak{m}} \leq (\mathrm{e}_0)_\sharp \boldsymbol{\pi} \leq C\tilde{\mathfrak{m}}$ for some $c, C > 0$.*

proof Define the convex function $u_q : [0, \infty) \to \mathbb{R}$ as

$$u_q(z) := \frac{z^{3-q} - (3-q)z}{(3-q)(2-q)}, \qquad \text{if } q \neq 2, 3,$$
$$u_2(z) := z \log z - z,$$
$$u_3(z) := z - \log z.$$

If $q = 2$ put $\rho_0 := \tilde{c}e^{-g}$ otherwise find $a \in \mathbb{R}$ such that $1 + (a - g)(2 - q) > b > 0$ $\mathfrak{m}$-a.e. and define

$$\rho_0 := \tilde{c}\Big(1 + \big(a - g\big)(2 - q)\Big)^{\frac{1}{2-q}},$$

where in any case $\tilde{c}$ is chosen so that $\int \rho_0 \, \mathrm{d}\tilde{\mathfrak{m}} = 1$. Notice that by construction it holds $c \leq \rho_0 \leq C$ $\mathfrak{m}$-a.e. for some $c, C > 0$, $u_q'(\rho_0) = -g + constant$ and, by the chain rule (2.16), $\rho_0 \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m}) \subset \mathrm{S}^p(X, \mathsf{d}, \tilde{\mathfrak{m}})$.

Now consider the gradient flow (ρ_t) of the Cheeger energy $\widetilde{\mathrm{Ch}}_p$ in $L^2(X, \tilde{\mathfrak{m}})$ starting from ρ_0. Part (ii) of Theorem 2.13 ensures that the densities ρ_t are non-negative and uniformly bounded in $L^\infty(X, \tilde{\mathfrak{m}})$, while (i) of the same theorem grants $\int \rho_t \, \mathrm{d}\tilde{\mathfrak{m}} = 1$ for any $t \geq 0$. Hence we have $\mu_t := \rho_t \tilde{\mathfrak{m}} \in \mathscr{P}(X)$, and the L^∞ bound on ρ_t in conjunction with (2.9) yields that the measures μ_t have finite q-moment for any $t \geq 0$. Furthermore, from (v) of Theorem 2.13 we also have that the curve $[0, 1] \ni t \mapsto \mu_t$ is q-absolutely continuous w.r.t. W_q. Hence we can use Theorem 2.1 to associate to (μ_t) a plan $\boldsymbol{\pi} \in \mathscr{P}(C([0, 1], X))$ concentrated on $AC^q([0, 1], X)$ satisfying (2.5).

We claim that $\boldsymbol{\pi}$ q-represents ∇g. Since we already know by construction that $c\tilde{\mathfrak{m}} \leq (\mathrm{e}_0)_\sharp \boldsymbol{\pi} \leq C\tilde{\mathfrak{m}}$ for some $c, C > 0$, this will give the thesis.

Since the ρ_t's are uniformly bounded and $\tilde{\mathfrak{m}} \leq \tilde{C}\mathfrak{m}$ for some $\tilde{C} > 0$, $\boldsymbol{\pi}$ has bounded compression. Using again (v) of Theorem 2.13 and (2.5) we get

$$\iint_0^t |\dot{\gamma}_s|^q \, \mathrm{d}s \, \mathrm{d}\boldsymbol{\pi}(\gamma) = \int_0^t |\dot{\mu}_s|^q \, \mathrm{d}s \leq \iint_0^t \frac{|D\rho_s|_w^p}{\rho_s^{q-1}} \, \mathrm{d}s \, \mathrm{d}\tilde{\mathfrak{m}} = \iint_0^t u_q''(\rho_s)|D\rho_s|_w^p \, \mathrm{d}s \, \mathrm{d}\tilde{\mathfrak{m}}.$$

Recall that $\rho_0 \in S^p(X, \mathsf{d}, \mathfrak{m})$ so that applying (2.29) we get

$$\|\boldsymbol{\pi}\|_q^q = \overline{\lim_{t \downarrow 0}} \frac{1}{t} \iint_0^t |\dot{\gamma}_s|^q \, \mathrm{d}s \leq \overline{\lim_{t \downarrow 0}} \frac{1}{t} \iint_0^t u_q''(\rho_s)|D\rho_s|_w^p \, \mathrm{d}s \, \mathrm{d}\tilde{\mathfrak{m}} = \int u_q''(\rho_0)|D\rho_0|_w^p \, \mathrm{d}\tilde{\mathfrak{m}}. \quad (3.27)$$

On the other hand, the convexity of u_q gives

$$\int \frac{g(\gamma_t) - g(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) = \int \frac{u_q'\big(\rho_0(\gamma_0)\big) - u_q'\big(\rho_0(\gamma_t)\big)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) = \frac{1}{t} \int u_q'(\rho_0)(\rho_0 - \rho_t) \, \mathrm{d}\tilde{\mathfrak{m}}$$

$$\geq \frac{1}{t} \int u_q(\rho_0) - u_q(\rho_t) \, \mathrm{d}\tilde{\mathfrak{m}},$$

so that using part (iii) and (iv) of Theorem 2.13 we get

$$\lim_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \geq \int u_q''(\rho_0)|D\rho_0|_w^p \, \mathrm{d}\tilde{\mathfrak{m}}. \quad (3.28)$$

Since $u_q''(\rho_0)|D\rho_0|_w^p = \rho_0|Dg|_w^p$ $\tilde{\mathfrak{m}}$-a.e., we have $\int u_q''(\rho_0)|D\rho_0|_w^p \, \mathrm{d}\tilde{\mathfrak{m}} = \||Dg|_w\|_{L^p((\mathrm{e}_0)_\sharp \boldsymbol{\pi})}^p$, thus the thesis follows from (3.28), (3.27). $\qquad \square$

Theorem 3.14 (Existence of plans representing ∇g) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $\tilde{\mathfrak{m}}$ as in (2.8) and (2.9), $p, q \in (1, \infty)$ conjugate exponents, $g \in S^p(X, \mathsf{d}, \mathfrak{m})$ and $\mu \in \mathscr{P}(X)$ a measure such that $\mu \leq C\tilde{\mathfrak{m}}$ for some $C > 0$.*

Then there exists $\boldsymbol{\pi} \in \mathscr{P}(C([0,1], X))$ which q-represents ∇g and such that $(\mathrm{e}_0)_\sharp \boldsymbol{\pi} = \mu$.

proof Let $\varphi : \mathbb{R} \to [0, 1]$ be defined by

$$\varphi(z) := \begin{cases} x - 2n & , \quad \text{if } x \in [2n, 2n+1), \text{ for some } n \in \mathbb{Z}, \\ 2n - x & , \quad \text{if } x \in [2n-1, 2n), \text{ for some } n \in \mathbb{Z}, \end{cases}$$

and $\psi := 1 - \varphi$. Clearly, both φ and ψ are 1-Lipschitz. Apply Lemma 3.13 above to $\varphi \circ g$ and $\psi \circ g$ to get plans $\boldsymbol{\pi}^1, \boldsymbol{\pi}^2$ q-representing $\nabla(\varphi \circ g), \nabla(\psi \circ g)$ respectively and with $c\tilde{\mathfrak{m}} \leq (\mathrm{e}_0)_\sharp \boldsymbol{\pi}^i \leq C\tilde{\mathfrak{m}}$, $i = 1, 2$, for some $c, C > 0$. Let $A := \{g^{-1}(\cup_n [2n, 2n+1))\} \subset X$, define $F^1, F^2 : C([0,1], X) \to \mathbb{R}$ by

$$F^1(\gamma) := \chi_A(\gamma_0) \frac{\mathrm{d}\mu}{\mathrm{d}(\mathrm{e}_0)_\sharp \boldsymbol{\pi}^1}(\gamma_0),$$

$$F^2(\gamma) := \chi_{X \setminus A}(\gamma_0) \frac{\mathrm{d}\mu}{\mathrm{d}(\mathrm{e}_0)_\sharp \boldsymbol{\pi}^2}(\gamma_0),$$

and notice that since $\frac{\mathrm{d}(\mathrm{e}_0)_\sharp \boldsymbol{\pi}^i}{\mathrm{d}\tilde{\mathfrak{m}}}$ is bounded from below, $i = 1, 2$, F^1 and F^2 are bounded. Also, we know from (2.14) that $\tilde{\mathfrak{m}}$-a.e. on $g^{-1}(\mathbb{Z})$ it holds $|Dg|_w = |D(-g)|_w = 0$. Now let $\boldsymbol{\pi} := F^1 \boldsymbol{\pi}^1 + F^2 \boldsymbol{\pi}^2$ and use (3.25) and (3.26) to conclude that $\boldsymbol{\pi}$ indeed q-represents ∇g. Since by construction it holds $(\mathrm{e}_0)_\sharp \boldsymbol{\pi} = \mu$, the proof is concluded. $\qquad \square$

3.3 Calculus rules

In this section we prove the basic calculus rules for $D^\pm f(\nabla g)$: the chain and Leibniz rules.

Proposition 3.15 (Chain rules) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p \in (1, \infty)$ and $f, g \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$. Also, let $\varphi : \mathbb{R} \to \mathbb{R}$ be locally Lipschitz with the following property: for any $x \in X$ there exists a neighborhood $x \in U_x \subset X$ and an open interval $I_x \subset \mathbb{R}$ such that $\mathfrak{m}(U_x \setminus f^{-1}(I_x)) = 0$ and $\varphi_{|I_x}$ is Lipschitz.*
Then for any $\varphi \circ f \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ and it holds

$$D^\pm(\varphi \circ f)(\nabla g) = \varphi' \circ f \ D^{\pm \mathrm{sign}(\varphi' \circ f)} f(\nabla g), \qquad \mathfrak{m} - a.e., \tag{3.29}$$

where the value of $\varphi' \circ f$ is defined arbitrarily at points x where φ is not differentiable at $f(x)$.
Similarly, assume that $\varphi : \mathbb{R} \to \mathbb{R}$ is locally Lipschitz and with the following property: for any $x \in X$ there exists a neighborhood $x \in U_x \subset X$ and an open interval $I_x \subset \mathbb{R}$ such that $\mathfrak{m}(U_x \setminus g^{-1}(I_x)) = 0$ and $\varphi_{|I_x}$ is Lipschitz.
Then $\varphi \circ g \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ and it holds

$$D^\pm f(\nabla(\varphi \circ g)) = \varphi' \circ g \ D^{\pm \mathrm{sign}(\varphi' \circ g)} f(\nabla g), \qquad \mathfrak{m} - a.e., \tag{3.30}$$

where the value of $\varphi' \circ g$ is defined arbitrarily at points x where φ is not differentiable at $g(x)$.

proof The fact that $\varphi \circ f$ (resp. $\varphi \circ g$) is in $\mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ is a direct consequence of the assumptions and formula (2.16). Also, thanks to the local nature of the claim it is sufficient to deal with the case $f, g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and $\varphi : \mathbb{R} \to \mathbb{R}$ Lipschitz.

We now prove (3.29). Let $\mathcal{N} \subset \mathbb{R}$ be the $\mathcal{L}^1$-negligible set of points of non-differentiability of φ. Then $\varphi(\mathcal{N})$ is also $\mathcal{L}^1$-negligible and by (2.14) we deduce that (3.29) holds $\mathfrak{m}$-a.e. on $f^{-1}(\mathcal{N})$ because both sides are 0. Now observe that by the very definition of $D^\pm f(\nabla g)$, the thesis is obvious if φ is affine. By the locality principle (2.15) and arguing as before to deal with non-differentiability points, the thesis is also true if φ is piecewise affine. For general φ, find a sequence of piecewise affine functions φ_n such that $\varphi'_n \to \varphi'$ $\mathcal{L}^1$-a.e.. Let $\mathcal{N}'$ be the $\mathcal{L}^1$-negligible set of z's such that either one of φ, φ_n is not differentiable at z or $\varphi'_n(z)$ does not converge at $\varphi'(z)$. With the same arguments used before we get that (3.29) holds $\mathfrak{m}$-a.e. on $f^{-1}(\mathcal{N}')$ because both sides are 0. On $X \setminus f^{-1}(N')$ we use the continuity property (3.9) and the chain rule (2.16) to get

$$\left| D^\pm(\varphi \circ f)(\nabla g) - D^\pm(\varphi_n \circ f)(\nabla g) \right| \leq |D((\varphi - \varphi_n) \circ f)|_w |Dg|_w = |\varphi' - \varphi'_n| \circ f |Df|_w |Dg|_w,$$

and by construction the rightmost term goes to 0 as $n \to \infty$ $\mathfrak{m}$-a.e. on $X \setminus f^{-1}(\mathcal{N}')$.

We pass to (3.30). Using (3.5) we can reduce ourselves to prove that $D^+ f(\nabla(\varphi \circ g)) = \varphi' \circ g \ D^{\mathrm{sign}(\varphi' \circ g)} f(\nabla g)$. Notice that the trivial identity

$$\frac{|D(ag + \varepsilon f)|^p_w - |D(ag)|^p_w}{p\varepsilon |D(ag)|^{p-2}_w} = a \frac{|D(g + \frac{\varepsilon}{a} f)|^p_w - |Dg|^p_w}{p \frac{\varepsilon}{a} |Dg|^{p-2}_w}, \qquad \mathfrak{m} - a.e.,$$

valid for any $\varepsilon, a \neq 0$ together with (3.29), yields (3.30) for φ linear. Hence (3.30) also holds for φ affine and, by locality, also for φ piecewise affine, where the same arguments as before are used to deal with the set of points of non-differentiability of φ.

Let $\tilde{J} \subset J$ be an interval where φ is Lipschitz, and (φ_n) a sequence of uniformly Lipschitz and piecewise affine functions such that $\varphi'_n \to \varphi'$ $\mathcal{L}^1$-a.e. on $\tilde{J}$. By construction, for every

$E \subset g^{-1}(\tilde{J})$ the sequence of functions $\varphi'_n \circ g \, D^{\mathrm{sign}(\varphi'_n \circ g)} f(\nabla g)|D(\varphi_n \circ g)|_w^{p-2}$ is dominated in $L^1(E, \mathfrak{m}|_E)$ and pointwise $\mathfrak{m}$-a.e. converges to $\varphi' \circ g \, D^{\mathrm{sign}(\varphi' \circ g)} f(\nabla g)|D(\varphi \circ g)|_w^{p-2}$. From the fact that $\int_E |D(\varphi_n \circ g - \varphi \circ g)|_w^p \, d\mathfrak{m} \to 0$ as $n \to \infty$ and the upper semicontinuity statement of Proposition 3.2 we deduce

$$
\begin{aligned}
\int_E D^+ f(\nabla(\varphi \circ g))|D(\varphi \circ g)|_w^{p-2} \, d\mathfrak{m} &\geq \varlimsup_{n \to \infty} \int_E D^+ f(\nabla(\varphi_n \circ g))|D(\varphi_n \circ g)|_w^{p-2} \, d\mathfrak{m} \\
&= \varlimsup_{n \to \infty} \int_E \varphi'_n \circ g \, D^{\mathrm{sign}(\varphi'_n \circ g)} f(\nabla g)|D(\varphi_n \circ g)|_w^{p-2} \, d\mathfrak{m} \\
&= \int_E \varphi' \circ g \, D^{\mathrm{sign}(\varphi' \circ g)} f(\nabla g)|D(\varphi \circ g)|_w^{p-2} \, d\mathfrak{m}.
\end{aligned}
$$

By the arbitrariness of $\tilde{J}$ and E we get

$$
D^+ f(\nabla(\varphi \circ g)) \geq \varphi' \circ g \, D^{\mathrm{sign}(\varphi' \circ g)} f(\nabla g), \qquad \mathfrak{m} - a.e.. \tag{3.31}
$$

Roughly said, the conclusion comes by applying this very same inequality with g replaced by $\varphi \circ g$ and φ replaced by φ^{-1}. To make this rigorous, assume at first that φ is in C^1_{loc}. Notice that equality in (3.31) holds $\mathfrak{m}$-a.e. on $g^{-1}(\{\varphi' = 0\})$. Now pick z such that $\varphi'(z) \neq 0$, say $\varphi'(z) > 0$ (the other case is similar) and let $(a, b) \ni z$ be an open interval such that $\inf_{(a,b)} \varphi' > 0$. Put $\tilde{g} := \min\{\max\{g, a\}, b\}$, notice that φ is invertible on $\tilde{g}(X)$ with C^1_{loc} inverse and use (3.31) to get

$$
D^+ f(\nabla \tilde{g}) = D^+ f(\nabla(\varphi^{-1} \circ (\varphi \circ \tilde{g}))) \geq (\varphi^{-1})' \circ (\varphi \circ \tilde{g}) D^+ f(\nabla(\varphi \circ \tilde{g})) = \frac{1}{\varphi' \circ \tilde{g}} D^+ f(\nabla(\varphi \circ \tilde{g})),
$$

$\mathfrak{m}$-a.e.. From the locality principle (2.15) and the arbitrariness of z we conclude that (3.30) holds for generic C^1_{loc} functions φ. The general case follows by approximating φ with a sequence (φ_n) of C^1_{loc} functions such that $\mathcal{L}^1(\{\varphi_n \neq \varphi\} \cup \{\varphi'_n \neq \varphi'\}) \to 0$ as $n \to \infty$ and using again the locality principle. $\qquad \square$

The following simple lemma will be useful in proving the Leibniz rule.

Lemma 3.16 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p, q \in (1, \infty)$ conjugate exponents, $f_1, f_2 \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m}) \cap L^\infty(X, \mathfrak{m})$, $g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and $\boldsymbol{\pi} \in \mathscr{P}(C([0,1], X))$ a plan q-representing ∇g.*
Then

$$
\lim_{t \downarrow 0} \int \left| \frac{(f_1(\gamma_t) - f_1(\gamma_0))(f_2(\gamma_t) - f_2(\gamma_0))}{t} \right| \, d\boldsymbol{\pi}(\gamma) = 0.
$$

proof Let $F_t^i(\gamma) := f_i(\gamma_t) - f_i(\gamma_0)$, $i = 1, 2$, $t \in (0, 1]$. From (3.19) we know that the family $\frac{F_t^i}{t}$ is dominated in $L^q(\boldsymbol{\pi})$, so that from the second part of Proposition 2.11 we get that the family $\frac{F_t^1}{t}$ is dominated in $L^1(\boldsymbol{\pi})$. By definition we also know that $\|F_t^2\|_{L^\infty(\boldsymbol{\pi})} \leq 2\|f_2\|_{L^\infty(\mathfrak{m})}$ for any $t \in (0, 1]$. Hence to conclude it is sufficient to check that $\boldsymbol{\pi}$-a.e. it holds $F_t^2 \to 0$ as $t \downarrow 0$. But this is obvious, because $\frac{F_t^2}{t}$ is dominated in $L^1(\boldsymbol{\pi})$ as well. $\qquad \square$

Proposition 3.17 (Leibniz rule) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p \in (1, \infty)$, $f_1, f_2 \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m}) \cap L^\infty_{\mathrm{loc}}(X, \mathfrak{m})$ and $g \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$.*
Then $\mathfrak{m}$-a.e. it holds

$$
\begin{aligned}
D^+(f_1 f_2)(\nabla g) &\leq f_1 D^{s_1} f_2(\nabla g) + f_2 D^{s_2} f_1(\nabla g), \\
D^-(f_1 f_2)(\nabla g) &\geq f_1 D^{-s_1} f_2(\nabla g) + f_2 D^{-s_2} f_1(\nabla g),
\end{aligned} \tag{3.32}
$$

where $s_i := \mathrm{sign} f_i$, $i = 1, 2$.

proof With a cut-off argument and from the locality property (3.8), we can assume that $f_1, f_2 \in S^p(X, \mathsf{d}, \mathfrak{m}) \cap L^\infty(X, \mathfrak{m})$ and $g \in S^p(X, \mathsf{d}, \mathfrak{m})$. Also, replacing f_1, f_2 with $|f_1|, |f_2|$ and using Proposition 3.15 above, we reduce to the case $f_1, f_2 \geq 0$.

Under this assumptions, we will prove that for any measure $\mu \in \mathscr{P}(X)$ with $\mu \leq C\tilde{\mathfrak{m}}$ for some $C > 0$ (with $\tilde{\mathfrak{m}}$ as in (2.8), (2.9)) it holds

$$\int D^+(f_1 f_2)(\nabla g)|Dg|_w^{p-2}\,\mathrm{d}\mu \leq \int \Big(f_1 D^+ f_2(\nabla g) + f_2 D^+ f_1(\nabla g) \Big)|Dg|_w^{p-2}\,\mathrm{d}\mu, \qquad (3.33)$$

which implies, since μ is arbitrary and $\mathfrak{m} \ll \tilde{\mathfrak{m}}$, the first inequality in (3.32). The other one will then follow putting $-g$ in place of g and using (3.5).

Fix such μ and notice that if $f_1 f_2 = 0$ μ-a.e., then (3.33) is obvious, thus we can assume $\int f_i\,\mathrm{d}\mu > 0$, $i = 1, 2$.

Fix $\varepsilon > 0$, put $g_\varepsilon := g + \varepsilon f_1 f_2 \in S^p(X, \mathsf{d}, \mathfrak{m})$ and let $\boldsymbol{\pi}^\varepsilon$ be a plan q-representing ∇g_ε with $(\mathrm{e}_0)_\sharp \boldsymbol{\pi} = \mu$ (Theorem 3.14, here q is the conjugate exponent of p) and recall (3.4), (3.18) to get

$$\int D^+(f_1 f_2)(\nabla g)|Dg|_w^{p-2}\,\mathrm{d}\mu \leq \int D^-(f_1 f_2)(\nabla g_\varepsilon)|Dg_\varepsilon|_w^{p-2}\,\mathrm{d}\mu$$

$$\leq \varliminf_{t\downarrow 0} \int \frac{f_1(\gamma_t)f_2(\gamma_t) - f_1(\gamma_0)f_2(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}^\varepsilon(\gamma)$$

$$\leq \varlimsup_{t\downarrow 0} \int \frac{f_1(\gamma_t)f_2(\gamma_t) - f_1(\gamma_0)f_2(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}^\varepsilon(\gamma).$$

Taking into account Lemma 3.16 we have

$$\varlimsup_{t\downarrow 0} \int \frac{f_1(\gamma_t)f_2(\gamma_t) - f_1(\gamma_0)f_2(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}^\varepsilon(\gamma)$$
$$= \varlimsup_{t\downarrow 0} \int f_1(\gamma_0)\frac{f_2(\gamma_t) - f_2(\gamma_0)}{t} + f_2(\gamma_0)\frac{f_1(\gamma_t) - f_1(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}^\varepsilon(\gamma)$$
$$\leq \varlimsup_{t\downarrow 0} \int f_1(\gamma_0)\frac{f_2(\gamma_t) - f_2(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}^\varepsilon(\gamma) + \varlimsup_{t\downarrow 0} \int f_2(\gamma_0)\frac{f_1(\gamma_t) - f_1(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}^\varepsilon(\gamma).$$

From (3.26) we know that the plans $\boldsymbol{\pi}_i^\varepsilon := \frac{1}{\int f_i\,\mathrm{d}\mu}f_i \circ \mathrm{e}_0\boldsymbol{\pi}^\varepsilon$, $i = 1, 2$, q-represent ∇g_ε with $(\mathrm{e}_0)_\sharp \boldsymbol{\pi}_i^\varepsilon = \frac{1}{\int f_i\,\mathrm{d}\mu}f_i \circ \mathrm{e}_0\,\mu$. Hence using again (3.18) we obtain

$$\varlimsup_{t\downarrow 0} \int f_1(\gamma_0)\frac{f_2(\gamma_t) - f_2(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}^\varepsilon(\gamma) \leq \int f_1 D^+ f_2(\nabla g_\varepsilon)|Dg_\varepsilon|_w^{p-2}\,\mathrm{d}\mu,$$

$$\varlimsup_{t\downarrow 0} \int f_2(\gamma_0)\frac{f_1(\gamma_t) - f_1(\gamma_0)}{t}\,\mathrm{d}\boldsymbol{\pi}^\varepsilon(\gamma) \leq \int f_2 D^+ f_1(\nabla g_\varepsilon)|Dg_\varepsilon|_w^{p-2}\,\mathrm{d}\mu.$$

In summary, we proved that for any $\varepsilon > 0$ it holds

$$\int D^+(f_1 f_2)(\nabla g)|Dg|_w^{p-2}\,\mathrm{d}\mu \leq \int f_1 D^+ f_2(\nabla g_\varepsilon)|Dg_\varepsilon|_w^{p-2} + f_2 D^+ f_1(\nabla g_\varepsilon)|Dg_\varepsilon|_w^{p-2}\,\mathrm{d}\mu.$$

Notice now that the integrand in the right hand side is dominated in $L^1(X, \mu)$ and use (3.4) in considering its limit as $\varepsilon \downarrow 0$. The thesis follows. $\qquad\square$

Remark 3.18 (The infinitesimally strictly convex case) If $(X, \mathsf{d}, \mathfrak{m})$ is q-infinitesimally strictly convex, q being the conjugate exponent of p, then under the same assumptions of Propositions 3.15 and 3.17 we have the following $\mathfrak{m}$-a.e. equalities:

$$D(\varphi \circ f)(\nabla g) = \varphi' \circ f \; Df(\nabla g), \tag{3.34}$$

$$Df(\nabla(\varphi \circ g)) = \varphi' \circ g \; Df(\nabla g), \tag{3.35}$$

$$D(f_1 f_2)(\nabla g) = f_1 Df_2(\nabla g) + f_2 Df_1(\nabla g). \tag{3.36}$$

∎

Remark 3.19 It is natural to question whether it is valid a Leibniz rule of a form like

$$Df(\nabla(g_1 g_2)) = g_1 \, Df(\nabla g_2) + g_2 \, Df(\nabla g_1), \qquad \forall f, g_1, g_2 \text{ "smooth"} \tag{3.37}$$

provided possibly we replace the equality with an inequality and do appropriate sign choices in $D^\pm$.

This is well known to be false even in a smooth world. For instance, in the flat normed setting considered in Section 1.1, the above equality holds for any couple of smooth f, g if and only if the norm comes from a scalar product. Indeed, recall that $\nabla g = \mathrm{Dual}^{-1}(Dg)$, so that (3.37) holds for any f, g_1, g_2 if and only if

$$\mathrm{Dual}^{-1}(D(g_1 g_2)) = g_1 \mathrm{Dual}^{-1}(Dg_2) + g_2 \mathrm{Dual}^{-1}(Dg_1).$$

By the Leibniz rule for the differentials, we know that the left hand side is equal to $\mathrm{Dual}^{-1}(g_1 Dg_2 + g_2 Dg_1)$. Therefore, since Dual^{-1} is always 1-homogeneous, (3.37) is true if and only if Dual^{-1} is linear. That is, if and only if the norm $\| \cdot \|$ comes from a scalar product (compare this fact with the Leibniz rule (4.16) obtained in Section 4.3 under the assumption that the space is infinitesimally Hilbertian). ∎

Remark 3.20 (The different roles of horizontal and vertical derivation) In developing this abstract differential calculus, we proved the chain rule as a consequence of the 'vertical' derivations and the Leibniz rule as a byproduct of the 'horizontal' ones. We chose this path to present a first direct application of both the approaches and to show how they relate, but actually this is unnecessary, because the Leibniz rule can also be proved via the 'vertical' approach only. Indeed, notice that as in the proof of Proposition 3.17 we can reduce to prove that

$$D^+(f_1 f_2)(\nabla g) \le f_1 D^+ f_2(\nabla g) + f_2 D^+ f_1(\nabla g), \qquad \mathfrak{m} - a.e.,$$

for positive $f_1, f_2 \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m}) \cap L^\infty(X, \mathfrak{m})$ and $g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$. Then use the chain rule[1] (3.29) and the convexity and positive 1-homogeneity of $f \mapsto D^+ f(\nabla g)$ to get

$$D^+(f_1 f_2)(\nabla g) = f_1 f_2 \, D^+(\log(f_1 f_2))(\nabla g) \le f_1 f_2 \left(D^+(\log(f_1))(\nabla g) + D^+(\log(f_2))(\nabla g) \right)$$
$$= f_1 D^+ f_2(\nabla g) + f_2 D^+ f_1(\nabla g), \qquad \mathfrak{m} - a.e..$$

On the other hand, a calculus can also be developed starting from the 'horizontal' approach only (but this would be much more cumbersome to do w.r.t. Definition 3.1).

---

[1] I saw this approach to the proof of the Leibniz rule at the page http://planetmath.org/?op=getobj&from=objects&id=8438, where it is acknowledged to W. Buck

Therefore the point is not which one of the two approaches is the best, because they produce an equally powerful calculus. The point is that they produce the same calculus. Notice that the 'horizontal' derivation is common when working with the Wasserstein geometry, while the 'vertical' one is typical of the L^p world. So that every time these two geometries are put in relation, in a way or another the same principle underlying Theorem 3.10 comes into play. For instance, the relation between horizontal and vertical derivation was already present, although well hidden, in the first proof in a non-smooth context of the equivalence between the gradient flow of Ch_2 in $L^2(X, \mathfrak{m})$ and the one of the relative entropy in the Wasserstein space $(\mathscr{P}_2(X), W_2)$, provided in the case of Alexandrov spaces with lower curvature bounds in [22].

In several situations, one specific approach works better than the other, and their interaction has fruitful consequences. For instance, in getting the Laplacian comparison out of the $\mathsf{CD}(K, N)$ condition (see the introduction and Theorem 5.14) it is crucial to take an horizontal derivative, which is not surprising given that the $\mathsf{CD}(K, N)$ condition is given in terms of the Wasserstein geometry: the same result seems hardly achievable working with the vertical approach only. On the other hand, the vertical approach is needed if one wants to know under which conditions the limit of

$$\int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma),$$

exists as $t \downarrow 0$, where here $f \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and $\boldsymbol{\pi}$ q-represents ∇g for some $g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$. In other words, if we want to know up to which extent we can consider plans $\boldsymbol{\pi}$ as differentiation operators. Using Theorem 3.10 we know that this is the case as soon as $(X, \mathsf{d}, \mathfrak{m})$ is q-infinitesimally strictly convex, a fact which would otherwise be not at all obvious from the horizontal approach only. Also, consider the following question: let $f_1, f_2 \in \mathrm{S}^2(X, \mathsf{d}, \mathfrak{m})$ and $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2$ plans of bounded compression with $(\mathrm{e}_0)_\sharp \boldsymbol{\pi}^1 = (\mathrm{e}_0)_\sharp \boldsymbol{\pi}^2 =: \mu$ such that $\boldsymbol{\pi}_i$ 2-represents ∇f_i, $i = 1, 2$. Assume that the seminorm $\| \cdot \|_{\mathrm{S}^2}$ satisfies the parallelogram rule. Can we deduce that the equality

$$\lim_{t \downarrow 0} \int \frac{f_1(\gamma_t) - f_1(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}_2(\gamma) = \lim_{t \downarrow 0} \int \frac{f_2(\gamma_t) - f_2(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}_1(\gamma),$$

holds? This is expected, as the fact that $\| \cdot \|_{\mathrm{S}^2}$ satisfies the parallelogram rule should mean that we can identify differentials and gradients, so that both the expressions above should be equal to something like

$$\int \nabla f_1 \cdot \nabla f_2 \, \mathrm{d}\mu.$$

This is indeed the case, as we will see in the Chapter 4.3, and once again this will follow from the 'vertical' approach, rather than from the 'horizontal' one (studies in directions have been firstly made in [6]). ∎

Remark 3.21 (Comparison with Cheeger's differential structure) In the seminal paper [13], Cheeger built a differential structure and defined the differential of Lipschitz functions in a completely different way. Working under the assumption that the measure is doubling and that the space supports a weak-local Poincaré inequality, he proved (among other things) that the space has a finite dimensional cotangent space, in the following sense. There exists $N \in \mathbb{N}$, a measurable decomposition $X = \sqcup_{i \in \mathbb{N}} U_i$ and Lipschitz functions $\mathbf{x}_i : U_i \to \mathbb{R}^{N_i}$,

$N_i \leq N$, such that for any $i \in \mathbb{N}$ and any Lipschitz function $f : X \to \mathbb{R}$ there exists a unique - up to negligible sets - map $df_i : U_i \to \mathbb{R}^{N_i}$ such that

$$\lim_{y \to x} \left| \frac{f(y) - f(x) - df_i(x) \cdot \big(\mathbf{x}_i(y) - \mathbf{x}_i(x)\big)}{\mathsf{d}(y,x)} \right| = 0, \qquad \text{for } \mathfrak{m} - a.e. \ x \in U_i. \tag{3.38}$$

This result can be interpreted as: the U_i are coordinate charts, the map $\mathbf{x}_i$ provides the coordinate maps of the chart and formula (3.38) is the first order Taylor expansion of f in this chart.

Relying on this result, he then interpreted (i.e. defined) the function df_i as the differential of the Lipschitz map f on the set U_i. From the construction, it is obvious that it satisfies the Leibniz rule in the following form, sharper than (3.32):

$$d(f_1 f_2)_i = f_1(df_2)_i + f_2(df_1)_i. \tag{3.39}$$

In particular, this holds without any sort of infinitesimal strict convexity. Yet, even in Cheeger's paper the issue of possibly not strictly convex norms in tangent spaces arose, in particular in connection with the surjectivity of a certain natural map on tangent cones. This issue was worked out using the finite dimensionality of the cotangent spaces to define equivalent well-behaved norms: his resulting Theorem 13.4 in [13] should be compared with our, weaker but valid in higher generality, Theorem A.6.

Using the tools described in Appendix A (see Theorems A.4 and A.6) one can check that (3.32) is a consequence of (3.39), which means in particular that the two constructions are compatible.

The fact that formula (3.32) is in form of inequalities rather than being an equality, is due to the fact that with the approach we proposed 'we don't really know who is the differential of a function before applying it to a gradient', and gradients in general are not uniquely defined. On the contrary, Cheeger's approach allows for a more intrinsic definition of differential, but it passes through the finite dimensionality of the cotangent space, a property which on arbitrary spaces certainly is not satisfied. ∎

4 Laplacian

4.1 Definition and basic properties

Having in mind the discussion made in Section 1.1 and the tools developed in the previous chapter, we will now discuss the definition of Laplacian. We start noticing that Proposition 2.6 allows to define the Sobolev class $S_{\text{loc}}^p(\Omega)$ in the following way:

Definition 4.1 (The Sobolev class $S_{\text{loc}}^p(\Omega)$) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $\Omega \subset X$ an open set and $p \in (1, \infty)$. The space $S_{\text{loc}}^p(\Omega)$ is the space of Borel functions $g : \Omega \to \mathbb{R}$ such that $g\chi \in S_{\text{loc}}^p(X, \mathsf{d}, \mathfrak{m})$ for any Lipschitz function $\chi : X \to [0,1]$ such that $\mathsf{d}(\text{supp}(\chi), X \setminus \Omega) > 0$, where $g\chi$ is taken 0 by definition on $X \setminus \Omega$.*

Here and in the following for $A, B \subset X$, we denote by $\mathsf{d}(A, B)$ the value of $\inf \mathsf{d}(x, y)$, where the infimum is taken among all $x \in A$ and $y \in B$.

It is immediate to verify that this is a good definition, and that for $g \in S_{\text{loc}}^p(\Omega)$ the function $|Dg|_w \in L_{\text{loc}}^p(\Omega, \mathfrak{m}_{|\Omega})$ is well defined, thanks to (2.15), by

$$|Dg|_w := |D(g\chi)|_w, \qquad \mathfrak{m} - a.e. \ \text{ on } \{\chi = 1\}, \tag{4.1}$$

where $\chi : X \to [0,1]$ is any Lipschitz function such that $\mathsf{d}(\mathrm{supp}(\chi), X \setminus \Omega) > 0$. Using the locality property (3.8) it is easy to see that the objects $D^{\pm}f(\nabla g)$ are well defined $\mathfrak{m}$-a.e. on Ω and that the chain rules (3.29), (3.30) and the Leibniz rule (3.32) hold $\mathfrak{m}$-a.e. on Ω for functions $f, g \in \mathrm{S}^p_{\mathrm{loc}}(\Omega)$, $p \in (1, \infty)$.

To define the Laplacian of g we will need to ask for an integrability of $|Dg|_w$ which on non-proper spaces is slightly stronger than $L^p_{\mathrm{loc}}(\Omega, \mathfrak{m}_{|\Omega})$:

Definition 4.2 (The class $\mathrm{Int}(\Omega)$ and 'internal' Sobolev spaces) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) and $\Omega \subset X$ an open set. The class $\mathrm{Int}(\Omega)$ is the collection of all open sets $\Omega' \subset \Omega$ such that*

i) Ω' is bounded,

ii) $\mathsf{d}(\Omega', X \setminus \Omega) > 0$,

iii) $\mathfrak{m}(\Omega') < \infty$.

For $p \in (1, \infty)$ we define the space $\mathrm{S}^p_{\mathrm{int}}(\Omega)$ of functions Sobolev internally in Ω as

$$\mathrm{S}^p_{\mathrm{int}}(\Omega) := \Big\{ g \in \mathrm{S}^p_{\mathrm{loc}}(\Omega) \ : \ \int_{\Omega'} |Dg|^p_w \, \mathrm{d}\mathfrak{m} < \infty, \ \forall \Omega' \in \mathrm{Int}(\Omega) \Big\}.$$

We also introduce the class $\mathrm{Test}(\Omega)$ of test functions in Ω as follows.

Definition 4.3 (Test functions) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) and $\Omega \subset X$ an open set. The space $\mathrm{Test}(\Omega)$ is defined by*

$$\mathrm{Test}(\Omega) := \Big\{ f \in \mathrm{LIP}(X) : \ \mathrm{supp}(f) \subset \Omega', \ \ for \ some \ \ \Omega' \in \mathrm{Int}(\Omega) \}.$$

Recall that if $f : X \to \mathbb{R}$ is a Lipschitz function, then $f \in \mathrm{S}^p_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ for any $p \in (1, \infty)$ and the p-minimal weak upper gradient $|Df|_w$ is uniformly bounded by $\mathrm{Lip}(f)$.

Clearly, for $g \in \mathrm{S}^p_{\mathrm{int}}(\Omega)$ and $f \in \mathrm{Test}(\Omega)$ it holds $g + \varepsilon f \in \mathrm{S}^p_{\mathrm{int}}(\Omega)$ for any $\varepsilon \in \mathbb{R}$. Thus for $g \in \mathrm{S}^p_{\mathrm{int}}(\Omega)$ and $f \in \mathrm{Test}(\Omega)$, thanks to (4.1) the functions $D^{\pm}f(\nabla g)$ are $\mathfrak{m}$-a.e. defined on Ω by

$$D^{\pm}f(\nabla g) := D^{\pm}f(\nabla(g\chi)), \qquad \text{on } \{\chi = 1\}, \tag{4.2}$$

for any $\chi : X \to [0,1]$ Lipschitz with $\mathrm{supp}(\chi) \subset \Omega'$ for some $\Omega' \in \mathrm{Int}(\Omega)$. From the bound (3.2) and the fact that $\mathfrak{m}(\mathrm{supp}(f)) < \infty$ we have $D^{\pm}f(\nabla g) \in L^1(\Omega, \mathfrak{m}_{|\Omega})$.

We are now ready to give the definition of distributional Laplacian.

Definition 4.4 (Laplacian) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $\Omega \subset X$ an open set and $g : \Omega \to \mathbb{R}$ a Borel function. We say that g is in the domain of the Laplacian in Ω, and write $g \in D(\boldsymbol{\Delta}, \Omega)$, provided $g \in \mathrm{S}^p_{\mathrm{int}}(\Omega)$ for some $p > 1$ and there exists a Radon measure μ on Ω such that for any $f \in \mathrm{Test}(\Omega) \cap L^1(\Omega, |\mu|)$ it holds*

$$-\int_\Omega D^+ f(\nabla g) \, \mathrm{d}\mathfrak{m} \leq \int_\Omega f \, \mathrm{d}\mu \leq -\int_\Omega D^- f(\nabla g) \, \mathrm{d}\mathfrak{m}. \tag{4.3}$$

In this case we write $\mu \in \boldsymbol{\Delta}g_{|\Omega}$. If $\Omega = X$ we simply write $g \in D(\boldsymbol{\Delta})$ and $\mu \in \boldsymbol{\Delta}g$.

We are using the bold font for the Laplacian to emphasize that in our discussion it will always be a measure, possibly multivalued.

Remark 4.5 As for the objects $|Df|_w$ and $D^\pm f(\nabla g)$, the choice of the Sobolev exponent p might affect the definition of Laplacian, but in our writing we will omit such explicit dependence (see also Remark 2.5). In this direction, notice that we require in the definition that the function g belongs to $\mathrm{S}^p_{\mathrm{int}}(\Omega)$ for some $p > 1$: the exponent p is then the one chosen to compute $|D(g + \varepsilon f)|_w$, which in turn is the tool used to introduce $D^\pm f(\nabla g)$. ∎

Remark 4.6 (A coincidence about a priori regularity) We already observed in Section 1.1 that in a non-Riemannian situation, in order to define the distributional Laplacian of a function g, we need to impose some a priori Sobolev regularity on it. The minimal one is then that - in the appropriate local sense - $g \in \mathrm{S}^1(\Omega)$. Yet, we didn't define the Sobolev class S^1, because its definition requires some care (for recent results in this direction, see [1]), so that we just asked for g to locally belong to S^p for some $p > 1$.

Notice that not only this allows to write down $D^\pm f(\nabla g)$, but is in perfect duality with the integrability of $|Df|_w$ for the smoothest functions available in this setting, i.e. Lipschitz f's, because as we already remarked it holds $|Df|_w \in L^\infty$.

Here there is an interesting coincidence. Definition 4.4 differs from the one of distributional Laplacian on the Euclidean setting $\mathbb{R}^d$, as in this framework in order for the expression $\Delta g = \mu$ to make sense, the only a priori condition that is imposed on g is to be locally in L^1. Yet, standard regularity theory gives that if $\Delta g = \mu$ in the sense of distributions, then the distributional gradient ∇g is a function which locally belongs to L^p for any $p \in [1, 1 + \frac{1}{d})$. Therefore if we are in a Euclidean setting and we 'don't know the dimension', we are anyway sure that $\Delta g = \mu$ implies $\nabla g \in L^p_{\mathrm{loc}}(X, \mathfrak{m})$ for some $p > 1$.

This is precisely the a priori condition on g we asked for in Definition 4.4. Which means that we are, in practice, not asking more than what is already true in the classical situation. In particular, our definition reduces to the standard one in a Euclidean setting. ∎

A drawback of Definition 4.4 is that in general the Laplacian might be non-unique: in Proposition 4.9 below we will actually see that explicit counterexamples to uniqueness can be produced as soon as the space is not 2-infinitesimally strictly convex. Yet, it is immediate to verify that the Laplacian is one-homogeneous, i.e.

$$g \in D(\boldsymbol{\Delta}, \Omega), \ \mu \in \boldsymbol{\Delta} g_{|\Omega} \qquad \Rightarrow \qquad \lambda g \in D(\boldsymbol{\Delta}, \Omega), \ \lambda \mu \in \boldsymbol{\Delta}(\lambda g)_{|\Omega}, \quad \forall \lambda \in \mathbb{R}, \qquad (4.4)$$

invariant by addition of a constant, i.e.

$$g \in D(\boldsymbol{\Delta}, \Omega), \ \mu \in \boldsymbol{\Delta} g_{|\Omega} \qquad \Rightarrow \qquad c + g \in D(\boldsymbol{\Delta}, \Omega), \ \mu \in \boldsymbol{\Delta}(c + g)_{|\Omega}, \quad \forall c \in \mathbb{R},$$

that the set $\boldsymbol{\Delta} g_{|\Omega}$ is convex and that it is weakly closed in the following sense: if $\mu_n \in \boldsymbol{\Delta} g_{|\Omega}$ for any $n \in \mathbb{N}$ and μ is a Radon measure on Ω such that $f \in \mathrm{Test}(\Omega) \cap L^1(\Omega, |\mu|)$ implies $f \in L^1(\Omega, |\mu_n|)$ for n large enough and

$$\lim_{n \to \infty} \int f \, \mathrm{d}\mu_n = \int f \, \mathrm{d}\mu,$$

then $\mu \in \boldsymbol{\Delta} g_{|\Omega}$ (see also the proof of Proposition 4.15).

Another direct consequence of the definition is the following 'global-to-local' property

$$\widetilde{\Omega} \subset \Omega \text{ open sets}, \ g \in D(\boldsymbol{\Delta}, \Omega), \ \mu \in \boldsymbol{\Delta} g_{|\Omega} \qquad \Rightarrow \qquad g \in D(\boldsymbol{\Delta}, \widetilde{\Omega}), \ \mu_{|\widetilde{\Omega}} \in \boldsymbol{\Delta} g_{|\widetilde{\Omega}}, \quad (4.5)$$

and the fact that measures in $\boldsymbol{\Delta} g_{|\Omega}$ are always concentrated on $\mathrm{supp}(\mathfrak{m})$.

By the definition of $D^\pm f(\nabla g)$, we easily get the following sufficient condition for the Laplacian to be unique.

Proposition 4.7 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $\Omega \subset X$ an open set, $p \in (1, \infty)$, $g \in \mathrm{S}^p_{\mathrm{int}}(\Omega) \cap D(\mathbf{\Delta}, \Omega)$. Assume that $(X, \mathsf{d}, \mathfrak{m})$ is q-infinitesimally strictly convex, q being the conjugate exponent of p.*

Then $\mathbf{\Delta} g_{|\Omega}$ contains only one measure.

proof By the definition (4.2) and (3.11), we deduce that for any $f \in \mathrm{Test}(\Omega)$ it holds $D^+ f(\nabla g) = D^- f(\nabla g)$ $\mathfrak{m}$-a.e. on Ω. The thesis follows. $\qquad\square$

The definition of Laplacian we just proposed is compatible with those available on Alexandrov spaces with curvature bounded from below, see Remark 4.25.

We conclude this introduction by comparing our definition with the one appeared in [4]. Recall the definition given in (2.26) of the Cheeger energy functional $\mathrm{Ch}_2 : L^2(X, \mathfrak{m}) \to [0, \infty]$, and that Ch_2 is convex and lower semicontinuous.

Definition 4.8 (Laplacian as defined in [4]) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) and $g \in L^2(X, \mathfrak{m})$. We say that g is in the domain of the Laplacian provided $\mathrm{Ch}_2(g) < \infty$ and the subdifferential $\partial^- \mathrm{Ch}_2(g)$ of Ch_2 at g is non-empty. In this case the Laplacian of g is the element of minimal $L^2(X, \mathfrak{m})$ norm in (the closed convex set) $-\partial^- \mathrm{Ch}_2(g)$.*

Proposition 4.9 (Compatibility) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $g \in L^2(X, \mathfrak{m})$ and $h \in L^2(X, \mathfrak{m})$. Assume that $\mathrm{Ch}_2(g) < \infty$ and $-h \in \partial^- \mathrm{Ch}_2(g)$. Then $g \in D(\mathbf{\Delta})$ and $h\mathfrak{m} \in \mathbf{\Delta} g$.*

proof Fix $f \in L^2(X, \mathfrak{m})$ with $\mathrm{Ch}_2(f) < \infty$ and notice that the definition of subdifferential yields

$$\mathrm{Ch}_2(g) - \varepsilon \int f h \, \mathrm{d}\mathfrak{m} \leq \mathrm{Ch}_2(g + \varepsilon f), \qquad \forall \varepsilon \in \mathbb{R}.$$

Hence for $\varepsilon > 0$ it holds

$$- \int f h \, \mathrm{d}\mathfrak{m} \leq \int \frac{|D(g + \varepsilon f)|^2_w - |Dg|^2_w}{2\varepsilon} \, \mathrm{d}\mathfrak{m}.$$

Passing to the limit as $\varepsilon \downarrow 0$ and noticing that the integrand at the right hand side is dominated, we get

$$- \int f h \, \mathrm{d}\mathfrak{m} \leq \int D^+ f(\nabla g) \, \mathrm{d}\mathfrak{m}.$$

In particular, this holds for $f \in \mathrm{Test}(X) \subset \{\mathrm{Ch}_2 < \infty\} \cap L^1(X, |h|\mathfrak{m})$. Replacing f with $-f$ we also get the other inequality in (4.3) and the conclusion. $\qquad\square$

Therefore from the perspective of Definition 4.4 there is nothing special about the element of minimal norm in $-\partial^- \mathrm{Ch}_2(g)$, as any of its elements is (the density of) an admissible distributional Laplacian. This is not so surprising, as in [4] the choice of the element of minimal norm was done essentially only for cosmetic reasons, as it allowed to identify uniquely who the Laplacian was, and to write the equation of the gradient flow of Ch_2 in the form $\frac{\mathrm{d}}{\mathrm{d}t} f_t = \Delta f_t$. Yet, the weak integration by part rules provided in [4] are valid, as the proof shows, for all the elements in $-\partial^- \mathrm{Ch}_2(g)$.

The main technical difference between the approach in [4] and the current one is then - beside the integrability assumption on g and $|Dg|_w$ - that in [4] the Laplacian is well defined 'only when it belongs to L^2', while here we are giving a meaning to the abstract distributional Laplacian, allowing it to be a measure.

Remark 4.10 It is natural to ask if the converse of Proposition 4.9 holds, i.e. if the following is true: suppose that $g \in D(\mathrm{Ch}_2) \cap D(\mathbf{\Delta})$ and that for some $\mu \in \mathbf{\Delta} g$ it holds $\mu \ll \mathfrak{m}$ with density $h \in L^2(X, \mathfrak{m})$. Can we deduce that $-h \in \partial^- \mathrm{Ch}_2(g)$? Without further assumptions, it is unclear to us whether this is true or not, the point being that we would like to deduce from

$$\mathrm{Ch}_2(g) - \int fh \, \mathrm{d}\mathfrak{m} \leq \mathrm{Ch}_2(g + f), \qquad \forall f \in \mathrm{Test}(X),$$

that

$$\mathrm{Ch}_2(g) - \int fh \, \mathrm{d}\mathfrak{m} \leq \mathrm{Ch}_2(g + f), \qquad \forall f \in L^2(X, \mathfrak{m}).$$

This implication seems to rely on a density result of Lipschitz functions in $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ finer than the one expressed in Theorem 2.8. Indeed, we would like to know that for any $f, g \in D(\mathrm{Ch}_2)$ there exists a sequence of Lipschitz functions (f_n) converging to f in $L^2(X, \mathfrak{m})$ such that $\mathrm{Ch}_2(g + f_n)$ converges to $\mathrm{Ch}_2(g + f)$ (Theorem 2.8 ensures this fact only when g is itself Lipschitz).

If $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is uniformly convex and $\mathfrak{m}$ is finite on bounded sets, then such approximation result is indeed true, because Corollary 2.9 ensures the density of $\mathrm{Test}(X)$ in $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$, see for instance the argument in Proposition 4.24 for the case in which $W^{1,2}$ is Hilbert. ∎

4.2 Calculus rules

In this section we collect the basic calculus properties of the Laplacian.

Proposition 4.11 (Chain rule) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $\Omega \subset X$ an open set, $g \in D(\mathbf{\Delta}, \Omega)$. Assume that g is Lipschitz on Ω' for any $\Omega' \in \mathrm{Int}(\Omega)$ and let $\varphi : g(\Omega) \to \mathbb{R}$ a $C^{1,1}_{\mathrm{loc}}$ map. Then $\varphi \circ g \in D(\mathbf{\Delta}, \Omega)$ and for any $\mu \in \mathbf{\Delta} g_{|\Omega}$ it holds*

$$\mathbf{\Delta}(\varphi \circ g)_{|\Omega} \ni \tilde{\mu} := \varphi' \circ g \, \mu + \varphi'' \circ g \, |Dg|_w^2 \mathfrak{m}. \tag{4.6}$$

proof Notice that since g is continuous, the expression $\varphi' \circ g \, \mu$ makes sense and defines a locally finite measure. Similarly, $\varphi'' \circ g \, |Dg|_w^2 \in L^\infty_{\mathrm{loc}}(\Omega, \mathfrak{m}_{|\Omega})$, so that $\tilde{\mu}$ is a locally finite measure and the statement makes sense.

More precisely, $\varphi' \circ g$ is Lipschitz on Ω' and $\varphi'' \circ g \, |Dg|_w^2$ is bounded on Ω' for any $\Omega' \in \mathrm{Int}(\Omega)$. Therefore if $f \in \mathrm{Test}(\Omega) \cap L^1(\Omega, |\tilde{\mu}|)$, we have $f\varphi' \circ g \in \mathrm{Test}(\Omega) \cap L^1(|\mu|)$. Fix such f, recall the chain rules (3.29), (3.30), the Leibniz rule (3.32) and (3.6) to get that $\mathfrak{m}$-a.e. on Ω it holds

$$D^+ f \nabla(\varphi \circ g) = \varphi' \circ g D^{\mathrm{sign}\varphi' \circ g} f \nabla g \geq D^+(f\varphi' \circ g)(\nabla g) - f D^{\mathrm{sign}f}(\varphi' \circ g)(\nabla g)$$

$$= D^+(f\varphi' \circ g)(\nabla g) - f\varphi'' \circ g D^{\mathrm{sign}(f\varphi'' \circ g)}(g)(\nabla g)$$

$$= D^+(f\varphi' \circ g)(\nabla g) - f\varphi'' \circ g |Dg|_w^2.$$

Integrating we obtain $\int D^+ f(\nabla(\varphi \circ g)) \, \mathrm{d}\mathfrak{m} \geq - \int f \, \mathrm{d}\tilde{\mu}$. Replacing f with $-f$ we conclude. □

Remark 4.12 The assumption that g was Lipschitz in Ω' was needed to ensure that $f \in \mathrm{Test}(\Omega)$ implies $f\varphi' \circ g \in \mathrm{Test}(\Omega)$, so that from assumptions one could deduce

$$\int D^+(f\varphi' \circ g)(\nabla g) \, \mathrm{d}\mathfrak{m} \geq - \int f \, \mathrm{d}\mu.$$

Yet, it would be more natural to have the chain rule to hold under the only assumption that $g \in S^2_{\text{int}}(\Omega) \cap C(\Omega)$, because these are the minimal regularity requirements needed for the right hand side of (4.6) to make sense.

To get the chain rule under this weaker assumption on g, a suitable generalization of the integration by parts rule is needed, which looks linked to the uniform convexity of $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$. See for instance the proof of Lemma 4.26 for the case of proper spaces such that $W^{1,2}$ is Hilbert. ∎

The following proposition provides a sufficient condition on g which ensures that it is in the domain of the Laplacian and gives a bound on elements of $\mathbf{\Delta} g_{|\Omega}$. Technicalities apart, it can be seen as a rewording of the statement 'a non-negative distribution is always a non-negative measure' valid in a Euclidean context.

Proposition 4.13 (Laplacian existence and comparison) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) and assume also that (X, d) is proper. Let $\Omega \subset X$ be an open set, $g \in S^p_{\text{int}}(\Omega)$, $p \in (1, \infty)$, and $\tilde{\mu}$ a Radon measure on Ω. Assume that for any non-negative $f \in \text{Test}(\Omega) \cap L^1(\Omega, |\tilde{\mu}|)$ we have*

$$- \int_\Omega D^- f(\nabla g) \, \mathrm{d}\mathfrak{m} \leq \int_\Omega f \, \mathrm{d}\tilde{\mu}. \tag{4.7}$$

Then $g \in D(\mathbf{\Delta}, \Omega)$ and for any $\mu \in \mathbf{\Delta} g_{|\Omega}$ it holds $\mu \leq \tilde{\mu}$.

proof Consider the real valued map $\text{Test}(\Omega) \ni f \mapsto T(f) := - \int_\Omega D^- f(\nabla g) \, \mathrm{d}\mathfrak{m}$. From Proposition 3.2 we know that it satisfies

$$T(\lambda f) = \lambda T(f), \qquad \forall f \in \text{Test}(\Omega), \ \lambda \geq 0,$$
$$T(f_1 + f_2) \leq T(f_1) + T(f_2), \qquad \forall f_1, f_2 \in \text{Test}(\Omega).$$

Hence by the Hahn-Banach theorem there exists a linear map $L : \text{Test}(\Omega) \to \mathbb{R}$ such that $L(f) \leq T(f)$ for any $f \in \text{Test}(\Omega)$. By (3.5) we then have

$$- \int_\Omega D^+ f(\nabla g) \, \mathrm{d}\mathfrak{m} \leq L(f) \leq - \int_\Omega D^- f(\nabla g) \, \mathrm{d}\mathfrak{m}, \qquad \forall f \in \text{Test}(\Omega).$$

Our goal is to prove that L can be represented as integral w.r.t. some measure μ. By (4.7) we get

$$\int_\Omega f \, \mathrm{d}\tilde{\mu} - L(f) \geq 0, \qquad \forall f \in \text{Test}(\Omega), \ f \geq 0.$$

Fix a compact set $K \subset \Omega$ and a function $\chi_K \in \text{Test}(\Omega)$ such that $0 \leq \chi_K \leq 1$ everywhere and $\chi_K = 1$ on K. Let $V_K \subset \text{Test}(\Omega)$ be the set of those functions with support contained in K and observe that for any non-negative $f \in V_K$, the fact that $(\max f)\chi_K - f$ is in $\text{Test}(\Omega)$ and non-negative yields

$$L(f) = -L((\max f)\chi_K - f) + L((\max f)\chi_K) \geq - \int (\max f)\chi_K - f \, \mathrm{d}\tilde{\mu} + L((\max f)\chi_K)$$
$$\geq - (\max f)\big(\tilde{\mu}(\text{supp}(\chi_K)) + L(\chi_K)\big).$$

Thus for a generic $f \in V_K$ it holds

$$L(f) = L(f^+ - f^-) = L(f^+) - L(f^-) \leq \int_\Omega f^+ \, \mathrm{d}\tilde{\mu} + (\max f^-)\big(\tilde{\mu}(\text{supp}(\chi_K)) + L(\chi_K)\big)$$
$$\leq (\max |f|)\Big(\tilde{\mu}(K) + \tilde{\mu}(\text{supp}(\chi_K)) + L(\chi_K)\Big),$$

i.e. $L : V_K \to \mathbb{R}$ is continuous w.r.t. the sup norm. Hence it can be extended to a (unique, by the density of Lipschitz functions in the uniform norm) linear bounded functional on the set $C_K \subset C(X)$ of continuous functions with support contained in K. Since K was arbitrary, by the Riesz theorem we get that there exists a Radon measure μ such that

$$L(f) = \int f \, \mathrm{d}\mu, \qquad \forall f \in \mathrm{Test}(\Omega).$$

Thus $g \in D(\boldsymbol{\Delta}, \Omega)$ and $\mu \in \boldsymbol{\Delta}g_{|\Omega}$.

Finally, from (4.7) it is immediate to get that for any $\mu' \in \boldsymbol{\Delta}g_{|\Omega}$ it holds $\mu' \leq \tilde{\mu}$. $\qquad \square$

Remark 4.14 In the proof of Proposition 4.13 we used the Hahn-Banach theorem, which means that some weak form of the Axiom of Choice (but still stronger than the Axiom of Dependent Choice) is required in the argument. In full generality, it seems hard to avoid such use of Hahn-Banach.

Yet, if furthermore $(X, \mathsf{d}, \mathfrak{m})$ is q-infinitesimally strictly convex and $g \in \mathrm{S}^p_{\mathrm{int}}(\Omega)$ - p, q being conjugate exponents - then the map T is linear and the proof goes on without any use of Choice. This is the situation we will work on when proving our main Laplacian comparison result in Theorem 5.14. $\qquad \blacksquare$

We now turn to the stability of $\boldsymbol{\Delta}g$ under convergence of g. Since in our definition of Laplacian we required some Sobolev regularity of g, in order to get stability we need to impose some sort of Sobolev convergence as well. This is clearly a much stronger assumption than the usual convergence in L^1_{loc} which is sufficient in the Euclidean context, but looks unavoidable in this general non-linear setting.

Proposition 4.15 (Stability) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $\Omega \subset X$ an open set and $p \in (1, \infty)$. Let $(g_n) \subset \mathrm{S}^p_{\mathrm{int}}(\Omega)$ be a sequence and $g \in \mathrm{S}^p_{\mathrm{int}}(\Omega)$ be such that $\int_{\Omega'} |D(g_n - g)|_w \, \mathrm{d}\mathfrak{m} \to 0$ for every $\Omega' \in \mathrm{Int}(\Omega)$. Assume also that $g_n \in D(\boldsymbol{\Delta}, \Omega)$ for every $n \in \mathbb{N}$, let $\mu_n \in \boldsymbol{\Delta}g_{n|\Omega}$ and suppose that for some locally finite measure μ on Ω it holds*

$$f \in \mathrm{Test}(\Omega) \cap L^1(\Omega, |\mu|) \qquad \Rightarrow \qquad \begin{cases} f \in L^1(\Omega, |\mu_n|) \text{ for } n \text{ large enough and} \\[2mm] \displaystyle\lim_{n \to \infty} \int f \, \mathrm{d}\mu_n = \int f \, \mathrm{d}\mu. \end{cases} \tag{4.8}$$

Then $g \in D(\boldsymbol{\Delta}, \Omega)$ and $\mu \in \boldsymbol{\Delta}g_{|\Omega}$.

Note: if $(X, \mathsf{d}, \mathfrak{m})$ is proper and the measures μ_n are equibounded - in total variation norm - on compact subsets of Ω, then the convergence in (4.8) is nothing but weak convergence in duality with $C_c(\Omega)$.

proof The argument is similar to that of the proof of Proposition 3.2. It is easy to check that for any $f \in \mathrm{Test}(\Omega)$, $g \in \mathrm{S}^p_{\mathrm{int}}(\Omega)$ it holds

$$\int D^+ f(\nabla g) \, \mathrm{d}\mathfrak{m} = \inf_{\varepsilon > 0} \int |Dg|_w \frac{|D(g + \varepsilon f)|_w - |Dg|_w}{\varepsilon} \, \mathrm{d}\mathfrak{m},$$
$$\int D^- f(\nabla g) \, \mathrm{d}\mathfrak{m} = \sup_{\varepsilon < 0} \int |Dg|_w \frac{|D(g + \varepsilon f)|_w - |Dg|_w}{\varepsilon} \, \mathrm{d}\mathfrak{m}.$$

By our assumptions we have that for any $f \in \text{Test}(\Omega)$ it holds $\int_{\text{supp}(f)} |D(g_n-g)|_w \, d\mathfrak{m} \to 0$, the sequence $\frac{|D(g_n+\varepsilon f)|_w - |Dg_n|_w}{\varepsilon}$ is uniformly bounded in $L^\infty(\Omega, \mathfrak{m}_{|\Omega})$ (by $\text{Lip}(f)$) and converges to $\frac{|D(g+\varepsilon f)|_w - |Dg|_w}{\varepsilon}$ in $L^1(\Omega, \mathfrak{m}_{|\Omega})$ for any $\varepsilon \neq 0$. It easily follows that

$$\overline{\lim_{n\to\infty}} \int_\Omega D^+ f(\nabla g_n) \, d\mathfrak{m} \leq \int_\Omega D^+ f(\nabla g) \, d\mathfrak{m},$$
$$\underline{\lim_{n\to\infty}} \int_\Omega D^- f(\nabla g_n) \, d\mathfrak{m} \geq \int_\Omega D^- f(\nabla g) \, d\mathfrak{m}. \tag{4.9}$$

If furthermore $f \in L^1(\Omega, |\mu|)$, (4.8) ensures that $f \in L^1(\Omega, |\mu_n|)$ for large n, so that from $\mu_n \in \mathbf{\Delta} g_{n|\Omega}$ we have

$$-\int D^+ f(\nabla g_n) \, d\mathfrak{m} \leq \int f \, d\mu_n \leq -\int D^- f(\nabla g_n) \, d\mathfrak{m}, \qquad \forall n \gg 0.$$

Using (4.9) and (4.8) we can pass to the limit in these inequalities and get the thesis. $\qquad\square$

The next statement shows that the definition of Laplacian is compatible with local minimizers of the Cheeger energy Ch_2.

Proposition 4.16 Let $(X, d, \mathfrak{m})$ be as in (2.7), $\Omega \subset X$ an open set and $g \in S^2_{\text{int}}(\Omega)$ with $\int_\Omega |Dg|^2_w \, d\mathfrak{m} < \infty$. Assume that it holds

$$\int_\Omega |Dg|^2_w \, d\mathfrak{m} \leq \int_\Omega |D(g+h)|^2_w \, d\mathfrak{m}, \qquad \forall h \in S^2(X, d, \mathfrak{m}) \text{ such that } \text{supp}(h) \subset \Omega. \tag{4.10}$$

Then $g \in D(\mathbf{\Delta}, \Omega)$ and $0 \in \mathbf{\Delta} g_{|\Omega}$

proof Notice that for $f \in \text{Test}(\Omega)$ we certainly have $\text{supp}(f) \subset \Omega$ and $f \in S^2(X, d, \mathfrak{m})$. Thus for $f \in \text{Test}(\Omega)$ and $\varepsilon \in \mathbb{R}$, (4.10) yields

$$\int_\Omega |D(g+\varepsilon f)|^2_w \, d\mathfrak{m} \geq \int_\Omega |Dg|^2_w \, d\mathfrak{m}.$$

Therefore

$$\int_\Omega \frac{|D(g+\varepsilon f)|^2_w - |Dg|^2_w}{2\varepsilon} \, d\mathfrak{m} \geq 0, \qquad \forall \varepsilon > 0,$$
$$\int_\Omega \frac{|D(g+\varepsilon f)|^2_w - |Dg|^2_w}{2\varepsilon} \, d\mathfrak{m} \leq 0, \qquad \forall \varepsilon < 0.$$

Letting $\varepsilon \downarrow 0$ and $\varepsilon \uparrow 0$ we get the thesis. $\qquad\square$

Another expected property of the Laplacian is the 'local-to-global' property. Shortly said, one would like a theorem of the following form: if for some measure μ on $\Omega_1 \cup \Omega_2$ it holds $\mathbf{\Delta} g_{|\Omega_i} \ni \mu_{|\Omega_i}$, $i = 1, 2$, then $\mathbf{\Delta} g_{|\Omega_1 \cup \Omega_2} \ni \mu$. Unfortunately, it is not clear - to us - whether this is true in the general case or not (actually, this is not so obvious even in normed spaces, when the norm is not strictly convex). Yet, at least for infinitesimally strictly convex and proper spaces the result holds, and follows by a standard partition of the unit argument.

Proposition 4.17 (Local-to-global) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) and assume also that (X, d) is proper. Let $p \in (1, \infty)$ and q the conjugate exponent, assume that $(X, \mathsf{d}, \mathfrak{m})$ is q-infinitesimally strictly convex, let $\Omega \subset X$ be an open set and $\{\Omega_i\}_{i \in I}$ a family of open sets such that $\Omega = \cup_i \Omega_i$. Let $g \in \mathrm{S}^p_{\mathrm{int}}(\Omega)$, with $g \in D(\boldsymbol{\Delta}, \Omega_i)$ for every $i \in I$ and let μ_i be the only element of $\boldsymbol{\Delta} g_{|\Omega_i}$. Then*

$$\mu_i|_{\Omega_i \cap \Omega_j} = \mu_j|_{\Omega_i \cap \Omega_j}, \qquad \forall i, j \in I, \tag{4.11}$$

$g \in D(\boldsymbol{\Delta}, \Omega)$ and the measure μ on Ω defined by

$$\mu|_{\Omega_i} := \mu_i, \qquad \forall i \in I, \tag{4.12}$$

is the only element of $\boldsymbol{\Delta} g_{|\Omega}$.

proof Since (X, d) is proper, for any $\tilde{\Omega} \subset X$ open and Radon measure ν on $\tilde{\Omega}$ and any $f \in \mathrm{Test}(\tilde{\Omega})$, it holds $f \in L^1(\tilde{\Omega}, |\nu|)$, because $\mathrm{supp}(f)$ is compact.

Let $i, j \in I$, $f \in \mathrm{Test}(\Omega_i \cap \Omega_j)$ and notice that by definition it holds

$$- \int_{\Omega_i \cap \Omega_j} f \, \mathrm{d}\mu_i = \int_{\Omega_i \cap \Omega_j} Df(\nabla g) \, \mathrm{d}\mathfrak{m} = - \int_{\Omega_i \cap \Omega_j} f \, \mathrm{d}\mu_j,$$

which yields (4.11). In particular, the measure μ is well defined by (4.12).

Fix $f \in \mathrm{Test}(\Omega)$. Since $\mathrm{supp}(f)$ is compact, there exists a finite set $I_f \subset I$ of indexes such that $\mathrm{supp}(f) \subset \cup_{i \in I_f} \Omega_i$. From the fact that (X, d) is proper it is easy to build a family $\{\chi_i\}_{i \in I_f}$ of Lipschitz functions such that $\sum_{i \in I_f} \chi_i \equiv 1$ on $\mathrm{supp}(f)$ and $\mathrm{supp}(\chi_i)$ is compact and contained in Ω_i for any $i \in I_f$. Hence $f\chi_i \in \mathrm{Test}(\Omega_i)$ for any $i \in I_f$ and taking into account the linearity of the differential expressed in Corollary 3.4, we have

$$\int Df(\nabla g) \, \mathrm{d}\mathfrak{m} = \int D\Big(\sum_{i \in I_f} \chi_i f\Big)(\nabla g) \, \mathrm{d}\mathfrak{m} = \sum_{i \in I_f} \int D(f\chi_i)(\nabla g) \, \mathrm{d}\mathfrak{m}$$

$$= - \sum_{i \in I_f} \int f\chi_i \, \mathrm{d}\mu_i = - \int f \, \mathrm{d}\Big(\sum_{i \in I_f} \chi_i \mu_i\Big),$$

which is the thesis. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We conclude the section discussing the effect of a change of the reference measure. As for the locality, the correct formula shows up under the assumption that the space is q-infinitesimally strictly convex.

Proposition 4.18 (Change of the reference measure) *Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric measure space and $V : X \to \mathbb{R}$ a locally Lipschitz function, which is Lipschitz when restricted to bounded sets. Define the measure $\mathfrak{m}' := e^{-V}\mathfrak{m}$ and let $\boldsymbol{\Delta}'$ be the Laplacian in $(X, \mathsf{d}, \mathfrak{m}')$, where in Definition 4.4 we replace $\mathfrak{m}$ with $\mathfrak{m}'$. Let $\Omega \subset X$ be an open set, $g \in D(\boldsymbol{\Delta}) \cap \mathrm{S}^p_{\mathrm{int}}(\Omega)$ for some $p \in (1, \infty)$ and assume that $(X, \mathsf{d}, \mathfrak{m})$ is q-infinitesimally strictly convex, q being the conjugate exponent of p. Then $g \in D(\boldsymbol{\Delta}', \Omega)$ and the measure*

$$\mu' := e^{-V}\mu - DV(\nabla g)e^{-V}\mathfrak{m},$$

is the only element in $\boldsymbol{\Delta}' g_{|\Omega}$, where here μ is the only element of $\boldsymbol{\Delta} g_{|\Omega}$ (Proposition 4.7).

proof Let Test$'(\Omega)$ be defined as Test(Ω) with the measure $\mathfrak{m}'$ replacing $\mathfrak{m}$ (the role of the measure is in the definition of Int(Ω)).

Since e^{-V} and $|DV|_w$ are locally bounded, μ' is a locally finite measure and the statement makes sense. For $f \in \text{Test}'(\Omega) \cap L^1(\Omega, |\mu'|)$, the fact that supp$(f)$ is bounded yields that the restriction of V to supp(f) is Lipschitz and bounded. It easily follows that $fe^{-V} \in \text{Test}(\Omega) \cap L^1(\Omega, |\mu|)$, so that from the chain rule (3.34) and the Leibniz rule (3.36) we get

$$\int f \, \mathrm{d}\mu' = \int f e^{-V} \, \mathrm{d}\mu - \int f DV(\nabla g)e^{-V} \, \mathrm{d}\mathfrak{m}$$

$$= -\int D(fe^{-V})(\nabla g) - f D(e^{-V})(\nabla g) \, \mathrm{d}\mathfrak{m} = -\int e^{-V} Df(\nabla g) \, \mathrm{d}\mathfrak{m},$$

which is the thesis. $\qquad\qquad\square$

4.3 The linear case

In this section we introduce a sufficient, and in most cases also necessary, condition in order for the Laplacian to be linear, and analyze its properties also in connection with Ricci curvature bounds.

Definition 4.19 (Infinitesimally Hilbertian spaces) *We say that $(X, \mathsf{d}, \mathfrak{m})$ is infinitesimally Hilbertian provided it is as in (2.7) and the seminorm $\| \cdot \|_{\mathrm{S}^2(X,\mathsf{d},\mathfrak{m})}$ on $\mathrm{S}^2(X, \mathsf{d}, \mathfrak{m})$ satisfies the parallelogram rule.*

In particular, infinitesimally Hilbertian spaces are 2-infinitesimally strictly convex, hence for any $f, g \in \mathrm{S}^2_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ the object $Df(\nabla g)$ is well defined $\mathfrak{m}$-a.e., and for $g \in D(\boldsymbol{\Delta}) \cap \mathrm{S}^2(X, \mathsf{d}, \mathfrak{m})$ the set $\boldsymbol{\Delta}g$ contains only one element, which - abusing a bit the notation - we will denote by $\boldsymbol{\Delta}g$.

Remark 4.20 (About stability) In the class of compact and normalized (i.e. $\mathfrak{m} \in \mathscr{P}(X)$) metric measure spaces, the class of infinitesimally Hilbertian spaces is definitively *not* closed w.r.t. measured Gromov-Hausdorff convergence. Actually it is dense, because any compact normalized space is the mGH limit of a sequence of discrete spaces, and on discrete spaces Sobolev analysis trivializes, i.e. $|Df|_w \equiv 0$ for any f. The fact that infinitesimal Hilbertianity is not a closed condition w.r.t. mGH convergence is not at all surprising, because the former is a first order condition on the space, while the latter is a zeroth order convergence.

It is worth underlying, however, that the class of spaces which are both infinitesimally Hilbertian and $\mathsf{CD}(K, N)$ for some $K \in \mathbb{R}$, $N \in (1, \infty]$, *is* closed w.r.t. mGH convergence. This can be seen - at the very heuristic level - as the stability of a first order condition w.r.t. a zeroth order convergence under a uniform second order bound (the curvature bound). The proof of this stability passes through the identification of the gradient flow of Ch_2 in L^2 with the one of the relative entropy in the Wasserstein space (proved in [4]) and the stability of gradient flows of a sequence of K-geodesically convex functionals which Γ-converges (proved in [19]) along the following lines:

   i) $(X, \mathsf{d}, \mathfrak{m})$ is infinitesimally Hilbertian if and only if Ch_2 is a Dirichlet form (obvious, see Proposition 4.22 below and the discussion thereafter for the details)

ii) Ch_2 is a Dirichlet form if and only if its gradient flow linearly depends on the initial datum (obvious)

iii) The gradient flow of Ch_2 in L^2 coincides with the gradient flow of the relative entropy in the Wasserstein space (Theorem 9.3 of [4])

iv) If $(X_n, \mathsf{d}_n, \mathfrak{m}_n)$ converges to $(X, \mathsf{d}, \mathfrak{m})$ in the mGH sense, then the relative entropies Γ-converge (this is the heart of the proof of stability of the $\mathsf{CD}(K, \infty)$ condition proved, up to minor technical differences, by Lott-Villani in [30] and Sturm in [46])

v) If a sequence of K-geodesically convex functionals Γ-converges to a limit functional, then gradient flows of the approximating sequence converge to gradient flows of the limit functional (Theorem 21 of [19])

So that in the end we know that on the limit space the gradient flow of the relative entropy linearly depends on the initial datum and we conclude using again $(iii), (ii), (i)$.

We remark that the combination of (iv) and (v) above is sufficient to prove that *a compact smooth Finsler manifold is the mGH-limit of a sequence of Riemannian manifolds with Ricci curvature uniformly bounded from below if and only if it is Riemannian itself.* To prove it just recall that the heat flow on a Finsler manifold is linear if and only if it is Riemannian and that $(iv), (v)$ yield that the heat flow passes to the limit under a uniform lower Ricci bound.

The fact that Finsler manifolds cannot arise as limits of Riemannian manifolds with a uniform Ricci bound from below is certainly not a new result, at least under the additional assumption that the dimension of the approximating sequence is uniformly bounded: in [14], [15], [16] Cheeger and Colding proved the same thing (and much more). The proof we propose has the advantage of being simple and of avoiding the use of the almost splitting theorem, although clearly to (hope to) recover the results in [14], [15], [16] in the non-smooth context a lot of work has still to be done.

It is natural to think that the right condition to add to $\mathsf{CD}(K, N)$ spaces in order to enforce a Riemannian-like behavior on small scales and to rule out Finsler geometries, is infinitesimal Hilbertianity. Yet, from the technical point of view it is not so clear whether this is sufficient or not to derive the expected properties of the heat flow (e.g., K-contractivity of W_2 along two flows): the first attempt in this direction has been done in [6], where infinitesimal Hilbertianity was enforced with a notion of K-geodesic convexity of the entropy stronger than the standard one (see Definition 4.30 below and Theorem 4.31 for a - non-complete - overview of the results of [6]). Recent results by Rajala (see [40]) suggest that infinitesimal Hilbertianity plus $\mathsf{CD}(K, \infty)$ are actually sufficient to recover all the properties obtained in [6], without any strengthening of the curvature condition, but the situation, as of today, is not yet completely clear[2]. ∎

We start proving that infinitesimal Hilbertianity can be checked by looking at the symmetry of the map $\mathrm{S}^2 \ni f, g \mapsto Df(\nabla g)$. The proof follows the same lines of Section 4.3 of [6].

Proposition 4.21 *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be as in* (2.7). *Then it is an infinitesimally Hilbertian space if and only if it is 2-infinitesimally strictly convex and for every* $f, g \in \mathrm{S}^2_{\mathrm{loc}}(X, \mathsf{d}, \mathfrak{m})$ *it holds*

$$Df(\nabla g) = Dg(\nabla f), \qquad \mathfrak{m} - a.e.. \tag{4.13}$$

---

[2]this turned out to be true: in the recent paper [2] it has been shown that infinitesimal Hilbertianity plus $\mathsf{CD}(K, \infty)$ is sufficient to deduce all the properties of the heat flow stated in [6]

proof Assume at first that the space is 2-infinitesimally strictly convex and that (4.13) holds. Fix $f, g \in S^2(X, d, m)$ and notice that

$$\|g + f\|_{S^2}^2 - \|g\|_{S^2}^2 = \||D(g+f)|_w\|_{L^2}^2 - \||Dg|_w\|_{L^2}^2 = \int_0^1 \frac{d}{dt} \||D(g+tf)|_w\|_{L^2}^2 \, dt$$

$$= 2 \iint_0^1 Df(\nabla(g+tf)) \, dt \, dm \overset{(4.13)}{=} 2 \iint_0^1 D(g+tf)(\nabla f) \, dt \, dm$$

$$= 2 \int Dg(\nabla f) \, dm + \|f\|_{S^2}^2,$$

having used the linearity of the differential (Corollary 3.4) in the last step. Exchange f with $-f$ and add up to conclude.

We pass to the converse implication. With a cut-off and truncation argument we can reduce to the case $f, g \in S^2(X, d, m) \cap L^\infty(X, m)$. The fact that $\|\cdot\|_{S^2}$ satisfies the parallelogram rule easily yields

$$\frac{\|g + \varepsilon f\|_{S^2}^2 - \|g\|_{S^2}^2}{\varepsilon} = \frac{\|f + \varepsilon g\|_{S^2}^2 - \|f\|_{S^2}^2}{\varepsilon} + O(\varepsilon), \qquad \forall f, g \in S^2(X, d, m),$$

so that by the very definition of $Df(\nabla g)$ and (3.6) we get

$$\int Df(\nabla g) \, dm = \int Dg(\nabla f) \, dm, \qquad \forall f, g \in S^2(X, d, m). \tag{4.14}$$

We want to pass from this integrated equality, to the pointwise statement (4.13). The thesis will follow if we prove that

$$S^2(X, d, m) \cap L^\infty(X, m) \ni f \qquad \mapsto \qquad \int h|Df|_w^2 \, dm \ \in \mathbb{R}$$

is a quadratic form for any $h \in S^2(X, d, m) \cap L^\infty(X, m)$. Fix such h and use (3.6), the Leibniz rule (3.36) and the chain rule to get

$$\int h|Df|_w^2 \, dm = \int h Df(\nabla f) \, dm = \int D(hf)(\nabla f) - f Dh(\nabla f) \, dm$$

$$= \int D(hf)(\nabla f) \, dm - \frac{1}{2} \int Dh(\nabla(f^2)) \, dm$$

$$= \int D(hf)(\nabla f) \, dm - \frac{1}{2} \int D(f^2)(\nabla h) \, dm,$$

having used (4.14) in the last step. Now recall from Corollary 3.4 that the map $S^2(X, d, m) \ni g \mapsto \int Dg(\nabla h) \, dm$ is linear, hence the map $S^2(X, d, m) \cap L^\infty(X, m) \ni f \mapsto \int D(f^2)(\nabla h) \, dm$ is a quadratic form. Similarly, from the fact that both the maps $S^2(X, d, m) \cap L^\infty(X, m) \ni g \mapsto \int D(hg)(\nabla f) \, dm$ and $S^2(X, d, m) \ni g \mapsto \int D(hf)(\nabla g) \, dm = \int Dg(\nabla(hf)) \, dm$ are linear, we get that $S^2(X, d, m) \cap L^\infty(X, m) \ni f \mapsto \int D(hf)(\nabla f) \, dm$ is a quadratic form. $\qquad \square$

On infinitesimally Hilbertian spaces we will denote $Df(\nabla g)$ by $\nabla f \cdot \nabla g$, in order to highlight the symmetry of this object. A direct consequence of (4.13) and the linearity of the differential expressed in Corollary 3.4, is the bilinearity of $\nabla f \cdot \nabla g$, i.e.

$$\nabla(\alpha_1 f_1 + \alpha_2 f_2) \cdot \nabla g = \alpha_1 \nabla f_1 \cdot \nabla g + \alpha_2 \nabla f_2 \cdot \nabla g, \qquad \forall f_1, f_2, g \in S^2_{loc}, \ \alpha_1, \alpha_2 \in \mathbb{R},$$

$$\nabla f \cdot \nabla(\beta_1 g_1 + \beta_2 g_2) = \beta_1 \nabla f \cdot \nabla g_1 + \beta_2 \nabla f \cdot \nabla g_2, \qquad \forall f, g_1, g_2 \in S^2_{loc}, \ \beta_1, \beta_2 \in \mathbb{R}, \tag{4.15}$$

and from (3.36) we also get

$$\nabla(f_1 f_2) \cdot \nabla g = f_1 \nabla f_2 \cdot \nabla g + f_2 \nabla f_1 \cdot \nabla g, \qquad \forall f_1, f_2 \in \mathrm{S}^2_{\mathrm{loc}} \cap L^\infty_{\mathrm{loc}}, \ g \in \mathrm{S}^2_{\mathrm{loc}},$$
$$\nabla f \cdot \nabla(g_1 g_2) = g_1 \nabla f \cdot \nabla g_2 + g_2 \nabla f \cdot \nabla g_1, \qquad \forall f \in \mathrm{S}^2_{\mathrm{loc}}, \ g_1, g_2 \in \mathrm{S}^2_{\mathrm{loc}} \cap L^\infty_{\mathrm{loc}}, \tag{4.16}$$

all the equalities being intended $\mathfrak{m}$-a.e..

The following proposition collects some different characterizations of infinitesimally Hilbertian spaces.

Proposition 4.22 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7). Then the following are equivalent.*

i) *$(X, \mathsf{d}, \mathfrak{m})$ is infinitesimally Hilbertian.*

ii) *For any $\Omega \subset X$ open with $\mathfrak{m}(\partial\Omega) = 0$ the space $(\overline{\Omega}, \mathsf{d}, \mathfrak{m}_{|\Omega})$ is infinitesimally Hilbertian.*

iii) *$W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is an Hilbert space.*

iv) *The Cheeger energy $\mathrm{Ch}_2 : L^2 \to [0, \infty]$ is a quadratic form, i.e.*

$$\mathrm{Ch}_2(f + g) + \mathrm{Ch}_2(f - g) = 2\Big(\mathrm{Ch}_2(f) + \mathrm{Ch}_2(g)\Big), \qquad \forall f, g \in L^2(X, \mathfrak{m}). \tag{4.17}$$

proof
(i) $\Rightarrow$ (ii) Follows from Proposition 2.6 and the equivalence stated in Proposition 4.21.
(ii) $\Rightarrow$ (i) Obvious: just take $\Omega := X$.
(i) $\Rightarrow$ (iii) Obviously follows from the definition of the $W^{1,2}$ norm.
(iii) $\Rightarrow$ (i) We know that $W^{1,2}(X, \mathsf{d}, \mathfrak{m}) \ni f \mapsto \||Df|_w\|^2_{L^2}$ is a quadratic form, and we want to prove that the same its true for its extension to $\mathrm{S}^2(X, \mathsf{d}, \mathfrak{m})$. Arguing exactly as in the proof of Proposition 4.21 above we have that

$$W^{1,2}(X, \mathsf{d}, \mathfrak{m}) \ni f \qquad \mapsto \qquad |Df|^2_w \in L^1(X, \mathfrak{m}), \qquad \text{is a quadratic form.} \tag{4.18}$$

Now use the local finiteness of $\mathfrak{m}$ and the Lindelöf property of (X, d) to build an increasing sequence (K_n) of compact sets such that $\mathfrak{m}(X \setminus \cup_n K_n) = 0$ and an increasing sequence (χ_n) of Lipschitz bounded functions such that $\mathfrak{m}(\mathrm{supp}(\chi_n)) < \infty$ and $\chi_n \equiv 1$ on K_n for every $n \in \mathbb{N}$. Then for every $f \in \mathrm{S}^2(X, \mathsf{d}, \mathfrak{m}) \cap L^\infty(X, \mathfrak{m})$ and $n \in \mathbb{N}$ it holds $f\chi_n \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$. Since $|Df|_w = |D(f\chi_n)|_w$ $\mathfrak{m}$-a.e. on K_n, from (4.18) and letting $n \to \infty$ we get that $\mathrm{S}^2(X, \mathsf{d}, \mathfrak{m}) \cap L^\infty(X, \mathfrak{m}) \ni f \mapsto |Df|^2_w \in L^1(X, \mathfrak{m})$ is a quadratic form as well. With a truncation argument we get that also $\mathrm{S}^2(X, \mathsf{d}, \mathfrak{m}) \ni f \mapsto |Df|^2_w \in L^1$ is a quadratic form. Integrate to conclude.
(iii) $\Leftrightarrow$ (iv) The formula $\|f\|^2_{W^{1,2}} = \|f\|^2_{L^2} + 2\mathrm{Ch}_2(f)$ shows that Ch_2 satisfies the parallelogram rule if and only if so does the $W^{1,2}$ norm. $\qquad\square$

A property of particular importance of infinitesimally Hilbertian spaces is that the Laplacian is linear, as explained in the following proposition. We will denote by $D_{\mathrm{ifm}}(\boldsymbol{\Delta}, \Omega) \subset D(\boldsymbol{\Delta}, \Omega)$ the set of those g's whose Laplacian has 'internally finite mass', i.e. for any $\mu \in \boldsymbol{\Delta} g_{|\Omega}$ it holds $|\mu|(\Omega') < \infty$ for any $\Omega' \in \mathrm{Int}(\Omega)$. If (X, d) is proper, then $D_{\mathrm{ifm}}(\boldsymbol{\Delta}, \Omega) = D(\boldsymbol{\Delta}, \Omega)$.

Proposition 4.23 (Linearity of the Laplacian) *Let $(X, \mathsf{d}, \mathfrak{m})$ be an infinitesimally Hilbertian space. Then for any $\Omega \subset X$ open, the set $D_{\mathrm{ifm}}(\boldsymbol{\Delta}, \Omega) \cap \mathrm{S}^2_{\mathrm{int}}(\Omega)$ is a vector space and for $g \in D_{\mathrm{ifm}}(\boldsymbol{\Delta}, \Omega) \cap \mathrm{S}^2_{\mathrm{int}}(\Omega)$, $\boldsymbol{\Delta} g_{|\Omega}$ is single valued and linearly depends on g.*

proof Fix $\Omega \subset X$ open. Being $(X, \mathsf{d}, \mathfrak{m})$ 2-infinitesimally strictly convex, by Proposition 4.7 we know that $\boldsymbol{\Delta} g_{|\Omega}$ is single valued for any $g \in \mathrm{S}^2_{\mathrm{int}}(\Omega) \cap D(\boldsymbol{\Delta}, \Omega)$. With a cut-off argument we deduce that (4.15) is satisfied also for functions in $\mathrm{S}^2_{\mathrm{int}}(\Omega)$. Pick $g_1, g_2 \in D_{\mathrm{ifm}}(\boldsymbol{\Delta}, \Omega) \cap \mathrm{S}^2_{\mathrm{int}}(\Omega)$, $\beta_1, \beta_2 \in \mathbb{R}$. Observe that $|\boldsymbol{\Delta} g_i|(\mathrm{supp}(f)) < \infty$ for any $f \in \mathrm{Test}(\Omega)$, so that in particular $\mathrm{Test}(\Omega) \subset L^1(\Omega, |\boldsymbol{\Delta} g_{|\Omega}|)$, $i = 1, 2$. Also, it holds $\mathrm{Test}(\Omega) \subset \mathrm{S}^2(X, \mathsf{d}, \mathfrak{m})$. Let $f \in \mathrm{Test}(\Omega)$ and conclude with

$$\int_\Omega f \, \mathrm{d}(\beta_1 \boldsymbol{\Delta} g_1{}_{|\Omega} + \beta_2 \boldsymbol{\Delta} g_2{}_{|\Omega}) = \beta_1 \int_\Omega f \, \mathrm{d}\boldsymbol{\Delta} g_1{}_{|\Omega} + \beta_2 \int_\Omega f \, \mathrm{d}\boldsymbol{\Delta} g_2{}_{|\Omega}$$
$$= -\beta_1 \int_\Omega \nabla f \cdot \nabla g_1 \, \mathrm{d}\mathfrak{m} - \beta_2 \int_\Omega \nabla f \cdot \nabla g_2 \, \mathrm{d}\mathfrak{m}$$
$$= - \int_\Omega \nabla f \cdot \nabla (\beta_1 g_1 + \beta_2 g_2) \, \mathrm{d}\mathfrak{m}.$$

$\square$

The fact that on infinitesimally Hilbertian spaces the Cheeger energy Ch_2 satisfies (4.17) yields that the formula

$$\mathcal{E}(f, g) := \mathrm{Ch}_2(f + g) - \mathrm{Ch}_2(f) - \mathrm{Ch}_2(g),$$

defines a symmetric bilinear form on $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$. The chain rule (2.16) gives that the form is Markovian and the semicontinuity of Ch_2 means that the form is closed. Hence, $\mathcal{E}$ is (more precisely, can be extended to) a Dirichlet form on $L^2(X, \mathfrak{m})$. As a consequence of the locality principle (2.15), $\mathcal{E}$ is strongly local. The generator of $\mathcal{E}$ will be denoted by Δ, so that $\Delta : D(\Delta) \subset L^2(X, \mathfrak{m}) \to L^2(X, \mathfrak{m})$ and Δ and its domain $D(\Delta)$ are defined by

$$g \in D(\Delta), \ h = \Delta g \qquad \text{if and only if} \qquad \mathcal{E}(f, g) = -\int fh \, \mathrm{d}\mathfrak{m}, \qquad \forall f \in D(\mathcal{E}). \quad (4.19)$$

Notice that $D(\mathcal{E}) = D(\mathrm{Ch}_2) = W^{1,2}(X, \mathsf{d}, \mathfrak{m})$.

In the following proposition we compare the various definitions of 'the Laplacian of g exists and is in L^2'. Notice that the results are sharper than those of Proposition 4.9.

Proposition 4.24 (Compatibility) *Let $(X, \mathsf{d}, \mathfrak{m})$ be an infinitesimally Hilbertian space, $g \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ and $h \in L^2(X, \mathfrak{m})$. Consider the following statements:*

i) $g \in D(\boldsymbol{\Delta})$ and $\boldsymbol{\Delta} g = h\mathfrak{m}$,

ii) $h \in -\partial^- \mathrm{Ch}_2(g)$,

iii) $g \in D(\Delta)$ and $\Delta g = h$.

Then

$$(ii) \quad \Leftrightarrow \quad (iii) \quad \Rightarrow \quad (i),$$

and if $\mathfrak{m}$ is finite on bounded sets it also holds

$$(i) \quad \Rightarrow \quad (ii), (iii).$$

proof
(iii) $\Rightarrow$ **(ii)**. We need to prove that for any $f \in L^2(X, \mathfrak{m})$ it holds

$$\mathrm{Ch}_2(g) - \int (f - g)h \, \mathrm{d}\mathfrak{m} \leq \mathrm{Ch}_2(f). \quad (4.20)$$

If $f \notin D(\mathrm{Ch}_2) = D(\mathcal{E})$ there is nothing to prove. Otherwise $f - g$ is in $D(\mathcal{E})$ and we can use the definition of Δ to get

$$-\int (f-g)h\,d\mathfrak{m} = \mathcal{E}(f-g,g) = \mathcal{E}(f,g) - \mathcal{E}(g,g) \leq \frac{1}{2}\mathcal{E}(f,f) - \frac{1}{2}\mathcal{E}(g,g),$$

which is (4.20).

$(\mathbf{ii}) \Rightarrow (\mathbf{iii})$. Pick $f \in D(\mathcal{E})$ and notice that by definition of $\partial^- \mathrm{Ch}_2(g)$ it holds

$$\mathrm{Ch}_2(g) - \int \varepsilon f h\,d\mathfrak{m} \leq \mathrm{Ch}_2(g + \varepsilon f), \qquad \forall \varepsilon \in \mathbb{R}.$$

Conclude observing that $\mathrm{Ch}_2(g + \varepsilon f) = \mathrm{Ch}_2(g) + \varepsilon^2 \mathrm{Ch}_2(f) + \mathcal{E}(f,g)$, dividing by ε and letting $\varepsilon \downarrow 0$ and $\varepsilon \uparrow 0$.

$(\mathbf{ii}) \Rightarrow (\mathbf{i})$. This is a particular case of Proposition 4.9.

$(\mathbf{i}) \Rightarrow (\mathbf{iii})$. Here we assume that $\mathfrak{m}$ is finite on bounded sets. We know that

$$-\int \nabla f \cdot \nabla g\,d\mathfrak{m} = \int f h\,d\mathfrak{m}, \qquad \forall f \in \mathrm{Test}(X) \cap L^1(X, |h|\mathfrak{m})$$

and we want to conclude that the same is true for any $f \in W^{1,2}(X,\mathsf{d},\mathfrak{m})$. Pick $f \in W^{1,2}(X,\mathsf{d},\mathfrak{m})$ and assume for the moment that $\mathrm{supp}(f)$ is bounded. Let χ be a Lipschitz bounded function with bounded support and identically 1 on $\mathrm{supp}(f)$. Also, let (f_n) be a sequence of Lipschitz functions converging to f in $W^{1,2}(X,\mathsf{d},\mathfrak{m})$ (Corollary 2.9). Then $f_n\chi \to f$ in $W^{1,2}(X,\mathsf{d},\mathfrak{m})$ and $f_n\chi \in \mathrm{Test}(X) \cap L^1(|h|\mathfrak{m})$ for any $n \in \mathbb{N}$. Thus passing to the limit in $-\int \nabla(f_n\chi) \cdot \nabla g\,d\mathfrak{m} = \int f_n\chi h\,d\mathfrak{m}$ we get $-\int \nabla f \cdot \nabla g\,d\mathfrak{m} = \int f h\,d\mathfrak{m}$ for any $f \in W^{1,2}(X,\mathsf{d},\mathfrak{m})$ with bounded support.

To achieve the general case, let (χ_n) be an increasing sequence of non-negative, 1-Lipschitz functions with bounded support and such that $\chi_n \equiv 1$ on $B_n(x_0)$, where $x_0 \in X$ is some fixed point, for any $n \in \mathbb{N}$. Then fix an arbitrary $f \in W^{1,2}(X,\mathsf{d},\mathfrak{m})$, notice that $f\chi_n \in W^{1,2}(X,\mathsf{d},\mathfrak{m})$ as well and has bounded support. We claim that $f\chi_n \to f$ in $W^{1,2}(X,\mathsf{d},\mathfrak{m})$. Indeed, by the dominate convergence theorem we immediately get that $\|f - f\chi_n\|_{L^2} \to 0$ as $n \to \infty$. Also, we have $|D(f - f\chi_n)|_w \to 0$ $\mathfrak{m}$-a.e. as $n \to \infty$ and $|D(f - f\chi_n)|_w \leq |Df|_w|1 - \chi_n| + |f|$ $\mathfrak{m}$-a.e., so that again by dominate convergence we get $|D(f - f\chi_n)|_w \to 0$ in $L^2(X,\mathfrak{m})$ as $n \to \infty$. Conclude by letting $n \to \infty$ in

$$-\int \nabla(f\chi_n) \cdot \nabla g\,d\mathfrak{m} = \int f\chi_n h\,d\mathfrak{m}.$$

$\square$

Remark 4.25 (Compatibility with the theory of Alexandrov spaces) In [27], using key results of [37] and [38], it has been proved that finite dimensional Alexandrov spaces with curvature bounded from below are infinitesimally Hilbertian.

On these spaces there are four - compatible - ways of defining a Laplacian:

i) To use the fact that $W^{1,2}$ is Hilbert to build a natural Dirichlet form and then considering its generator.

ii) To use the DC structure of the space to define the Laplacian of DC^1 functions on $X \setminus S_\delta$ by direct calculus. Recall that DC functions are functions which are locally the Difference of Concave functions, DC^1 functions are also C^1, that a DC chart is a chart made of maps with DC components and that Perelman showed in [38] that 'most' of an Alexandrov space (i.e. all of it with the exception of a small singular set S_δ) can be covered by DC charts.

iii) To use the Riemannian structure of the space to give a meaning to the integration by parts formula for functions which are not necessarly DC.

iv) To equip these spaces with their natural Hausdorff measure $\mathcal{H}^n$ and use the definition given here. Notice that the resultant metric measure space is always doubling and supports a 1-1 Poincaré inequality, hence in this case there is no ambiguity due to the choice of the Sobolev exponent p in the definition of distributional Laplacian (see also Remarks 2.5 and 4.5).

For the first two approaches see [27]. Due to Proposition 4.24, the approach (i) is clearly in line with (iv). The same is true for the approach (ii), because (iv) clearly reduces to the standard definition of distributional Laplacian on Riemannian manifolds equipped with a metric tensor locally uniformly elliptic, which is the case for DC manifolds.

The approach (iii) was introduced by Petrunin ([39]) for the case of DC functions. To define the Laplacian of these functions, he used the integration by parts rule: for g which is DC on some open domain Ω he proposed

$$\int_\Omega f \, \mathrm{d}\boldsymbol{\Delta} g := - \int_\Omega \langle \operatorname{grad} f, \operatorname{grad} g \rangle \, \mathrm{d}\mathcal{H}^n, \qquad \forall f : \overline{\Omega} \to \mathbb{R} \text{ Lipschitz with } f_{|\partial\Omega} = 0, \quad (4.21)$$

where n is the dimension and $\mathcal{H}^n$ the corresponding Hausdorff measure and the object $\langle \operatorname{grad} f, \operatorname{grad} g \rangle$ is defined $\mathcal{H}^n$-a.e. via the use of DC charts. He used this definition of Laplacian to prove that for DC functions the measure $\boldsymbol{\Delta}$ actually exists and that it is non-negative for convex functions.

To see that (4.21) defines the same measure as the one given in (4.4), start observing that the requirements '$f_{|\partial\Omega} = 0$' and 'supp$(f) \subset\subset \Omega$' are easily seen to be interchangeable in this setting, so that it is sufficient to prove that $\mathcal{H}^n$-a.e. it holds $\langle \operatorname{grad} f, \operatorname{grad} g \rangle = \nabla f \cdot \nabla g$. This follows from the fact that $\mathcal{H}^n$-almost all the space can be covered by DC charts, and that on these charts the two objects are clearly the same. $\blacksquare$

In the next two lemmas we prove that the higher regularity of g and $\boldsymbol{\Delta} g$ is, the wider is the class of functions for which the integration by part rules holds.

Lemma 4.26 *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be a proper infinitesimally Hilbertian space,* $\Omega \subset X$ *an open set and* $g \in D(\boldsymbol{\Delta}, \Omega) \cap \mathrm{S}^2_{\mathrm{int}}(\Omega)$. *Then for every* $\psi \in \mathrm{S}^2(X, \mathsf{d}, \mathfrak{m}) \cap C_c(X)$ *with support contained in* Ω *it holds*

$$- \int_\Omega \nabla\psi \cdot \nabla g \, \mathrm{d}\mathfrak{m} = \int_\Omega \psi \, \mathrm{d}\boldsymbol{\Delta} g_{|\Omega}.$$

proof Since (X, d) is proper, supp(ψ) is compact and therefore $|\boldsymbol{\Delta} g_{|\Omega}|(\operatorname{supp}(\psi)) < \infty$. Similarly, we have $\mathfrak{m}(\operatorname{supp}(\psi)) < \infty$ and thus $\psi \in L^2(X, \mathfrak{m})$. Hence $\psi \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ and from

Corollary 2.9 we know there exists a sequence $(\psi_n) \subset W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ of Lipschitz functions converging to ψ in $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$. Define $\psi^{t,+}, \psi^{t,-} : X \to \mathbb{R}$ by

$$\psi^{t,+}(x) := \inf_y \psi(y) + \frac{\mathsf{d}^2(x,y)}{2t}, \qquad \psi^{t,-}(x) := \sup_y \psi(y) - \frac{\mathsf{d}^2(x,y)}{2t}.$$

It is not hard to show that since $\psi \in C_b(X)$, the functions $\psi^{t,+}, \psi^{t,-}$ are Lipschitz, equibounded and it holds $\psi^{t,+}(x) \uparrow \psi(x)$, $\psi^{t,-}(x) \downarrow \psi(x)$ as $t \downarrow 0$ for any $x \in X$ (see for instance Chapter 3 of [4]).

Put $\psi_{n,t} := \min\{\max\{\psi_n, \psi^{t,+}\}, \psi^{t,-}\}$ and observe that $\psi_{n,t}$ is Lipschitz for any $n \in \mathbb{N}$, $t > 0$. Let $\chi \in \text{Test}(\Omega)$ be identically 1 on $\text{supp}(\psi)$, and consider the functions $\chi\psi_{n,t} \in \text{Test}(\Omega)$.

Since $\psi^{t,+}, \psi^{t,-}, \chi$ are Lipschitz, an application of the dominate convergence theorem ensures that for any $t > 0$ the sequence $n \mapsto \psi_{n,t}$ converges to ψ in energy $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ as $n \to \infty$. Thus this convergence is also w.r.t. the $W^{1,2}$ norm and from $\text{supp}(\chi\psi_{n,t}) \subset \text{supp}(\chi)$, $\mathfrak{m}(\text{supp}(\chi)) < \infty$ we get

$$\lim_{n\to\infty} \int_\Omega \nabla(\chi\psi_{n,t}) \cdot \nabla g \, \mathrm{d}\mathfrak{m} = \int_\Omega \nabla\psi \cdot \nabla g \, \mathrm{d}\mathfrak{m}, \qquad \forall t > 0. \tag{4.22}$$

By construction $\psi_{n,t}$ is uniformly bounded in n, t and pointwise converges to ψ as $t \to 0$ uniformly on n. Taking into account that $\text{supp}(\chi\psi_{n,t}) \subset \text{supp}(\chi)$, $|\boldsymbol{\Delta}g_{|\Omega}|(\text{supp}(\chi)) < \infty$ and the dominate convergence theorem we get

$$\lim_{t\downarrow 0} \int_\Omega \chi\psi_{n,t} \, \mathrm{d}\boldsymbol{\Delta}g_{|\Omega} = \int \psi \, \mathrm{d}\boldsymbol{\Delta}g_{|\Omega}, \qquad \text{uniformly on } n. \tag{4.23}$$

Since $\chi\psi_{n,t} \in \text{Test}(\Omega)$, (4.22) and (4.23) together with a diagonalization argument give the thesis. $\qquad\square$

Lemma 4.27 *Let $(X, \mathsf{d}, \mathfrak{m})$ be a proper infinitesimally Hilbertian space, $\Omega \subset X$ an open set and $g \in D(\boldsymbol{\Delta}, \Omega) \cap \mathrm{S}^2_{\mathrm{int}}(\Omega)$ with $\boldsymbol{\Delta}g_{|\Omega} \ll \mathfrak{m}$, $\boldsymbol{\Delta}g_{|\Omega} = h\mathfrak{m}$ and $h \in L^2_{\mathrm{loc}}(\Omega, \mathfrak{m}_{|\Omega})$. Then for every $\psi \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ with bounded support contained in Ω it holds*

$$-\int_\Omega \nabla\psi \cdot \nabla g \, \mathrm{d}\mathfrak{m} = \int_\Omega \psi h \, \mathrm{d}\mathfrak{m}.$$

proof By Corollary 2.9 we know there exists a sequence $(\psi_n) \subset W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ of Lipschitz functions converging to ψ in $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$. Let $\chi \in \text{Test}(\Omega)$ be identically 1 on $\text{supp}(\psi)$ and notice that $\chi\psi_n \in \text{Test}(\Omega) \cap L^1(\Omega, |h|\mathfrak{m})$ and $\chi\psi_n \to \psi$ in $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$. Hence we can pass to the limit in

$$-\int_\Omega \nabla(\chi\psi_n) \cdot \nabla g \, \mathrm{d}\mathfrak{m} = \int_\Omega \chi\psi_n h \, \mathrm{d}\mathfrak{m},$$

and get the thesis. $\qquad\square$

Proposition 4.28 (Chain rule) *Let $(X, \mathsf{d}, \mathfrak{m})$ be an infinitesimally Hilbertian metric measure space, $\Omega \subset X$ an open set, $g \in D(\boldsymbol{\Delta}, \Omega) \cap \mathrm{S}^2_{\mathrm{int}}(\Omega)$, $I \subset \mathbb{R}$ an open set such that $\mathfrak{m}(g^{-1}(\mathbb{R} \setminus I)) = 0$ and $\varphi \in C^{1,1}_{\mathrm{loc}}(I)$. Then the following holds.*

i) *Assume that* $g_{|\Omega'}$ *is Lipschitz for every* $\Omega' \in \mathrm{Int}(\Omega)$. *Then*

$$\varphi \circ g \in D(\mathbf{\Delta}, \Omega), \qquad and \qquad \mathbf{\Delta}(\varphi \circ g)_{|\Omega} = \varphi' \circ g\, \mathbf{\Delta}g_{|\Omega} + \varphi'' \circ g\, |Dg|_w^2 \mathfrak{m}_{|\Omega}. \quad (4.24)$$

ii) *Assume that* (X, d) *is proper and* $g \in C(\Omega)$. *Then* (4.24) *holds.*

iii) *Assume that* (X, d) *is proper,* $g \in L^2_{\mathrm{loc}}(X, \mathfrak{m})$ *and* $\mathbf{\Delta}g_{|\Omega} \ll \mathfrak{m}$ *with* $\dfrac{\mathrm{d}\mathbf{\Delta}g_{|\Omega}}{\mathrm{d}\mathfrak{m}} \in L^2_{\mathrm{loc}}(\Omega, \mathfrak{m}_{|\Omega})$. *Then* (4.24) *holds.*

proof
(**i**) This is a particular case of Proposition 4.11.
(**ii**) Let $\Omega' \in \mathrm{Int}(\Omega)$ and observe that since $\overline{\Omega'}$ is compact, $g(\overline{\Omega'})$ is compact as well and thus φ'' is bounded on $g(\Omega')$. It easily follows that the formula $\tilde{\mu} := \varphi' \circ g\, \mathbf{\Delta}g_{|\Omega} + \varphi'' \circ g\, |Dg|_w^2 \mathfrak{m}_{|\Omega}$ defines a locally finite measure on Ω, so that the statement makes sense.

Now notice that $\mathrm{Test}(\Omega) \subset L^1(\Omega, |\tilde{\mu}|)$, pick $f \in \mathrm{Test}(\Omega)$ and use (4.16) to get

$$\begin{aligned}
\nabla f \cdot \nabla(\varphi \circ g) = \varphi' \circ g\, \nabla f \cdot \nabla g &= \nabla(f\varphi' \circ g) \cdot \nabla g - f\, \nabla(\varphi' \circ g) \cdot \nabla g \\
&= \nabla(f\varphi' \circ g) \cdot \nabla g - f\varphi'' \circ g|Dg|_w^2 \qquad \mathfrak{m} - a.e.,
\end{aligned} \quad (4.25)$$

which by integration gives

$$-\int \nabla f \cdot \nabla(\varphi \circ g)\, \mathrm{d}\mathfrak{m} = -\int \nabla(f\varphi' \circ g) \cdot \nabla g\, \mathrm{d}\mathfrak{m} + \int f\varphi'' \circ g|Dg|_w^2\, \mathrm{d}\mathfrak{m}. \quad (4.26)$$

Hence to conclude it is sufficient to show that

$$-\int \nabla(f\varphi' \circ g) \cdot \nabla g\, \mathrm{d}\mathfrak{m} = \int f\varphi' \circ g\, \mathrm{d}\mathbf{\Delta}g_{|\Omega}. \quad (4.27)$$

This is a consequence of Lemma 4.26 applied to $\psi := f\varphi' \circ g$, which by our assumptions belongs to $\mathrm{S}^2(X, \mathsf{d}, \mathfrak{m}) \cap C_c(\Omega)$.
(**iii**) From the assumptions we know that $\varphi' \circ g \in L^2_{\mathrm{loc}}(\Omega, \mathfrak{m}_{|\Omega})$ and $\varphi'' \circ g \in L^\infty_{\mathrm{loc}}(\Omega, _{|\Omega})$. Therefore, since $\mathbf{\Delta}g_{|\Omega} \ll \mathfrak{m}$ with L^2_{loc} density, the formula $\tilde{\mu} := \varphi' \circ g\, \mathbf{\Delta}g_{|\Omega} + \varphi'' \circ g\, |Dg|_w^2 \mathfrak{m}_{|\Omega}$ defines a locally finite measure on Ω and the statement makes sense. As before, we have $\mathrm{Test}(\Omega) \subset L^1(\Omega, |\tilde{\mu}|)$. With the same computations done in (4.25), (4.26) we reduce to show that (4.27) holds also in this case. This is a consequence of Lemma 4.27 applied to $\psi := f\varphi' \circ g$, which belongs to $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ and has bounded support contained in Ω. $\qquad \square$

The next theorem shows that in this linear case, the Leibniz rule holds also for the Laplacian, a property which, being consequence of the second identity in (4.16), is not available in the general non-linear setting. As for the chain rule, the theorem is stated in 3 different ways, depending on the assumptions on g_i, $\mathbf{\Delta}g_{i|\Omega}$ and the space.

Theorem 4.29 (Leibniz rule for the Laplacian) *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be an infinitesimally Hilbertian space,* $\Omega \subset X$ *an open set and* $g_1, g_2 \in D(\mathbf{\Delta}, \Omega) \cap \mathrm{S}^2_{\mathrm{int}}(\Omega)$. *Then the following is true.*

i) *If* g_1, g_2 *are Lipschitz on* Ω' *for every* $\Omega' \in \mathrm{Int}(\Omega)$, *and* $g_1, g_2 \in D_{\mathrm{ifm}}(\mathbf{\Delta}, \Omega)$, *then* $g_1 g_2 \in D(\mathbf{\Delta}, \Omega)$ *and*

$$\mathbf{\Delta}(g_1 g_2)_{|\Omega} = g_1 \mathbf{\Delta}g_{2|\Omega} + g_2 \mathbf{\Delta}g_{1|\Omega} + 2\nabla g_1 \cdot \nabla g_2\, \mathfrak{m}. \quad (4.28)$$

ii) If (X, d) *is proper and* $g_1, g_2 \in C(\Omega)$, *then* $g_1 g_2 \in D(\mathbf{\Delta}, \Omega)$ *and* (4.28) *holds.*

iii) If (X, d) *is proper,* $g_1, g_2 \in L^2_{\mathrm{loc}}(\Omega, \mathfrak{m}_{|\Omega}) \cap L^\infty_{\mathrm{loc}}(\Omega, \mathfrak{m}_{|\Omega})$ *and* $\mathbf{\Delta} g_{i|\Omega} \ll \mathfrak{m}$ *with* $L^2_{\mathrm{loc}}(\Omega, \mathfrak{m}_{|\Omega})$ *density,* $i = 1, 2$, *then* $g_1 g_2 \in D(\mathbf{\Delta}, \Omega)$ *and* (4.28) *holds.*

proof
(**i**) Being Lipschitz, g_1, g_2 are bounded on Ω' for any $\Omega' \in \mathrm{Int}(\Omega)$, hence $g_1 g_2 \in \mathrm{S}^2_{\mathrm{int}}(\Omega)$. It is also clear that the right hand side of (4.28) defines a locally finite measure μ on Ω, so that the statement makes sense. The fact that $|\mathbf{\Delta} g_{i|\Omega}|(\Omega')$ is finite for every $\Omega' \in \mathrm{Int}(\Omega)$, $i = 1, 2$, grants that $\mathrm{Test}(\Omega) \subset L^1(\Omega, |\mathbf{\Delta} g_{i|\Omega}|)$, $i = 1, 2$, and $\mathrm{Test}(\Omega) \subset L^1(\Omega, |\mu|)$. To conclude, pick $f \in \mathrm{Test}(\Omega)$, notice that $fg_1, fg_2 \in \mathrm{Test}(\Omega)$ and take the Leibniz rule (4.16) into account to get

$$\nabla f \cdot \nabla(g_1 g_2) = g_1 \nabla f \cdot \nabla g_2 + g_2 \nabla f \cdot \nabla g_1 = \nabla(fg_1) \cdot \nabla g_2 + \nabla(fg_2) \cdot \nabla g_1 - 2f \, \nabla g_1 \cdot \nabla g_2, \tag{4.29}$$

which integrated gives the thesis.
(**ii**) As before, the right hand side of (4.28) defines a locally finite measure μ, and as before $g_1 g_2 \in \mathrm{S}^2_{\mathrm{int}}(\Omega)$, $\mathrm{Test}(\Omega) \subset L^1(\Omega, |\mathbf{\Delta} g_{i|\Omega}|)$, $i = 1, 2$, and $\mathrm{Test}(\Omega) \subset L^1(\Omega, |\mu|)$. Pick $f \in \mathrm{Test}(\Omega)$ and notice that with the same computations done in (4.29) the thesis follows if we show that

$$\int_\Omega \nabla(fg_1) \cdot \nabla g_2 \, \mathrm{d}\mathfrak{m} = -\int_\Omega fg_1 \, \mathrm{d}\mathbf{\Delta} g_{2|\Omega},$$
$$\int_\Omega \nabla(fg_2) \cdot \nabla g_1 \, \mathrm{d}\mathfrak{m} = -\int_\Omega fg_2 \, \mathrm{d}\mathbf{\Delta} g_{1|\Omega}. \tag{4.30}$$

These are a consequence of Lemma 4.26 applied to $\psi := fg_1$, $g := g_2$ and $\psi := fg_2$, $g := g_1$ respectively.
(**iii**) Same as (*ii*) with Lemma 4.27 in place of Lemma 4.26 to justify (4.30) $\qquad\qquad\square$

In the discussion we made up to now, the Laplacian has always been understood by means of the integration by parts formula in Definition 4.4. On a Euclidean setting there is at least another completely different approach: to look at the behavior of the heat flow for small times. It is therefore natural to try to understand whether a similar approach is possible also in the non-smooth context. Although it is always possible to define the heat flow as the gradient flow of Ch_2 in L^2, in practice, in order for it to have at least some of the regularization properties available in a Euclidean context, it seems better to restrict the attention to spaces with *Riemannian Ricci curvature bounded from below*. This class of spaces was introduced in [6] as a strengthening of the standard $\mathsf{CD}(K, \infty)$ class which rules out Finsler geometries. On such spaces the heat flow can be equivalently regarded as gradient flow of Ch_2 in L^2 or as gradient flow of the relative entropy in $(\mathscr{P}_2(X), W_2)$, it is linear, self adjoint and has some very general regularization properties, see Theorem 4.31 below.

Recall that for $\mathfrak{m} \in \mathscr{P}(X)$ the relative entropy functional $\mathrm{Ent}_\mathfrak{m} : \mathscr{P}(X) \to [0, +\infty]$ is defined as

$$\mathrm{Ent}_\mathfrak{m}(\mu) := \begin{cases} \displaystyle\int \rho \log \rho \, \mathrm{d}\mathfrak{m}, & \text{if } \mu = \rho\mathfrak{m}, \\ +\infty, & \text{otherwise.} \end{cases}$$

By $D(\mathrm{Ent}_\mathfrak{m})$ we denote the set of μ's such that $\mathrm{Ent}_\mathfrak{m}(\mu) < \infty$.

Definition 4.30 (Riemannian Ricci curvature bounds) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) with $\mathfrak{m} \in \mathscr{P}(X)$ and $K \in \mathbb{R}$. We say that it is a $\mathsf{RCD}(K, \infty)$ space provided it is infinitesimally Hilbertian and for any $\mu, \nu \in D(\mathrm{Ent}_\mathfrak{m}) \cap \mathscr{P}_2(X)$ there exists $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu, \nu)$ such that the inequality*

$$\mathrm{Ent}_\mathfrak{m}((\mathrm{e}_t)_\sharp \boldsymbol{\pi}_F) \leq (1-t)\mathrm{Ent}_\mathfrak{m}((\mathrm{e}_0)_\sharp \boldsymbol{\pi}_F) + t\mathrm{Ent}_\mathfrak{m}((\mathrm{e}_1)_\sharp \boldsymbol{\pi}_F) - \frac{K}{2}t(1-t)W_2^2\big((\mathrm{e}_0)_\sharp \boldsymbol{\pi}_F, (\mathrm{e}_1)_\sharp \boldsymbol{\pi}_F\big),$$
(4.31)

holds for any $t \in [0,1]$, where $F : \mathrm{Geo}(X) \to \mathbb{R}$ is any Borel non-negative bounded function such that $\int F \, \mathrm{d}\boldsymbol{\pi} = 1$ and $\boldsymbol{\pi}_F := F\boldsymbol{\pi}$.

Notice that the usual K-geodesic convexity of the entropy is enforced by asking it along all weighted plans $\boldsymbol{\pi}_F$ in (4.31)[3].

Let $(X, \mathsf{d}, \mathfrak{m})$ be a $\mathsf{RCD}(K, \infty)$ space. Denote by $\mathsf{H}_t : L^2(X, \mathfrak{m}) \to L^2(X, \mathfrak{m})$ the gradient flow of Ch_2 (which clearly exists and is unique), i.e. for any $f \in L^2(X, \mathfrak{m})$ the curve $t \mapsto \mathsf{H}_t(f)$ is the gradient flow of Ch_2 in $L^2(X, \mathfrak{m})$ starting from f. Similarly, denote by $\mathscr{H}_t : D(\mathrm{Ent}_\mathfrak{m}) \cap \mathscr{P}_2(X) \to D(\mathrm{Ent}_\mathfrak{m}) \cap \mathscr{P}_2(X)$ the gradient flow of $\mathrm{Ent}_\mathfrak{m}$ in $(\mathscr{P}_2(X), W_2)$ (uniqueness has been established in [19], existence in [3], [19], [6] for, respectively, compact / locally compact / extended Polish spaces).

In the next theorem we collect those results proved in [4], [6] which we will use in the rest of the section.

Theorem 4.31 (Properties of the heat flow) *Let $K \in \mathbb{R}$ and $(X, \mathsf{d}, \mathfrak{m})$ a $\mathsf{RCD}(K, \infty)$ space. Then the following hold.*

i) *Coincidence of the two flows Let $f \in L^2(X, \mathfrak{m})$ be such that $f\mathfrak{m} \in \mathscr{P}_2(X)$. Then for any $t \geq 0$ it holds $(\mathsf{H}_t f)\mathfrak{m} = \mathscr{H}_t(f\mathfrak{m})$.*

ii) *Linearity and extendibility H_t is a linear semigroup which can be uniquely extended to a continuous semigroup on $L^1(X, \mathfrak{m})$, still denoted by H_t. Similarly, $\mathscr{H}_t$ is a linear semigroup which can be uniquely extended to a continuous linear semigroup, still denoted by $\mathscr{H}_t$, on the space $\mathcal{M}(\mathrm{supp}(\mathfrak{m}))$ of Radon measures with finite mass on $\mathrm{supp}(\mathfrak{m})$, where continuity is meant in duality with $C_b(\mathrm{supp}(\mathfrak{m}))$.*

iii) *Regularization properties For $t > 0$, $\mathscr{H}_t$ maps $\mathcal{M}(\mathrm{supp}(\mathfrak{m}))$ on $D(\mathrm{Ent}_\mathfrak{m})$ and H_t maps contractively $L^\infty(X, \mathfrak{m})$ in $C_b(\mathrm{supp}(\mathfrak{m}))$, with $L^\infty \to \mathrm{Lip}$ regularization:*

$$\mathrm{Lip}(\mathsf{H}_t f) \leq \frac{\|f\|_{L^\infty}}{\sqrt{2 \int_0^t e^{2Ks} \, \mathrm{d}s}}.$$
(4.32)

iv) *Self-adjointness For any $f \in L^\infty(X, \mathfrak{m})$, $\mu \in \mathcal{M}(\mathrm{supp}(\mathfrak{m}))$ and $t > 0$ it holds*

$$\int f \, \mathrm{d}\mathscr{H}_t \mu = \int \mathsf{H}_t f \, \mathrm{d}\mu.$$
(4.33)

v) *Compatibility with the Dirichlet form theory The intrinsic distance $\mathsf{d}_\mathcal{E}$ associated to $\mathcal{E}$ coincides with d.*

---

[3]in the recent paper [2] it is shown that this assumption is in fact redundant, as it follows by the standard $\mathsf{CD}(K, \infty)$ condition plus infinitesimal Hilbertianity

vi) *Validity of the Bakry-Emery curvature condition* For any $f \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ *it holds*

$$|D(\mathsf{H}_t f)|_w^2 \le e^{-2Kt} \mathsf{H}_t(|Df|_w^2), \qquad \mathfrak{m} - a.e. \ \forall t \ge 0. \tag{4.34}$$

As expected, the Laplacian commutes with the heat flow:

Proposition 4.32 ($\boldsymbol{\Delta}\mathsf{H}_t g = \mathscr{H}_t \boldsymbol{\Delta} g$) *Let* $K \in \mathbb{R}$, $(X, \mathsf{d}, \mathfrak{m})$ *a* $\mathsf{RCD}(K, \infty)$ *space and* $g \in W^{1,2}(X, \mathsf{d}, \mathfrak{m}) \cap D(\boldsymbol{\Delta})$ *with* $\|\boldsymbol{\Delta} g\|_{\mathrm{TV}} < \infty$. *Then for every* $t \ge 0$ *it holds*

$$\boldsymbol{\Delta}\big(\mathsf{H}_t g\big) = \mathscr{H}_t\big(\boldsymbol{\Delta} g\big). \tag{4.35}$$

proof Taking into account (4.33) and the definition of $\boldsymbol{\Delta}$ we need to prove that for any $f \in \mathrm{Test}(X)$ it holds

$$\int \nabla f \cdot \nabla \mathsf{H}_t g \, \mathrm{d}\mathfrak{m} = \int \nabla \mathsf{H}_t f \cdot \nabla g \, \mathrm{d}\mathfrak{m}.$$

Given that $f, g \in D(\mathcal{E})$, such equality is a standard consequence of the self-adjointness of the evolution semigroup associated to a Dirichlet form, we omit the details. $\qquad\square$

Having a well defined heat flow allows, as we discussed before, to give a definition of Laplacian based on its infinitesimal behavior.

Definition 4.33 (A variant of the definition of Laplacian) *Let* $K \in \mathbb{R}$, $(X, \mathsf{d}, \mathfrak{m})$ *be an* $\mathsf{RCD}(K, \infty)$ *space and* $g \in L^1(X, \mathfrak{m})$. *We say that* $g \in D(\tilde{\boldsymbol{\Delta}})$ *provided there exists a locally finite measure* μ *such that*

$$\lim_{t \downarrow 0} \int f \frac{\mathsf{H}_t g - g}{t} \, \mathrm{d}\mathfrak{m} = \int f \, \mathrm{d}\mu,$$

for any $f \in C_b(X) \cap L^1(X, |\mu|)$ *with bounded support. In this case we write* $\tilde{\boldsymbol{\Delta}} g = \mu$ *(notice that μ is certainly unique).*

Given that the heat semigroup is linear, it is trivial that $D(\tilde{\boldsymbol{\Delta}}) \subset L^1$ is a vector space and that $\tilde{\boldsymbol{\Delta}}$ is linear. Our first task is to compare this version of the Laplacian with the one previously introduced in Definition 4.4. Notice that Proposition 4.24 grants that

$$g \in D(\Delta), \quad h = \Delta g \quad \Leftrightarrow \quad g \in W^{1,2}(X, \mathsf{d}, \mathfrak{m}) \cap D(\boldsymbol{\Delta}), \quad \boldsymbol{\Delta} g = h\mathfrak{m},$$
$$\Leftrightarrow \quad g \in W^{1,2}(X, \mathsf{d}, \mathfrak{m}) \cap D(\tilde{\boldsymbol{\Delta}}), \quad \tilde{\boldsymbol{\Delta}} g = h\mathfrak{m}.$$

The following statement generalizes such result to situations where the Laplacian is a measure.

Proposition 4.34 *Let* $K \in \mathbb{R}$, $(X, \mathsf{d}, \mathfrak{m})$ *a* $\mathsf{RCD}(K, \infty)$ *space and* $g \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$. *Assume that* $g \in D(\tilde{\boldsymbol{\Delta}})$. *Then* $g \in D(\boldsymbol{\Delta})$ *and* $\boldsymbol{\Delta} g = \tilde{\boldsymbol{\Delta}} g$. *Viceversa, if* $g \in D(\boldsymbol{\Delta})$ *and* $\|\boldsymbol{\Delta} g\|_{\mathrm{TV}} < \infty$, *then* $g \in D(\tilde{\boldsymbol{\Delta}})$ *and* $\tilde{\boldsymbol{\Delta}} g = \boldsymbol{\Delta} g$.

proof We know that $\mathsf{H}_s g \to g$ in $L^2(X, \mathfrak{m})$ as $s \downarrow 0$ and, from the lower semicontinuity of Ch_2 and the fact that Ch_2 decreases along the heat flow, $\mathrm{Ch}_2(\mathsf{H}_s g) \to \mathrm{Ch}_2(g)$ as $s \to 0$. Hence, since $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is Hilbert, we deduce $\mathsf{H}_s g \to g$ in $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ as $s \to 0$.

Let $f \in \mathrm{Test}(X) \cap L^1(X, |\tilde{\boldsymbol{\Delta}} g|) \subset W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ and notice that

$$\int f \frac{\mathsf{H}_t g - g}{t} \, \mathrm{d}\mathfrak{m} = \frac{1}{t} \iint_0^t f \Delta \mathsf{H}_s g \, \mathrm{d}s \, \mathrm{d}\mathfrak{m} = -\frac{1}{t} \int_0^t \int \nabla f \cdot \nabla \mathsf{H}_s g \, \mathrm{d}\mathfrak{m} \, \mathrm{d}s$$
$$= -\int \nabla f \cdot \nabla g \, \mathrm{d}\mathfrak{m} + \mathrm{Rem}_t,$$

where the reminder term Rem_t is bounded by

$$\left|\mathrm{Rem}_t\right| \leq \||Df|_w\|_{L^2(X,\mathfrak{m})} \frac{1}{t} \int_0^t \||D(\mathsf{H}_s g - g)|_w\|_{L^2(X,\mathfrak{m})}\, \mathrm{d}s.$$

The first part of the statement follows.

For the second, fix $f \in C_b(X)$ and use (4.35) and (4.33) to get

$$\int f \frac{\mathsf{H}_t g - g}{t}\, \mathrm{d}\mathfrak{m} = \frac{1}{t} \int_0^t \int f \mathrm{d}\boldsymbol{\Delta}\mathsf{H}_s g\, \mathrm{d}s = \int \left(\frac{1}{t} \int_0^t \mathsf{H}_s f\, \mathrm{d}s \right) \mathrm{d}\boldsymbol{\Delta}g.$$

By (iii) of Theorem 4.31 we have $\|\mathsf{H}_s f\|_{L^\infty} \leq \|f\|_{L^\infty}$, thus by the dominate convergence theorem, to conclude it is sufficient to show that $\mathsf{H}_s f(x) \to f(x)$ for any $x \in \mathrm{supp}(\mathfrak{m})$ as $s \downarrow 0$ (recall that $|\boldsymbol{\Delta}g|(X \setminus \mathrm{supp}(\mathfrak{m})) = 0$). Using again (4.33) we have

$$\mathsf{H}_s f(x) = \int \mathsf{H}_s f(y)\, \mathrm{d}\delta_x(y) = \int f(y)\, \mathrm{d}\mathscr{H}_s(\delta_x)(y),$$

and the conclusion follows from the weak continuity statement in (ii) of Theorem 4.31. $\qquad\square$

The following is a natural variant of Proposition 4.13. Notice that the use of the theory of semigroups allows us to conclude that the $\underline{\lim}$ at the left hand side of (4.36) is always a limit. The compactness assumption can be dropped at least if the measure is doubling (in this case, also a weak local Poincaré inequality holds, see [41]).

Proposition 4.35 *Let $K \in \mathbb{R}$, $(X, \mathsf{d}, \mathfrak{m})$ a compact $\mathsf{RCD}(K, \infty)$ space, $g \in L^1(X, \mathfrak{m})$ and μ a Radon measure on X. Assume that*

$$\varliminf_{t \downarrow 0} \int f \frac{\mathsf{H}_t g - g}{t}\, \mathrm{d}\mathfrak{m} \geq \int f\, \mathrm{d}\mu, \qquad \forall f \in C(X),\ f \geq 0. \tag{4.36}$$

Then $g \in D(\tilde{\boldsymbol{\Delta}})$ and $\tilde{\boldsymbol{\Delta}}g \geq \mu$.

proof It is obvious that (4.36) holds also with $\mu|_{\mathrm{supp}(\mathfrak{m})}$ in place of μ and that it implies $\mu|_{X \setminus \mathrm{supp}(\mathfrak{m})} \leq 0$, thus we can assume $\mathrm{supp}(\mathfrak{m}) = X$.

Define the functionals $L^\pm : C(X) \to \mathbb{R} \cup \{\pm\infty\}$ by

$$L^-(f) := \varliminf_{t \downarrow 0} \int f \frac{\mathsf{H}_t g - g}{t}\, \mathrm{d}\mathfrak{m}, \qquad L^+(f) := \varlimsup_{t \downarrow 0} \int f \frac{\mathsf{H}_t g - g}{t}\, \mathrm{d}\mathfrak{m}. \tag{4.37}$$

It is clear that both are positively 1-homogeneous, that L^- is concave and L^+ convex. Moreover, since $\int \mathsf{H}_t g\, \mathrm{d}\mathfrak{m} = \int g\, \mathrm{d}\mathfrak{m}$, it holds $L^\pm(f + c) = L^\pm(f)$ for any $f \in C(X)$, $c \in \mathbb{R}$.

Pick a generic $f \in C(X)$, notice that $f + \|f\|_{L^\infty} \geq 0$ so that from (4.36) we derive

$$L^-(f) = L^-(f + \|f\|_{L^\infty}) \geq \int f + \|f\|_{L^\infty}\, \mathrm{d}\mu \geq -2\|f\|_{L^\infty}\|\mu\|_{\mathrm{TV}}. \tag{4.38}$$

Then using the identity $L^-(-f) = -L^+(f)$ we get

$$-2\|f\|_{L^\infty}\|\mu\|_{\mathrm{TV}} \leq L^-(f) \leq L^+(f) \leq 2\|f\|_{L^\infty}\|\mu\|_{\mathrm{TV}}, \qquad \forall f \in C(X).$$

Hence the inequality $L^-(f) - L^-(g) \geq L^-(f - g) \geq -2\|f - g\|_{L^\infty}\|\mu\|_{\mathrm{TV}}$, and the similar one written with reversed roles, gives that $L^- : C(X) \to \mathbb{R}$ is Lipschitz. Similarly for L^+.

Now fix a non-negative function $\varphi \in C_c^\infty(0,1)$ such that $\int_0^\infty \varphi(t)\,\mathrm{d}t = 1$. For any $f \in$ LIP(X), $r > 0$ define $\tau_r f : X \to \mathbb{R}$ by

$$\tau_r f := \frac{1}{r} \int_0^\infty \mathsf{H}_s f\, \varphi\Big(\frac{s}{r}\Big)\,\mathrm{d}s.$$

From (4.34) it follows that $\mathrm{Lip}(\tau_r f) \leq \mathrm{Lip}(f) \int_0^\infty e^{-Ksr}\varphi(s)\,\mathrm{d}s$ for any $r > 0$ (see Section 6.1 of [6] for details), and therefore the $L^2(X,\mathfrak{m})$ convergence of $\tau_r f$ to f implies uniform convergence.

By standard semigroups computations we know that $\tau_r f \in D(\Delta)$ and that $\Delta \tau_r f = \frac{1}{r}\int_0^\infty \Delta \mathsf{H}_s f \varphi(\frac{s}{r})\,\mathrm{d}s$. Hence for any $r > 0$ we have

$$L^-(\tau_r f) = \int \Delta \tau_r f g\,\mathrm{d}\mathfrak{m} = L^+(\tau_r f).$$

Since L^-, L^+ are both continuous on $C(X)$, by letting $r \downarrow 0$ we deduce that $L^+(f) = L^-(f)$ for any $f \in$ LIP(X) and hence they coincide also on $C(X)$. The conclusion follows from the Riesz theorem. $\qquad\square$

The notion of Laplacian $\tilde{\boldsymbol{\Delta}}$ allows to prove the Bochner inequality with $N = \infty$ in compact RCD(K,∞) spaces[4]: shortly said, this will follow by the standard argument consisting in derivating at time $t = 0$ the Bakry-Emery condition (4.34). Recall that by Δ we intend the generator of the Dirichlet form $\mathcal{E}$, see (4.19).

Theorem 4.36 (Bochner inequality with $N = \infty$) *Let $K \in \mathbb{R}$, $(X, \mathsf{d}, \mathfrak{m})$ a compact* RCD(K,∞) *space and $g \in D(\Delta)$ with $\Delta g \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$. Then $|Dg|_w^2 \in D(\tilde{\boldsymbol{\Delta}})$ and it holds*

$$\tilde{\boldsymbol{\Delta}}\frac{|Dg|_w^2}{2} \geq \big(\nabla \Delta g \cdot \nabla g + K|Dg|_w^2\big)\mathfrak{m}. \tag{4.39}$$

Note: for any $f \in L^2(X, \mathfrak{m})$, $t > 0$, the function $g := \mathsf{H}_t f$ fulfills the hypothesis.

proof By standard semigroup computation we get that for g as in the hypothesis it holds

$$\lim_{t \downarrow 0} \frac{|D(\mathsf{H}_t g)|_w^2 - |Dg|_w^2}{t} = 2\nabla \Delta g \cdot \nabla g, \qquad \text{in } L^2(X, \mathfrak{m}), \tag{4.40}$$

we omit the details. Now recall that from (4.34) we know that it holds

$$|D(\mathsf{H}_t g)|_w^2 \leq e^{-2Kt}\mathsf{H}_t(|Dg|_w^2), \qquad \forall t \geq 0.$$

Notice that for $t = 0$ this is an equality, multiply both sides by a continuous and non-negative function $f : X \to \mathbb{R}$ to get

$$\int f \frac{|D(\mathsf{H}_t g)|_w^2 - |Dg|_w^2}{t}\,\mathrm{d}\mathfrak{m} \leq \int f\frac{e^{-2Kt}\mathsf{H}_t(|Dg|_w^2) - |Dg|_w^2}{t}\,\mathrm{d}\mathfrak{m}.$$

---

[4]in the recent paper [43], Savaré proved that for 'many' functions g as in the statement of Theorem 4.36 one has $|Dg|_w \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$, so that thanks to Proposition 4.34 in the left hand side we can write the Laplacian $\boldsymbol{\Delta}$ instead of $\tilde{\boldsymbol{\Delta}}$. The regularity result $|Dg|_w \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is non-trivial and follows from a generalization of works of Barky in the context of Dirichlet forms [9]

Using (4.40) we get

$$2 \int f \nabla \Delta g \cdot \nabla g \, d\mathfrak{m} \leq \varliminf_{t \downarrow 0} \int f \frac{e^{-2Kt} \mathsf{H}_t(|Dg|_w^2) - |Dg|_w^2}{t} \, d\mathfrak{m}.$$

since $\mathsf{H}_t(|Dg|_w^2)\mathfrak{m}$ weakly converges to $|Dg|_w^2\mathfrak{m}$, it is immediate to check that for the right hand side it holds

$$\varliminf_{t \downarrow 0} \int f \frac{e^{-2Kt} \mathsf{H}_t(|Dg|_w^2) - |Dg|_w^2}{t} \, d\mathfrak{m} = -2K \int f |Dg|_w^2 \, d\mathfrak{m} + \varliminf_{t \downarrow 0} \int f \frac{\mathsf{H}_t(|Dg|_w^2) - |Dg|_w^2}{t} \, d\mathfrak{m}$$

therefore the conclusion follows from Proposition 4.35 with $|Dg|_w^2$ in place of g and $\mu := 2(\nabla \Delta g \cdot \nabla g + K|Dg|_w^2)\mathfrak{m}$. $\qquad\square$

Remark 4.37 The derivation of the Bochner inequality with $N = \infty$ from the contraction estimate (4.34) is classical. What is interesting to observe is the link between (4.34) and the K-geodesic convexity of the entropy. Indeed, even on a smooth Riemannian setting, in deriving the former from the latter, the fact that the tangent space is Hilbert is used twice:

  i) A first time to deduce the K-contractivity of W_2 along two heat flows.

  ii) A second time to pass, by a duality argument, from the K-contractivity of W_2 to (4.34) (see e.g. the general argument in [26]).

It is known that on Finsler manifolds the K-convexity of the entropy does *not* imply the K-contractivity of W_2 along two heat flows (as proved by Ohta-Sturm in [36]), and the duality argument in (*ii*) works only for linear flows, so that when applied to the heat flow, it works only on Riemannian manifolds.

The fact that a duality argument is used twice, suggests that there should be a way to state and prove the Bochner inequality also on a Finsler setting satisfying the $\mathsf{CD}(K, \infty)$ condition. This is actually the case: more generally, the Bochner inequality, also with finite N, holds also on a Finsler structure, provided an appropriate reformulation of the term $\Delta \frac{|\nabla \varphi|^2}{2}$ is considered (see [35] and [20] for two different approaches). $\qquad\blacksquare$

5 Comparison estimates

In this chapter we recall the definition of metric measure spaces with Ricci curvature $\geq K$ and dimension $\leq N$, and prove that on these spaces the sharp Laplacian comparison estimates for the distance function holds, at least on the infinitesimally strictly convex case.

5.1 Weak Ricci curvature bounds

Here we recall the definition of $\mathsf{CD}(K, N)$, $N < \infty$, spaces given by Sturm in [47] and Lott-Villani in [30] (the latter reference deals with the case $K = 0$ only).

Let $u : [0, \infty) \to \mathbb{R}$ be a convex continuous and sublinear (i.e. $\lim_{z \to +\infty} \frac{u(z)}{z} = 0$) function satisfying $u(0) = 0$. Let $\mathcal{M}^+(X)$ the space of finite non-negative Radon measures on X. The *internal energy* functional $\mathcal{U} : \mathcal{M}^+(X) \to \mathbb{R} \cup \{+\infty\}$ associated to u is well defined by the formula

$$\mathcal{U}(\mu) := \int u(\rho) \, d\mathfrak{m}, \qquad \mu = \rho\mathfrak{m} + \mu^s, \ \mu^s \perp \mathfrak{m}.$$

Jensen's inequality ensures that if $\mathfrak{m}(\mathrm{supp}(\mu)) < \infty$, then $\mathcal{U}(\mu) > -\infty$. More generally, the functional $\mathcal{U}$ is lower semicontinuous in $\mathscr{P}(B) \subset \mathcal{M}^+(X)$ w.r.t. convergence in duality with $C_b(B)$, for any closed set B such that $\mathfrak{m}(B) < \infty$.

Functions u of interest for us are

$$u_N(z) := -z^{1-\frac{1}{N}}, \qquad N \in (1, \infty),$$

and we will denote the associated internal energies by $\mathcal{U}_N$ respectively.

For $N \in (1, \infty)$, $K \in \mathbb{R}$ the *distortion coefficients* $\tau_{K,N}^{(t)}(\theta)$ are functions $[0,1] \times [0, \infty) \ni (t, \theta) \mapsto \tau_{K,N}^{(t)}(\theta) \in [0, +\infty]$ defined by

$$
\tau_{K,N}^{(t)}(\theta) := \begin{cases}
+\infty, & \text{if } K\theta^2 \geq (N-1)\pi^2, \\[2mm]
t^{\frac{1}{N}} \left(\dfrac{\sin(t\theta\sqrt{K/(N-1)})}{\sin(\theta\sqrt{K/(N-1)})} \right)^{1-\frac{1}{N}}, & \text{if } 0 < K\theta^2 < (N-1)\pi^2, \\[4mm]
t, & \text{if } K\theta^2 = 0, \\[2mm]
t^{\frac{1}{N}} \left(\dfrac{\sinh(t\theta\sqrt{-K/(N-1)})}{\sinh(\theta\sqrt{-K/(N-1)})} \right)^{1-\frac{1}{N}}, & \text{if } K\theta^2 < 0.
\end{cases}
$$

Definition 5.1 (Weak Ricci curvature bound) *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be a boundedly finite metric measure space. We say that* $(X, \mathsf{d}, \mathfrak{m})$ *is a* $\mathsf{CD}(K, N)$ *space,* $K \in \mathbb{R}$, $N \in (1, \infty)$ *provided for any* $\mu, \nu \in \mathscr{P}(\mathrm{supp}(\mathfrak{m}))$ *with bounded support there exists* $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu, \nu)$ *such that*

$$\mathcal{U}_{N'}((\mathsf{e}_t)_\sharp \boldsymbol{\pi}) \leq -\int \tau_{K,N'}^{(1-t)}\big(\mathsf{d}(\gamma_0, \gamma_1)\big) \rho^{-\frac{1}{N'}}(\gamma_0) + \tau_{K,N'}^{(t)}\big(\mathsf{d}(\gamma_0, \gamma_1)\big) \eta^{-\frac{1}{N'}}(\gamma_1) \, \mathrm{d}\boldsymbol{\pi}(\gamma), \quad \forall t \in [0, 1], \tag{5.1}$$

for any $N' \geq N$, *where* $\mu = \rho\mathfrak{m} + \mu^s$ *and* $\nu = \eta\mathfrak{m} + \nu^s$, *with* $\mu^s, \nu^s \perp \mathfrak{m}$.

When dealing with a $\mathsf{CD}(K, N)$ space, we will always assume that $\mathrm{supp}(\mathfrak{m}) = X$ (this is dangerous only when discussing stability issues), and to avoid trivialities we also assume that X contains more than one point. In the following proposition we collect those basic properties of $\mathsf{CD}(K, N)$ spaces we will use later on.

Proposition 5.2 (Basic properties of $\mathsf{CD}(K, N)$ spaces) *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{CD}(K, N)$ *space,* $K \in \mathbb{R}$, $N < \infty$. *Then* (X, d) *is proper and geodesic,* $\mathfrak{m}$ *is doubling and* $(X, \mathsf{d}, \mathfrak{m})$ *supports a weak local 1-1 Poincaré inequality, i.e. for any bounded Borel function* $f : X \to \mathbb{R}$ *and any upper gradient* G *of* f *and* $R > 0$ *it holds*

$$\frac{1}{\mathfrak{m}(B_r(x))} \int_{B_r(x)} \left| f - \langle f \rangle_{B_r(x)} \right| \mathrm{d}\mathfrak{m} \leq C(R) r \frac{1}{\mathfrak{m}(B_{2r}(x))} \int_{B_{2r}(x)} G \, \mathrm{d}\mathfrak{m}, \qquad \forall 0 < r < R,$$

where $\langle f \rangle_{B_r(x)} := \frac{1}{\mathfrak{m}(B_r(x))} \int_{B_r(x)} f \, \mathrm{d}\mathfrak{m}$, *for some constant* $C(R)$ *depending only on* K, N *and the upper bound on the radius* $R > 0$. *Also, the Bishop-Gromov comparison estimates holds,*

i.e. for any $x \in \mathrm{supp}(\mathfrak{m})$ it holds

$$\frac{\mathfrak{m}(B_r(x))}{\mathfrak{m}(B_R(x))} \geq \begin{cases} \dfrac{\int_0^r \sin(t\sqrt{K/(N-1)})^{N-1}\,\mathrm{d}t}{\int_0^R \sin(t\sqrt{K/(N-1)})^{N-1}\,\mathrm{d}t} & \text{if } K > 0 \\[3mm] \dfrac{r^N}{R^N} & \text{if } K = 0 \\[3mm] \dfrac{\int_0^r \sinh(t\sqrt{K/(N-1)})^{N-1}\,\mathrm{d}t}{\int_0^R \sinh(t\sqrt{K/(N-1)})^{N-1}\,\mathrm{d}t} & \text{if } K < 0 \end{cases} \qquad (5.2)$$

for any $0 < r \leq R \leq \pi\sqrt{(N-1)/(\max\{K,0\})}$.

Finally, if $K > 0$ then $(\mathrm{supp}(\mathfrak{m}), \mathsf{d})$ is compact and with diameter at most $\pi\sqrt{\frac{N-1}{K}}$.

proof For the Poincaré inequality see [29] for the original argument requiring the non-branching condition and the more recent paper [41] for the same result without such assumption. For the other properties, see [47] or the final Chapter of [48]. $\qquad \square$

We conclude this introduction recalling that on general metric measure spaces $(X, \mathsf{d}, \mathfrak{m})$, given a Lipschitz function $f : X \to \mathbb{R}$, typically the objects $\mathrm{lip}(f), \mathrm{lip}^+(f), \mathrm{lip}^-(f), |Df|_w$ are all different one from each other, the only information available being $|Df|_w \leq \mathrm{lip}^\pm(f)$ $\mathfrak{m}$-a.e. and $\mathrm{lip}^\pm(f) \leq \mathrm{lip}(f)$ everywhere. On doubling spaces supporting a weak local 1-1 Poincaré inequality, the fine analysis done by Cheeger in [13] ensures that these objects can be identified and, as already recalled in Remark 2.5, that the p-minimal weak upper gradient does not depend on p.

In particular, we have the following:

Theorem 5.3 *Let $(X, \mathsf{d}, \mathfrak{m})$ be a $\mathsf{CD}(K, N)$ space, $K \in \mathbb{R}$, $N \in (1, \infty)$. Then we have*

$$|Df|_w == \mathrm{lip}^+(f) = \mathrm{lip}^-(f)\mathrm{lip}(f), \qquad \mathfrak{m} - a.e. \qquad \forall f : X \to \mathbb{R} \text{ locally Lipschitz.} \quad (5.3)$$

Furthermore, the p-minimal weak upper gradient of s Sobolev function does not depend on $p > 1$ and in particular also the objects $D^\pm f(\nabla g)$ and $\mathbf{\Delta}$ do not depend on p.

5.2 Variants of calculus rules

In the introduction we saw that the heart of the proof of Laplacian comparison estimates for the distance function on $\mathsf{CD}(K, N)$ space is to differentiate the internal energy functional $\mathcal{U}_N$ along a Wasserstein geodesic. To make this computation on a metric measure space is complicated by the lack of a change of variable formula, which prevents the use of the standard analytical tools. In this abstract setting, this sort of computation has been done for the first time in [6] via an analogous of Theorem 3.10 relating horizontal and vertical derivatives. In order to apply such theorem, it was important to be sure that the Wasserstein geodesic along which the derivative was computed was made of absolutely continuous measures with uniformly bounded densities (at least for t close to 0), so that the plan associated to it was of bounded compression. Geodesics of this kind were obtained by assuming some strong form of geodesic convexity for the relative entropy.

Our purpose here is to show that as a consequence of Theorem 5.3, in the finite dimensional case no form of strong geodesic convexity is needed for the internal energy functionals in

order for the argument to be carried out. Shortly said, the assumption that $\boldsymbol{\pi}$ has bounded compression can be replaced by $\mathcal{U}_N((\mathrm{e}_t)_\sharp\boldsymbol{\pi}) \to \mathcal{U}_N((\mathrm{e}_0)_\sharp\boldsymbol{\pi})$ as $t \downarrow 0$.

The arguments of this section are based on the following simple lemma.

Lemma 5.4 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $(\mu_n) \subset \mathscr{P}(X)$ a sequence of measures and $\mu \in \mathscr{P}(X)$. Assume that for some closed set B with $\mathfrak{m}(B) < \infty$ it holds $\mathrm{supp}(\mu_n) \subset B$ for every $n \in \mathbb{N}$, $\mathrm{supp}(\mu) \subset B$ and that (μ_n) weakly converges to μ in duality with $C_b(B)$. Assume furthermore that*

$$\mu \ll \mathfrak{m}, \qquad and \qquad \mathcal{U}_N(\mu_n) \to \mathcal{U}_N(\mu), \qquad as \ n \to \infty.$$

Then for every Borel bounded function $f : B \to \mathbb{R}$ it holds

$$\lim_{n\to\infty} \int f \, \mathrm{d}\mu_n = \int f \, \mathrm{d}\mu. \tag{5.4}$$

proof Write $\mu = \rho\mathfrak{m}$, $\mu_n = \rho_n\mathfrak{m} + \mu_n^s$ with $\mu_n^s \perp \mathfrak{m}$. We need to prove that $\mu_n^s(B) \to 0$ and that $\rho_n \to \rho$ in duality with $L^\infty(B, \mathfrak{m})$. We argue by contradiction and assume that there exists a sequence $c_n \uparrow \infty$ such that

$$\int_{\{\rho_n \geq c_n\}} \rho_n \, \mathrm{d}\mathfrak{m} + \mu_n^s(B) \geq C > 0, \qquad \forall n \in \mathbb{N}. \tag{5.5}$$

Define the measures $\nu_n := \rho_n \chi_{\{\rho_n \leq c_n\}}\mathfrak{m} \leq \mu_n$ and notice that since (μ_n) is tight, so is (ν_n). Hence up to pass to a subsequence, not relabeled, we can assume that (ν_n) converges to some Borel measure $\nu \in \mathcal{M}^+(B)$ in duality with $C_b(B)$. From

$$\left| \mathcal{U}_N(\mu_n) - \mathcal{U}_N(\nu_n) \right| = \int_{\{\rho_n \geq c_n\}} |u_N(\rho_n)| \, \mathrm{d}\mathfrak{m} \leq \left(\int_{\{\rho_n \geq c_n\}} \rho_n \, \mathrm{d}\mathfrak{m} \right)^{\frac{N-1}{N}} \mathfrak{m}(\{\rho_n \geq c_n\})^{\frac{1}{N}},$$

and Chebyshev's inequality $\mathfrak{m}(\{\rho_n \geq c_n\}) \leq \frac{1}{c_n} \downarrow 0$, we get $\varlimsup_{n\to\infty} |\mathcal{U}_N(\mu_n) - \mathcal{U}_N(\nu_n)| = 0$ and thus

$$\mathcal{U}_N(\nu) \leq \varliminf_{n\to\infty} \mathcal{U}_N(\nu_n) = \varliminf_{n\to\infty} \mathcal{U}_N(\mu_n) = \mathcal{U}_N(\mu). \tag{5.6}$$

Write $\nu = \eta\mathfrak{m} + \nu^s$, with $\nu^s \perp \mathfrak{m}$ and notice that by construction it holds $\nu \leq \mu$, so that $\nu^s = 0$ and $\eta \leq \rho$ $\mathfrak{m}$-a.e.. Also, since u_N is strictly decreasing, (5.6) forces $\eta = \rho$ $\mathfrak{m}$-a.e., so that $\nu(B) = \int \eta \, \mathrm{d}\mathfrak{m} = \int \rho \, \mathrm{d}\mathfrak{m} = 1$. But this is a contradiction, because from (5.5) we have $\nu_n(B) \leq 1 - C$ for every $n \in \mathbb{N}$, so that $\nu(B) \leq 1 - C$ as well. $\qquad\square$

Notice that as a consequence of the convexity of $\mathcal{U}_N$ (w.r.t. affine interpolation) and its weak lower semicontinuity, the following implication holds:

$$\left\{ \begin{array}{l} \boldsymbol{\pi} \in \mathscr{P}(\mathrm{Geo}(X)) \\ \mathrm{supp}((\mathrm{e}_t)_\sharp\boldsymbol{\pi}) \subset B, \quad \forall t \in [0,1], \ \text{ for some bounded closed set } B \text{ s.t. } \mathfrak{m}(B) < \infty \\ \mathcal{U}_N((\mathrm{e}_t)_\sharp\boldsymbol{\pi}) \to \mathcal{U}_N((\mathrm{e}_0)_\sharp\boldsymbol{\pi}), \quad \text{as } t \downarrow 0 \end{array} \right\}$$

$$\Rightarrow \qquad \mathcal{U}_N\left(\frac{1}{t} \int_0^t (\mathrm{e}_s)_\sharp\boldsymbol{\pi} \, \mathrm{d}s \right) \to \mathcal{U}_N((\mathrm{e}_0)_\sharp\boldsymbol{\pi}), \quad \text{as } t \downarrow 0$$

$$\tag{5.7}$$

A direct consequence of Lemma 5.4, is the following version of the metric Brenier theorem proved in [4]. Notice that the argument we use here is essentially the same used in [4].

Proposition 5.5 (A variant of the metric Brenier theorem) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $B \subset X$ a bounded closed set such that $\mathfrak{m}(B) < \infty$, $\mu_0, \mu_1 \in \mathscr{P}(B)$ and $N \in (1, \infty)$. Assume that $\mu_0 \ll \mathfrak{m}$, that there exists $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu_0, \mu_1)$ such that $\lim_{t\downarrow 0} \mathcal{U}_N((\mathrm{e}_t)_\sharp \boldsymbol{\pi}) = \mathcal{U}_N(\mu_0)$ and $\mathrm{supp}((\mathrm{e}_t)_\sharp \boldsymbol{\pi}) \subset B$ for any $t \in [0,1]$. Let φ be a Kantorovich potential associated to μ_0, μ_1 and assume that φ is Lipschitz on bounded subsets of X.*

Then for every $\tilde{\boldsymbol{\pi}} \in \mathrm{OptGeo}(\mu_0, \mu_1)$ it holds

$$\mathsf{d}(\gamma_0, \gamma_1) = \mathrm{lip}^+(\varphi)(\gamma_0), \qquad \tilde{\boldsymbol{\pi}} - a.e.\ \gamma. \tag{5.8}$$

proof Fix $x \in X$ and notice that for every $y \in \partial^c \varphi(x)$ it holds

$$\varphi(x) = \frac{\mathsf{d}^2(x, y)}{2} - \varphi^c(y),$$

$$\varphi(z) \leq \frac{\mathsf{d}^2(z, y)}{2} - \varphi^c(y), \qquad \forall z \in X.$$

Therefore $\varphi(z) - \varphi(x) \leq \frac{\mathsf{d}^2(z,y)}{2} - \frac{\mathsf{d}^2(x,y)}{2} \leq \mathsf{d}(x, z)\frac{\mathsf{d}(z,y)+\mathsf{d}(x,y)}{2}$, so that dividing by $\mathsf{d}(x, z)$ and letting $z \to x$, since $y \in \partial^c \varphi(x)$ is arbitrary we get

$$\mathrm{lip}^+(\varphi)(x) \leq \inf_{y \in \partial^c \varphi(x)} \mathsf{d}(x, y).$$

In particular, if $\tilde{\boldsymbol{\pi}}$ is an arbitrary plan in $\mathrm{OptGeo}(\mu_0, \mu_1)$, from the above inequality and the fact that $\gamma_1 \in \partial^c \varphi(\gamma_0)$ for $\tilde{\boldsymbol{\pi}}$-a.e. γ we deduce

$$\mathrm{lip}^+(\varphi)(\gamma_0) \leq \mathsf{d}(\gamma_0, \gamma_1), \qquad \tilde{\boldsymbol{\pi}} - a.e.\ \gamma.$$

Thus to conclude it is sufficient to show that

$$\int \mathrm{lip}^+(\varphi)^2 \, \mathrm{d}\mu_0 \geq W_2^2(\mu_0, \mu_1).$$

To this aim, let $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu_0, \mu_1)$ be as in the hypothesis, and use again the fact that φ is a Kantorovich potential to get that for any $t \in (0, 1]$ it holds

$$\varphi(\gamma_0) - \varphi(\gamma_t) \geq \frac{\mathsf{d}^2(\gamma_0, \gamma_1)}{2} - \frac{\mathsf{d}^2(\gamma_t, \gamma_1)}{2} = \mathsf{d}^2(\gamma_0, \gamma_1)(t - t^2/2), \qquad \boldsymbol{\pi} - a.e.\ \gamma. \tag{5.9}$$

Dividing by $\mathsf{d}(\gamma_0, \gamma_t) = t\mathsf{d}(\gamma_0, \gamma_1)$, squaring and integrating w.r.t. $\boldsymbol{\pi}$ we get

$$\varlimsup_{t\downarrow 0} \int \left(\frac{\varphi(\gamma_0) - \varphi(\gamma_t)}{\mathsf{d}(\gamma_0, \gamma_t)} \right)^2 \mathrm{d}\boldsymbol{\pi}(\gamma) \geq \int \mathsf{d}^2(\gamma_0, \gamma_1) \, \mathrm{d}\boldsymbol{\pi}(\gamma) = W_2^2(\mu_0, \mu_1).$$

Now notice that since φ is locally Lipschitz, $\mathrm{lip}^+(\varphi)$ is an upper gradient for φ, hence it holds

$$\int \left(\frac{\varphi(\gamma_0) - \varphi(\gamma_t)}{\mathsf{d}(\gamma_0, \gamma_t)} \right)^2 \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \int \frac{1}{t^2} \left(\int_0^t \mathrm{lip}^+(\varphi)(\gamma_s) \, \mathrm{d}s \right)^2 \mathrm{d}\boldsymbol{\pi}(\gamma)$$

$$\leq \frac{1}{t} \iint_0^t \mathrm{lip}^+(\varphi)^2(\gamma_s) \, \mathrm{d}s \, \mathrm{d}\boldsymbol{\pi}(\gamma) = \frac{1}{t} \int_0^t \int \mathrm{lip}^+(\varphi)^2 \, \mathrm{d}(\mathrm{e}_s)_\sharp \boldsymbol{\pi} \, \mathrm{d}s.$$

From our assumptions and (5.7) we get that $\mathcal{U}_N\left(\frac{1}{t} \int_0^t (\mathrm{e}_s)_\sharp \boldsymbol{\pi} \, \mathrm{d}s \right) \to \mathcal{U}_N(\mu)$ as $t \downarrow 0$. Also, notice that by hypothesis the Borel function $\mathrm{lip}^+(\varphi)^2$ is bounded on B. Thus we can apply Lemma 5.4 and get

$$\lim_{t\downarrow 0} \frac{1}{t} \int_0^t \int \mathrm{lip}^+(\varphi)^2 \, \mathrm{d}(\mathrm{e}_s)_\sharp \boldsymbol{\pi} \mathrm{d}s = \int \mathrm{lip}^+(\varphi)^2 \, \mathrm{d}\mu_0,$$

and the proof is completed. $\qquad \square$

The following proposition is the variant of Proposition 2.10: notice that we are assuming (5.3) to hold, replacing Sobolev functions with Lipschitz ones and the assumption of bounded compression with $(e_0)_\sharp \boldsymbol{\pi} \ll \mathfrak{m}$ and $\mathcal{U}_N((e_t)_\sharp \boldsymbol{\pi}) \to \mathcal{U}_N((e_0)_\sharp \boldsymbol{\pi})$ as $t \downarrow 0$.

Proposition 5.6 *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be as in* (2.7). *Assume that* (5.3) *holds and let* $\boldsymbol{\pi} \in \mathscr{P}(C([0, 1], X))$ *be such that* $(e_0)_\sharp \boldsymbol{\pi} \ll \mathfrak{m}$, *and* $\mathcal{U}_N((e_t)_\sharp \boldsymbol{\pi}) \to \mathcal{U}_N((e_0)_\sharp \boldsymbol{\pi})$ *as* $t \downarrow 0$ *for some* $N \in (1, \infty)$. *Assume furthermore that for some bounded and closed set* $B \subset X$ *with* $\mathfrak{m}(B) < \infty$ *it holds* $\mathrm{supp}((e_t)_\sharp \boldsymbol{\pi}) \subset B$ *for any* $t \in [0, 1]$.

Then for every function $f : X \to \mathbb{R}$ *Lipschitz on bounded sets and every* $p \in (1, \infty)$ *it holds*

$$\varlimsup_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \frac{\||Df|_w\|_{L^p(X, (e_0)_\sharp \boldsymbol{\pi})}^p}{p} + \frac{\|\boldsymbol{\pi}\|_q^q}{q}, \tag{5.10}$$

where q *is the conjugate exponent of* p.

proof If $\|\boldsymbol{\pi}\|_q = \infty$ there is nothing to prove. Thus we can assume $\|\boldsymbol{\pi}\|_q < \infty$ which in particular implies that for some $T \in (0, 1]$ the plan $(\mathrm{restr}_0^T)_\sharp \boldsymbol{\pi}$ is concentrated on absolutely continuous curves and that $(e_t)_\sharp \boldsymbol{\pi}$ weakly converges to $(e_0)_\sharp \boldsymbol{\pi}$ in duality with $C_b(B)$.

Since f is locally Lipschitz, $\mathrm{lip}(f)$ is an upper gradient for f and thus

$$\frac{f(\gamma_t) - f(\gamma_0)}{t} \leq \frac{1}{t} \int_0^t \mathrm{lip}(f)(\gamma_s)|\dot{\gamma}_s| \, \mathrm{d}s \leq \frac{\int_0^t \mathrm{lip}(f)^p(\gamma_s) \, \mathrm{d}s}{tp} + \frac{\int_0^t |\dot{\gamma}_s|^q \, \mathrm{d}s}{tq},$$

holds for any absolutely continuous curve γ and $t \in (0, 1]$. Integrate over $\boldsymbol{\pi}$ and let $t \downarrow 0$ to get

$$\varlimsup_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \frac{1}{p} \varlimsup_{t \downarrow 0} \frac{1}{t} \iint_0^t \mathrm{lip}(f)^p(\gamma_s) \, \mathrm{d}s \, \mathrm{d}\boldsymbol{\pi}(\gamma) + \frac{\|\boldsymbol{\pi}\|_q^q}{q}.$$

From our assumptions on $\boldsymbol{\pi}$ and (5.7) we get that $\mathcal{U}_N\left(\frac{1}{t} \int_0^t (e_s)_\sharp \boldsymbol{\pi} \, \mathrm{d}s\right) \to \mathcal{U}_N((e_0)_\sharp \boldsymbol{\pi})$. Since $(e_0)_\sharp \boldsymbol{\pi} \ll \mathfrak{m}$ and everything is taking place in the closed and bounded set B, we can apply Lemma 5.4 to the bounded Borel function $\mathrm{lip}(f)^p$ to get that

$$\varlimsup_{t \downarrow 0} \frac{1}{t} \iint_0^t \mathrm{lip}(f)^p(\gamma_s) \, \mathrm{d}s \, \mathrm{d}\boldsymbol{\pi}(\gamma) = \varlimsup_{t \downarrow 0} \int \mathrm{lip}(f)^p \, \mathrm{d}\left(\frac{1}{t} \int_0^t (e_s)_\sharp \boldsymbol{\pi} \, \mathrm{d}s\right) = \int \mathrm{lip}(f)^p \, \mathrm{d}(e_0)_\sharp \boldsymbol{\pi}.$$

Finally, we use (5.3) to replace $\mathrm{lip}(f)$ with $|Df|_w$. $\qquad\qquad\square$

This proposition suggest the following variant of Definition 3.7:

Definition 5.7 (Plans weakly representing gradients) *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be as in* (2.7), $p, q \in (1, \infty)$ *conjugate exponents,* $g : X \to \mathbb{R}$ *a function Lipschitz on bounded sets and* $\boldsymbol{\pi} \in \mathscr{P}(C([0, 1], X))$. *We say that* $\boldsymbol{\pi}$ *weakly* q-*represents* ∇g *provided:*

   i) $\|\boldsymbol{\pi}\|_q < \infty$,

   ii) $(e_0)_\sharp \boldsymbol{\pi} \ll \mathfrak{m}$ *and* $\mathcal{U}_N((e_t)_\sharp \boldsymbol{\pi}) \to \mathcal{U}_N((e_0)_\sharp \boldsymbol{\pi})$ *as* $t \downarrow 0$ *for some* $N \in (1, \infty)$,

   iii) for some bounded and closed set $B \subset X$ *with* $\mathfrak{m}(B) < \infty$ *it holds* $\mathrm{supp}((e_t)_\sharp \boldsymbol{\pi}) \subset B$ *for any* $t \in [0, 1]$,

   iv) it holds

$$\varliminf_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \geq \frac{\||Dg|_w\|_{L^p(X, (e_0)_\sharp \boldsymbol{\pi})}^p}{p} + \frac{\|\boldsymbol{\pi}\|_q^q}{q}. \tag{5.11}$$

Note: it is not the representation of ∇g to be weak in this definition, but the assumption that $\boldsymbol{\pi}$ has bounded compression that has been weakened into (ii) above.

As a direct consequence of the metric Brenier theorem 5.5 we have the following simple and important result.

Corollary 5.8 *With the same assumptions and notations of Proposition 5.5 above, assume furthermore that (5.3) holds. Then $\boldsymbol{\pi}$ weakly 2-represents $\nabla(-\varphi)$.*

proof Dividing (5.9) by t, integrating w.r.t. $\boldsymbol{\pi}$ and letting $t \downarrow 0$ we get

$$\varlimsup_{t \downarrow 0} \int \frac{\varphi(\gamma_0) - \varphi(\gamma_t)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \geq \int \mathsf{d}^2(\gamma_0, \gamma_1) \, \mathrm{d}\boldsymbol{\pi}(\gamma).$$

On the other hand, since $\boldsymbol{\pi}$ is concentrated on geodesics it holds

$$\frac{1}{t} \iint_0^t |\dot{\gamma}_s|^2 \, \mathrm{d}s \, \mathrm{d}\boldsymbol{\pi}(\gamma) = \int \mathsf{d}^2(\gamma_0, \gamma_1) \, \mathrm{d}\boldsymbol{\pi}(\gamma), \qquad \forall t \in (0, 1].$$

Finally, (5.8), (5.3) and the fact that $(\mathrm{e}_0)_\sharp \boldsymbol{\pi} \ll \mathfrak{m}$ give $\int |D\varphi|_w^2 \, \mathrm{d}(\mathrm{e}_0)_\sharp \boldsymbol{\pi} = \int \mathsf{d}^2(\gamma_0, \gamma_1) \, \mathrm{d}\boldsymbol{\pi}(\gamma)$. $\square$

We conclude with the variant of Theorem 3.10 which we will use in the proof of the Laplacian comparison estimates.

Proposition 5.9 (Variant of horizontal-vertical derivation) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) and assume that (5.3) holds. Let $p, q \in (1, \infty)$ be conjugate exponents and $N \in (1, \infty)$. Let $f, g : X \to \mathbb{R}$ be two functions Lipschitz continuous on bounded subsets of X and $\boldsymbol{\pi} \in \mathscr{P}(C([0, 1], X)$ a plan weakly q-representing ∇g.*

Then the same conclusions of Theorem 3.10 hold, i.e.

$$\int D^+ f(\nabla g) |Dg|_w^{p-2} \, \mathrm{d}(\mathrm{e}_0)_\sharp \boldsymbol{\pi} \geq \varlimsup_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma)$$

$$\geq \varliminf_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \geq \int D^- f(\nabla g) |Dg|_w^{p-2} \, \mathrm{d}(\mathrm{e}_0)_\sharp \boldsymbol{\pi}.$$
$$(5.12)$$

proof The proof is the same as the one of Theorem 3.10, with Proposition 5.6 replacing Proposition 2.10. Indeed notice that the assumptions of Proposition 5.6 are fulfilled, so that from (5.10) applied to the function $g + \varepsilon f$, $\varepsilon \in \mathbb{R}$, we get

$$\varlimsup_{t \downarrow 0} \int \frac{(g + \varepsilon f)(\gamma_t) - (g + \varepsilon f)(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \frac{\||D(g + \varepsilon f)|_w\|_{L^p(X, (\mathrm{e}_0)_\sharp \boldsymbol{\pi})}^p}{p} + \frac{\|\boldsymbol{\pi}\|_q^q}{q}, \qquad (5.13)$$

while from our assumptions on g and $\boldsymbol{\pi}$ we know that

$$\lim_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \geq \frac{\||Dg|_w\|_{L^p(X, (\mathrm{e}_0)_\sharp \boldsymbol{\pi})}^p}{p} + \frac{\|\boldsymbol{\pi}\|_q^q}{q}. \qquad (5.14)$$

Subtract (5.14) from (5.13) to get

$$\varlimsup_{t \downarrow 0} \varepsilon \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \int \frac{|D(g + \varepsilon f)|_w^p - |Dg|_w^p}{p} \, \mathrm{d}(\mathrm{e}_0)_\sharp \boldsymbol{\pi}.$$

The thesis follows. $\square$

5.3 Proof of Laplacian comparison

We are now ready to prove the main comparison result of this paper. The proof will come by combining a bound from below and a bound from above of the derivative of the internal energy along a geodesic.

We start with the bound from below, which is the one technically more involved. We will use all the machinery developed in the previous section. Let us introduce the pressure functionals $p_N : [0, \infty) \to \mathbb{R}$, by putting $p_N(z) := zu'_N(z) - u_N(z)$, i.e.

$$p_N(z) = \tfrac{1}{N} z^{1 - \frac{1}{N}}.$$

Proposition 5.10 (Bound from below on the derivative of the internal energy)
Let (X, d, m) be as in (2.7) such that (5.3) holds and m is finite on bounded sets.

Let $\mu = \mu_0 \in \mathscr{P}(X)$ be a measure with bounded support such that $\mu_0 \ll m$, say $\mu_0 = \rho m$. Assume that $\mathrm{supp}(\mu_0) = \overline{\Omega}$, with Ω open such that $m(\partial \Omega) = 0$. Assume also that the restriction of ρ to $\overline{\Omega}$ is Lipschitz and bounded from below by a positive constant. Also, let $\mu_1 \in \mathscr{P}(X)$ and $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu_0, \mu_1)$. Assume that $\mathrm{supp}((e_t)_\sharp \boldsymbol{\pi}) \subset \overline{\Omega}$ for every $t \in [0, 1]$ and that $\mathcal{U}_N((e_t)_\sharp \boldsymbol{\pi}) \to \mathcal{U}_N(\mu_0)$, as $t \downarrow 0$.

Then it holds

$$\varliminf_{t \downarrow 0} \frac{\mathcal{U}_N((e_t)_\sharp \boldsymbol{\pi}) - \mathcal{U}_N((e_0)_\sharp \boldsymbol{\pi})}{t} \geq -\int_\Omega D^+(p_N(\rho))(\nabla \varphi) \, dm, \tag{5.15}$$

where φ is any Kantorovich potential from μ_0 to μ_1 which is Lipschitz on bounded sets and the integrand $D^+(p_N(\rho))(\nabla \varphi)$ is defined according to Definition 3.1 in the space $(\overline{\Omega}, d, m_{|\Omega})$ (this remark is necessary because in general ρ, being discontinuous along $\partial \Omega$, does not belong to any $S^p(X, d, m)$).

proof We work on the space $(\overline{\Omega}, d, m_{|\Omega})$. Since $\rho, \rho^{-1} : \overline{\Omega} \to \mathbb{R}$ are bounded, the function $u'_N(\rho) : \overline{\Omega} \to \mathbb{R}$ is bounded. Thus for every $\nu \in \mathscr{P}(\overline{\Omega})$ absolutely continuous w.r.t. $m_{|\Omega}$, the convexity of u_N gives $\mathcal{U}_N(\nu) - \mathcal{U}_N(\mu) \geq \int_\Omega u'_N(\rho)(\frac{d\nu}{dm} - \rho) \, dm$. Then a simple approximation argument based on the continuity of ρ gives

$$\mathcal{U}_N(\nu) - \mathcal{U}_N(\mu) \geq \int_{\overline{\Omega}} u'_N(\rho) \, d\nu - \int_{\overline{\Omega}} u'_N(\rho) \, d\mu, \qquad \forall \nu \in \mathscr{P}(\overline{\Omega}).$$

Plugging $\nu := (e_t)_\sharp \boldsymbol{\pi}$, dividing by t and letting $t \downarrow 0$ we get

$$\varliminf_{t \downarrow 0} \frac{\mathcal{U}_N((e_t)_\sharp \boldsymbol{\pi}) - \mathcal{U}_N((e_0)_\sharp \boldsymbol{\pi})}{t} \geq \varliminf_{t \downarrow 0} \int \frac{u'_N(\rho) \circ e_t - u'_N(\rho) \circ e_0}{t} \, d\boldsymbol{\pi}.$$

Now observe that for any Lipschitz function $f : X \to \mathbb{R}$ and $x \in \Omega$, the values of $\mathrm{lip}(f)(x), \mathrm{lip}^\pm(f)(x)$ calculated in (X, d) and $(\overline{\Omega}, d)$ are the same. Taking into account Proposition 2.6 and the fact that $m(\partial \Omega) = 0$, we can conclude that (5.3) holds also in the space $(\overline{\Omega}, d, m_{|\Omega})$.

In particular, Corollary 5.8 is applicable and we have that $\boldsymbol{\pi}$ weakly 2-represents $\nabla(-\varphi)$. Then, since the maps $u'_N(\rho), \varphi : \overline{\Omega} \to \mathbb{R}$ are Lipschitz, we can apply Proposition 5.9 on the space $(\overline{\Omega}, d, m_{|\Omega})$ with $p = q = 2$ to get

$$\varliminf_{t \downarrow 0} \int \frac{u'_N(\rho) \circ e_t - u'_N(\rho) \circ e_0}{t} \, d\boldsymbol{\pi} \geq \int_\Omega D^-(u'_N(\rho))(\nabla(-\varphi)) \rho \, dm.$$

To conclude, notice that $\mathrm{p}'_N(z) = zu''_N(z)$ and apply twice the chain rule (3.29):

$$\int_\Omega D^-(u'_N(\rho))(\nabla(-\varphi))\rho\,\mathrm{d}\mathfrak{m} = -\int_\Omega \rho u''_N(\rho)D^+\rho(\nabla\varphi)\,\mathrm{d}\mathfrak{m} = -\int_\Omega D^+(\mathrm{p}_N(\rho))(\nabla\varphi)\,\mathrm{d}\mathfrak{m}.$$

$\square$

For $K \in \mathbb{R}$, $N \in (1,\infty)$, we introduce the functions $\tilde{\tau}_{K,N} : [0,\infty) \to \mathbb{R}$ by

$$\tilde{\tau}_{K,N}(\theta) := \begin{cases} \dfrac{1}{N}\left(1 + \theta\sqrt{K(N-1)}\cotan\left(\theta\sqrt{\frac{K}{N-1}}\right)\right), & \text{if } K > 0, \\[3mm] 1, & \text{if } K = 0, \\[3mm] \dfrac{1}{N}\left(1 + \theta\sqrt{-K(N-1)}\cotanh\left(\theta\sqrt{\frac{-K}{N-1}}\right)\right), & \text{if } K < 0, \end{cases}$$

so that $\tilde{\tau}_{K,N}(\theta) = \lim_{t\downarrow 0}\frac{\tau_{K,N}^{(1)}(\theta)-\tau_{K,N}^{(1-t)}(\theta)}{t}$ (provided $K\theta^2 < (N-1)\pi^2$ if $K > 0$). Notice that if $K \le 0$ then $\tilde{\tau}_{K,N}(\cdot)$ is bounded on bounded subsets of $[0,\infty)$, but for $K > 0$ it blows up as θ approaches $\pi\sqrt{(N-1)/K}$. This creates a potential complication about the integrability of $\tilde{\tau}_{K,N}(\mathsf{d}(\gamma_0,\gamma_1))$ w.r.t. optimal geodesic plans $\boldsymbol{\pi}$ which must be dealt with. This is the scope of the following lemma.

Lemma 5.11 *Let $(X,\mathsf{d},\mathfrak{m})$ be a $\mathsf{CD}(K,N)$ space, $K \in \mathbb{R}$, $N \in (1,\infty)$, and $\mu,\nu \in \mathscr{P}(X)$ with bounded supports and $\mu \le C\mathfrak{m}$ for some $C > 0$. Then for every $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu,\nu)$ the function $\gamma \mapsto \tilde{\tau}_{K,N}(\mathsf{d}(\gamma_0,\gamma_1))$ is in $L^1(\boldsymbol{\pi})$ and the convergence of $\frac{\tau_{K,N}^{(1)}(\mathsf{d}(\gamma_0,\gamma_1))-\tau_{K,N}^{(1-t)}(\mathsf{d}(\gamma_0,\gamma_1))}{t}$ to $\tilde{\tau}_{K,N}(\mathsf{d}(\gamma_0,\gamma_1))$ as $t \to 0$ is dominated in $L^1(\boldsymbol{\pi})$.*

proof If $K \le 0$ there is nothing to prove, thus let $K > 0$ and recall that in this case (X,d) is compact (Proposition 5.2). It is immediate to verify that the domination of $\left|\frac{\tau_{K,N}^{(1)}(\mathsf{d}(\gamma_0,\gamma_1))-\tau_{K,N}^{(1-t)}(\mathsf{d}(\gamma_0,\gamma_1))}{t}\right|$ is a consequence of the fact that $\gamma \mapsto \tilde{\tau}_{K,N}(\mathsf{d}(\gamma_0,\gamma_1))$ is in $L^1(\boldsymbol{\pi})$, thus we focus on this second problem.

If $\sup_{\gamma\in\mathrm{supp}(\boldsymbol{\pi})}\mathsf{d}(\gamma_0,\gamma_1) < \pi\sqrt{(N-1)/K}$ there is nothing to prove. Otherwise, taking into account the compactness of (X,d), there exists $\bar{\gamma} \in \mathrm{supp}(\boldsymbol{\pi})$ such that $\mathsf{d}(\gamma_0,\gamma_1) = \pi\sqrt{(N-1)/K} =: \mathsf{D}$. Put $x_E := \gamma_0$, $x_O := \gamma_1$. We claim[5] that for any $x \in X$ it holds

$$\mathsf{d}(x,x_E) + \mathsf{d}(x,x_O) = \mathsf{D}. \tag{5.16}$$

If not, for some $r \in (0,\mathsf{D})$ the two disjoint balls $B_r(x_E)$, $B_{\mathsf{D}-r}(x_O)$ would not cover some neighborhood of x, so that

$$\mathfrak{m}(B_r(x_E)) + \mathfrak{m}(B_{\mathsf{D}-r}(x_O)) = \mathfrak{m}\big(B_r(x_E) \cup B_{\mathsf{D}-r}(x_O)\big) < \mathfrak{m}(X). \tag{5.17}$$

However, taking into account the Bishop-Gromov volume comparison estimates (inequality (5.2)) we have

$$\mathfrak{m}(B_r(x_E)) + \mathfrak{m}(B_{\mathsf{D}-r}(x_O)) \ge \left(\frac{\int_0^r \sin(t\pi/\mathsf{D})^{N-1}\,\mathrm{d}t}{\int_0^{\mathsf{D}} \sin(t\pi/\mathsf{D})^{N-1}\,\mathrm{d}t} + \frac{\int_0^{\mathsf{D}-r} \sin(t\pi/\mathsf{D})^{N-1}\,\mathrm{d}t}{\int_0^{\mathsf{D}} \sin(t\pi/\mathsf{D})^{N-1}\,\mathrm{d}t}\right)\mathfrak{m}(X) = \mathfrak{m}(X),$$

---

[5]this claim has been proved by Ohta in [31], we are reporting here the proof for completeness

which contradicts (5.17). Hence our claim is true. The argument also yields

$$\mathfrak{m}(B_r(x_E)) = \frac{\int_0^r \sin(t\pi/\mathsf{D})^{N-1}\,\mathrm{d}t}{\int_0^{\mathsf{D}} \sin(t\pi/\mathsf{D})^{N-1}\,\mathrm{d}t}\mathfrak{m}(X), \qquad \forall r \in [0, \mathsf{D}] \tag{5.18}$$

Now we claim that for every $\gamma \in \mathrm{supp}(\boldsymbol{\pi})$ it holds either $\gamma_0 = x_E$ or $\gamma_1 = x_O$. Indeed, by cyclical monotonicity and (5.16) we have

$$\begin{aligned}
\mathsf{d}^2(\gamma_0, \gamma_1) + \mathsf{D}^2 &\leq \mathsf{d}^2(\gamma_0, x_O) + \mathsf{d}^2(x_E, \gamma_1) \\
&= \mathsf{D}^2 + \mathsf{d}^2(\gamma_0, x_E) - 2\mathsf{D}\mathsf{d}(\gamma_0, x_E) + \mathsf{d}^2(x_E, \gamma_1) \\
&\leq \mathsf{D}^2 + \mathsf{d}^2(\gamma_0, x_E) - 2\mathsf{d}(x_E, \gamma_1)\mathsf{d}(\gamma_0, x_E) + \mathsf{d}^2(x_E, \gamma_1) \\
&= \mathsf{D}^2 + \big(\mathsf{d}(x_E, \gamma_1) - \mathsf{d}(\gamma_0, x_E)\big)^2,
\end{aligned} \tag{5.19}$$

which yields $\mathsf{d}(\gamma_0, \gamma_1) \leq \big|\mathsf{d}(x_E, \gamma_1) - \mathsf{d}(\gamma_0, x_E)\big|$. Hence the triangle inequality forces the inequalities in (5.19) to be equalities. The claim follows observing that the second inequality in (5.19) is an equality if and only if either $\mathsf{d}(\gamma_0, x_E) = 0$ or $\mathsf{d}(\gamma_1, x_E) = \mathsf{D}$ and in this latter case, (5.16) yields $\mathsf{d}(\gamma_1, x_O) = 0$.

By assumption, $(\mathrm{e}_0)_\sharp \boldsymbol{\pi} \ll \mathfrak{m}$, so that, since $\mathfrak{m}$ has no atoms, for $\boldsymbol{\pi}$-a.e. γ it holds $\gamma_0 \neq x_E$ and what we just proved yields $\gamma_1 = x_O$ for $\boldsymbol{\pi}$-a.e. γ. Thus

$$\tilde{\tau}_{K,N}(\mathsf{d}(\gamma_0, \gamma_1)) = \tilde{\tau}_{K,N}(\mathsf{d}(\gamma_0, x_O)), \qquad \boldsymbol{\pi} - a.e.\ \gamma,$$

and, taking the assumption $\mu \leq C\mathfrak{m}$ and the defining formula of $\tilde{\tau}_{K,N}(\cdot)$ in mind, our thesis reduces to prove that

$$\int_{B_{\mathsf{D}/2}(x_E)} \frac{1}{\sin\left(\mathsf{d}(x, x_O)\sqrt{\frac{K}{N-1}}\right)}\,\mathrm{d}\mathfrak{m}(x) < \infty.$$

Let $T : X \to [0, \mathsf{D}]$ be given by $T(x) := \mathsf{d}(x, x_E)$ and observe that (5.18) gives

$$\mathrm{d}T_\sharp\mathfrak{m}(r) = \left(\frac{\sin(r\pi/\mathsf{D})^{N-1}}{\int_0^{\mathsf{D}} \sin(t\pi/\mathsf{D})^{N-1}\,\mathrm{d}t}\mathfrak{m}(X)\right)\mathrm{d}\mathcal{L}^1|_{[0,\mathsf{D}]}(r).$$

Therefore

$$\int_{B_{\mathsf{D}/2}(x_E)} \frac{1}{\sin\left(\mathsf{d}(x, x_O)\sqrt{\frac{K}{N-1}}\right)}\,\mathrm{d}\mathfrak{m}(x) = \int_{B_{\mathsf{D}/2}(x_E)} \frac{1}{\sin\left(\mathsf{d}(x, x_E)\frac{\pi}{\mathsf{D}}\right)}\,\mathrm{d}\mathfrak{m}(x)$$

$$= \frac{\mathfrak{m}(X)}{\int_0^{\mathsf{D}} \sin(t\pi/\mathsf{D})^{N-1}\,\mathrm{d}t}\int_0^{\mathsf{D}/2} \sin\left(r\frac{\pi}{\mathsf{D}}\right)^{N-2}\,\mathrm{d}r.$$

The conclusion follows from the fact that $N > 1$. $\qquad\qquad\square$

Proposition 5.12 (Bound from above on the derivative of the internal energy)
Let $(X, \mathsf{d}, \mathfrak{m})$ *be as in* (2.7), $K \in \mathbb{R}$ *and* $N \in (1, \infty)$. *Let* $\boldsymbol{\pi} \in \mathscr{P}(\mathrm{Geo}(X))$ *be such that* $(\mathrm{e}_i)_\sharp \boldsymbol{\pi}$ *has bounded support,* $i = 0, 1$, *and* $(\mathrm{e}_0)_\sharp \boldsymbol{\pi} \ll \mathfrak{m}$ *with bounded density. Assume also that* $\boldsymbol{\pi}$ *satisfies*

$$\mathcal{U}_N((\mathrm{e}_t)_\sharp\boldsymbol{\pi}) \leq -\int \tau_{K,N}^{(1-t)}\big(\mathsf{d}(\gamma_0, \gamma_1)\big)\rho^{-\frac{1}{N}}(\gamma_0)\,\mathrm{d}\boldsymbol{\pi}(\gamma), \qquad \forall t \in [0, 1], \tag{5.20}$$

where ρ is the density of $(\mathrm{e}_0)_\sharp \boldsymbol{\pi}$.

Then

$$\overline{\lim_{t \downarrow 0}} \frac{\mathcal{U}_N((\mathrm{e}_t)_\sharp \boldsymbol{\pi}) - \mathcal{U}_N((\mathrm{e}_0)_\sharp \boldsymbol{\pi})}{t} \leq \int \tilde{\tau}_{K,N}\big(\mathsf{d}(\gamma_0, \gamma_1)\big) \rho^{-\frac{1}{N}}(\gamma_0) \, \mathrm{d}\boldsymbol{\pi}(\gamma). \qquad (5.21)$$

proof From $\mathcal{U}_N((\mathrm{e}_0)_\sharp \boldsymbol{\pi}) = -\int \rho^{-\frac{1}{N}}(\gamma_0) \, \mathrm{d}\boldsymbol{\pi}(\gamma)$, and (5.20) we get

$$\frac{\mathcal{U}_N((\mathrm{e}_t)_\sharp \boldsymbol{\pi}) - \mathcal{U}_N((\mathrm{e}_0)_\sharp \boldsymbol{\pi})}{t} \leq \int \rho^{-\frac{1}{N}} \left(\frac{1 - \tau_{K,N}^{(1-t)}\big(\mathsf{d}(\gamma_0, \gamma_1)\big)}{t} \right) \mathrm{d}\boldsymbol{\pi}(\gamma),$$

and the conclusion follows by letting $t \downarrow 0$ and using the dominate convergence theorem and Lemma 5.11 to pass the limit inside the integral. $\qquad \square$

The combination of (5.15) and (5.21) will produce the Laplacian comparison estimate. Before concluding, however, we need a result which grants that Proposition 5.10 can be applied in a sufficiently large class of situations, the tricky part being to ensure that $\mathrm{supp}((\mathrm{e}_t)_\sharp \boldsymbol{\pi}) \subset \overline{\Omega}$ for any $t \in [0,1]$. This is the scope of the following lemma.

Lemma 5.13 *Let $(X, \mathsf{d}, \mathfrak{m})$ be a proper geodesic metric space with $\mathfrak{m}$ locally finite. Let $\varphi : X \to \mathbb{R}$ a locally Lipschitz c-concave function and $B \subset X$ a compact set. Then there exists another locally Lipschitz c-concave function $\tilde{\varphi} : X \to \mathbb{R}$ and a bounded open set $\Omega \supset B$ such that the following are true.*

i) $\tilde{\varphi} = \varphi$ on B.

ii) For any $x \in X$, the set $\partial^c \tilde{\varphi}(x)$ is non-empty.

iii) For any $x \in \Omega$, $y \in \partial^c \tilde{\varphi}(x)$ and $\gamma \in \mathrm{Geo}(X)$ connecting x to y it holds $\gamma_t \in \Omega$ for any $t \in [0,1]$.

iv) $\mathfrak{m}(\partial \Omega) = 0$.

proof By Proposition 2.2 we know that for any $x \in X$ the set $\partial^c \varphi(x)$ is non-empty and that the set $C \subset X$ defined by

$$C := \Big\{ y \in X \ : \ y \in \partial^c \varphi(x), \ \text{ for some } x \in B \Big\},$$

is compact. Define $\tilde{\varphi} : X \to \mathbb{R}$ by

$$\tilde{\varphi}(x) := \inf_{y \in C} \frac{\mathsf{d}^2(x,y)}{2} - \varphi^c(y).$$

By definition, $\tilde{\varphi}$ is c-concave, locally Lipschitz and $\tilde{\varphi} = \varphi$ on B. By Proposition 2.2 again we get that (ii) is fulfilled as well. Let $a := \sup_{x \in B} \tilde{\varphi}(x)$ and for $b > a$ define

$$\Omega_b := \{ x \ : \ \tilde{\varphi}(x) < b \}.$$

Clearly Ω_b is an open set which contains B. We claim that it is also bounded. Indeed since φ^c is c-concave and C is bounded, it holds $\sup_{y \in C} \varphi^c(y) < \infty$, so that the boundedness of

Ω_b follows from the inequality $\tilde{\varphi}(x) \geq \frac{\mathsf{d}^2(x,C)}{2} - \sup_C \varphi^c$. Now pick an arbitrary $x \in \Omega_b$ and $y \in \partial^c \tilde{\varphi}(x)$. From

$$\tilde{\varphi}(x) = \frac{\mathsf{d}^2(x,y)}{2} - \tilde{\varphi}^c(y),$$

$$\tilde{\varphi}(z) \leq \frac{\mathsf{d}^2(z,y)}{2} - \tilde{\varphi}^c(y), \qquad \forall z \in X,$$

we get that for any $z \in X$ such that $\mathsf{d}(z,y) \leq \mathsf{d}(x,y)$ it holds $\tilde{\varphi}(z) \leq \tilde{\varphi}(x)$. In particular, this applies to $z := \gamma_t$, with $\gamma \in \mathrm{Geo}(X)$ connecting x to y, thus (iii) is proved as well.

It remains to show that b can be chosen so that $\mathfrak{m}(\partial\Omega_b) = 0$. To achieve this, notice that for $b \neq b'$ it holds $\partial\Omega_b \cap \partial\Omega_{b'} = \emptyset$ so that the σ-finiteness of $\mathfrak{m}$ yields that for any $b > a$ except at most a countable number it holds $\mathfrak{m}(\partial\Omega_b) = 0$. $\qquad\square$

Notice that on $\mathsf{CD}(K,N)$ spaces, if φ is a locally Lipschitz c-concave function, then Proposition 5.5, (5.3) and Lemma 5.11 give that $\tilde{\tau}_{K,N}(|D\varphi|_w) \in L^1_{\mathrm{loc}}(X,\mathfrak{m})$.

In the following theorem and the discussion thereafter, we say that $(X,\mathsf{d},\mathfrak{m})$ is infinitesimally strictly convex provided for some $q \in (1,\infty)$ it is q-infinitesimally strictly convex. In this case, since locally Lipschitz functions belong to $\mathrm{S}^p_{\mathrm{loc}}(X,\mathsf{d},\mathfrak{m})$ for any $p \in (1,\infty)$, for any f,g locally Lipschitz it holds

$$D^+ f(\nabla g) = D^- f(\nabla g), \qquad \mathfrak{m} - a.e.,$$

and we will denote this common quantity, as usual, by $Df(\nabla g)$. In particular, on these spaces for g locally Lipschitz and in $D(\boldsymbol{\Delta})$, the set $\boldsymbol{\Delta}g$ contains only one measure, which abusing a bit the notation we will indicate as $\boldsymbol{\Delta}g$.

Theorem 5.14 (Comparison estimates) *Let* $K \in \mathbb{R}$, $N \in (1,\infty)$ *and* $(X,\mathsf{d},\mathfrak{m})$ *be a* $\mathsf{CD}(K,N)$ *space. Assume that* $(X,\mathsf{d},\mathfrak{m})$ *is infinitesimally strictly convex. Let* $\varphi : X \to \mathbb{R}$ *a locally Lipschitz c-concave function. Then*

$$\varphi \in D(\boldsymbol{\Delta}) \qquad and \qquad \boldsymbol{\Delta}\varphi \leq N \, \tilde{\tau}_{K,N}(|D\varphi|_w) \, \mathfrak{m}. \qquad (5.22)$$

proof Fix a compact set $B \subset X$ and use Lemma 5.13 to find a c-concave function $\tilde{\varphi}$ and a bounded open set Ω satisfying $(i),(ii),(iii),(iv)$ of the statement.

Consider the multivalued map $\overline{\Omega} \ni x \mapsto T(x) \subset \mathscr{P}(\overline{\Omega})$ defined by: $\mu \in T(x)$ if and only if $\mathrm{supp}(\mu) \subset \partial^c \tilde{\varphi}(x)$. Notice that (ii) of Lemma 5.13 ensures that $T(x) \neq \emptyset$ for any $x \in \overline{\Omega}$. Recall that $\mathscr{P}(\overline{\Omega})$ endowed with the weak topology given by the duality with $C(\overline{\Omega})$ is Polish and observe that since $\partial^c \tilde{\varphi}$ is closed, we easily get that the graph of T is a closed subset of $\overline{\Omega} \times \mathscr{P}(\overline{\Omega})$. Hence standard measurable selection arguments (see e.g. Theorem 6.9.3. in [11]) ensure that there exists a Borel map $x \mapsto \eta_x$ such that $\eta_x \in T(x)$ for any $x \in \overline{\Omega}$.

Fix a bounded function $\rho : X \to [0,\infty)$ such that $\int \rho \,\mathrm{d}\mathfrak{m} = 1$, $\rho \equiv 0$ on $X \setminus \overline{\Omega}$ and $\rho|_{\overline{\Omega}}$ is Lipschitz and bounded from below by a positive constant. Define $\mu,\nu \in \mathscr{P}(\overline{\Omega})$ by $\mu := \rho\mathfrak{m}$ and $\nu := \int \eta_x \,\mathrm{d}\mu(x)$. Since $(X,\mathsf{d},\mathfrak{m})$ is a $\mathsf{CD}(K,N)$ space we know that there exists $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu,\nu)$ such that (5.1) holds. In particular

$$\mathcal{U}_N((\mathrm{e}_t)_\sharp\boldsymbol{\pi}) \leq - \int \tau_{K,N}^{(1-t)}\big(\mathsf{d}(\gamma_0,\gamma_1)\big)\rho^{-\frac{1}{N}}(\gamma_0) \,\mathrm{d}\boldsymbol{\pi}(\gamma), \qquad \forall t \in [0,1]. \qquad (5.23)$$

Proposition 5.12 grants

$$\varlimsup_{t\downarrow 0} \frac{\mathcal{U}_N((\mathrm{e}_t)_\sharp\boldsymbol{\pi}) - \mathcal{U}_N(\mu)}{t} \leq \int \rho^{-\frac{1}{N}}(\gamma_0)\tilde{\tau}_{K,N}\big(\mathsf{d}(\gamma_0,\gamma_1)\big) \,\mathrm{d}\boldsymbol{\pi}(\gamma). \qquad (5.24)$$

By construction, $\tilde\varphi$ is a Kantorovich potential from μ to ν, and $\boldsymbol\pi \in \mathrm{OptGeo}(\mu, \nu)$. Therefore for any $\gamma \in \mathrm{supp}(\boldsymbol\pi)$ it holds $\gamma_1 \in \partial^c \tilde\varphi(\gamma_0)$. Using property (iii) of Lemma 5.13 we get that $\mathrm{supp}((e_t)_\sharp\boldsymbol\pi) \subset \overline\Omega$ for any $t \in [0,1]$.

Furthermore, using (5.23) and the semicontinuity of $\mathcal{U}_N$ on $\mathscr{P}(\overline\Omega)$ we get that $\mathcal{U}_N((e_t)_\sharp\boldsymbol\pi) \to \mathcal{U}_N(\mu)$ as $t \downarrow 0$, and thus Corollary 5.8 gives that $\boldsymbol\pi$ weakly 2-represents $\nabla(-\tilde\varphi)$.

Also, from (iv) of Lemma 5.13 we get $\mathfrak{m}(\partial\Omega) = 0$.

Therefore all the assumptions of Proposition 5.10 are satisfied (put μ in place of μ_0 and $\tilde\varphi$ in place of φ), and we get

$$\varliminf_{t \downarrow 0} \frac{\mathcal{U}_N((e_t)_\sharp\boldsymbol\pi) - \mathcal{U}_N((e_0)_\sharp\boldsymbol\pi)}{t} \geq -\int_\Omega D^+(\mathrm{p}_N(\rho))(\nabla\tilde\varphi)\,\mathrm{d}\mathfrak{m}. \tag{5.25}$$

From inequalities (5.24) and (5.25), the metric Brenier theorem 5.5, (5.3) and taking into account that $\mathrm{p}_N(z) = \frac{1}{N}z^{1-\frac{1}{N}}$ we deduce

$$-\frac{1}{N}\int_\Omega D^+(\rho^{1-\frac{1}{N}})(\nabla\tilde\varphi)\,\mathrm{d}\mathfrak{m} \leq \int \rho^{1-\frac{1}{N}}\,\tilde\tau_{K,N}(|D\tilde\varphi|_w)\,\mathrm{d}\mathfrak{m}. \tag{5.26}$$

Now let $f : X \to \mathbb{R}$ be a non-negative Lipschitz function such that $\mathrm{supp}(f) \subset B$ and, for $\varepsilon > 0$, define $\rho_\varepsilon : X \to [0,\infty)$ by $\rho_\varepsilon := c_\varepsilon(f + \varepsilon)^{\frac{N}{N-1}}\chi_{\overline\Omega}$, where c_ε is such that $\int \rho_\varepsilon\,\mathrm{d}\mathfrak{m} = 1$. Apply (5.26) to $\rho := \rho_\varepsilon$ to obtain

$$-\frac{1}{N}\int_\Omega D^+(f + \varepsilon)(\nabla\tilde\varphi)\,\mathrm{d}\mathfrak{m} \leq \int_\Omega (f + \varepsilon)\,\tilde\tau_{K,N}(|D\tilde\varphi|_w)\,\mathrm{d}\mathfrak{m}.$$

Since $(X, \mathsf{d}, \mathfrak{m})$ is infinitesimally strictly convex, with a cut-off argument and using (i) of Proposition 2.6 we have that $(\overline\Omega, \mathsf{d}, \mathfrak{m})$ is infinitesimally strictly convex as well. Let $\varepsilon \downarrow 0$ and recall that from (i) of Lemma 5.13 we have $\tilde\varphi \equiv \varphi$ on B to get

$$-\int_\Omega Df(\nabla\varphi)\,\mathrm{d}\mathfrak{m} \leq N\int f\,\tilde\tau_{K,N}(|D\varphi|_w)\,\mathrm{d}\mathfrak{m}. \tag{5.27}$$

By (ii) of Proposition 2.6 we have $\int_\Omega Df(\nabla\varphi)\,\mathrm{d}\mathfrak{m} = \int_X Df(\nabla\varphi)\,\mathrm{d}\mathfrak{m}$, in the sense that the integrand $Df(\nabla\varphi)$ can be equivalently computed in the space $(\overline\Omega, \mathsf{d}, \mathfrak{m}_{|\Omega})$ or in $(X, \mathsf{d}, \mathfrak{m})$ (actually, for φ in general it doesn't hold $\mathrm{supp}(\varphi) \subset \Omega$, but thanks to the locality property (3.8) and a cut-off argument, the claim is still true).

In summary, we proved that for any compact set $B \subset X$ and any $f : X \to [0,\infty)$ Lipschitz, non-negative and with $\mathrm{supp}(f) \subset B$ it holds

$$-\int Df(\nabla\varphi)\,\mathrm{d}\mathfrak{m} \leq N\int f\,\tilde\tau_{K,N}(|D\varphi|_w)\,\mathrm{d}\mathfrak{m}. \tag{5.28}$$

Since B is arbitrary, we just proved (5.28) for any non-negative $f \in \mathrm{Test}(X)$. The conclusion comes from Proposition 4.13. $\qquad\square$

Corollary 5.15 (Laplacian of the distance) *Let $K \in \mathbb{R}$, $N \in (1,\infty)$ and $(X, \mathsf{d}, \mathfrak{m})$ an infinitesimally strictly convex $\mathsf{CD}(K,N)$ space. For $x_0 \in X$ denote by $\mathsf{d}_{x_0} : X \to [0,\infty)$ the function $x \mapsto \mathsf{d}(x, x_0)$. Then*

$$\frac{\mathsf{d}_{x_0}^2}{2} \in D(\boldsymbol\Delta), \qquad with \qquad \boldsymbol\Delta\frac{\mathsf{d}_{x_0}^2}{2} \leq N\,\tilde\tau_{K,N}(\mathsf{d}_{x_0})\,\mathfrak{m} \qquad \forall x_0 \in X, \tag{5.29}$$

and

$$\mathsf{d}_{x_0} \in D(\boldsymbol{\Delta}, X \setminus \{x_0\}) \qquad with \qquad \boldsymbol{\Delta}\mathsf{d}_{x_0}|_{X \setminus \{x_0\}} \leq \frac{N \, \tilde{\tau}_{K,N}(\mathsf{d}_{x_0}) - 1}{\mathsf{d}_{x_0}} \, \mathfrak{m} \qquad \forall x_0 \in X. \quad (5.30)$$

proof (5.29) is a particular case of (5.22). To get (5.30) notice that since (X, d) is geodesic, it holds $\mathrm{lip}^-(\mathsf{d}_{x_0}) = 1$ on $X \setminus \{x_0\}$, so that by Theorem 5.3 we get $|D\mathsf{d}_{x_0}|_w = 1$ $\mathfrak{m}$-a.e. and from the chain rule (2.16) that $|D(\frac{\mathsf{d}_{x_0}^2}{2})|_w = \mathsf{d}_{x_0}$ $\mathfrak{m}$-a.e.. The conclusion comes from the chain rule of Proposition 4.11 with $\varphi(z) := \sqrt{2z}$. $\qquad\square$

Notice that in proving the Laplacian comparison estimates we showed that

$$(5.22) \qquad \Rightarrow \qquad (5.29)$$

In a smooth situation, also the converse implication holds. It is not so clear if the same holds in the abstract framework, because we didn't prove a property like

$$g_i \in D(\boldsymbol{\Delta}), \quad \boldsymbol{\Delta} g_i \leq \mu, \qquad \forall i \in I, \qquad \Rightarrow \qquad \inf_i g_i \in D(\boldsymbol{\Delta}), \quad \boldsymbol{\Delta}(\inf_i g_i) \leq \mu.$$

Remark 5.16 (Laplacian comparison from the $MCP(K,N)$ condition) If one is only interested in (5.29) and not in (5.22), then the Measure Contraction Property $MCP(K,N)$ (in conjunction with infinitesimal strict convexity) is sufficient to conclude. The proof is exactly the same: just interpolate from an arbitrary measure to δ_{x_0} using the Markov kernel given by the definition of $MCP(K,N)$. $\qquad\blacksquare$

Remark 5.17 (Laplacian comparison from the $\mathsf{CD}^*(K,N)$ condition) In [8] Bacher and Sturm introduced a variant of the $\mathsf{CD}(K,N)$ condition, called reduced curvature dimension condition and denoted by $\mathsf{CD}^*(K,N)$, where the distortion coefficients $\tau_{K,N}^{(t)}(\theta)$ are replaced by the coefficients $\sigma_{K,N}^{(t)}(\theta)$ defined by

$$\sigma_{K,N}^{(t)}(\theta) := \begin{cases} +\infty, & \text{if } K\theta^2 \geq N\pi^2, \\ \frac{\sin(t\theta\sqrt{K/N})}{\sin(\theta\sqrt{K/N})} & \text{if } 0 < K\theta^2 < N\pi^2, \\ t & \text{if } K\theta^2 = 0, \\ \frac{\sinh(t\theta\sqrt{K/N})}{\sinh(\theta\sqrt{K/N})} & \text{if } K\theta^2 < 0. \end{cases}$$

Notice that $\tau_{K,N}^{(t)}(\theta) = t^{\frac{1}{N}}(\sigma_{K,N-1}^{(t)}(\theta))^{1-\frac{1}{N}}$. The motivation for the introduction of this notion comes from the study of the local to global property. Indeed, while the question on whether the local validity of the $\mathsf{CD}(K,N)$ condition implies its global validity is still open, the same question for the $\mathsf{CD}^*(K,N)$ condition has affirmative answer, at least for non-branching spaces. Also, the $\mathsf{CD}^*(K,N)$ condition is stable w.r.t. mGH convergence and it is locally equivalent to the $\mathsf{CD}(K,N)$ condition (all this has been proved in [8]). The drawback of the $\mathsf{CD}^*(K,N)$ condition is that the geometric and functional inequalities which are its consequence sometime occur with non-sharp constants (although some positive recent development in this direction can be found in [12]).

Now, if we introduce the functions $\tilde{\sigma}_{K,N} : [0, \infty) \to \mathbb{R}$ by

$$\tilde{\sigma}_{K,N}(\theta) := \begin{cases} \theta\sqrt{\frac{K}{N}} \cotan\left(\theta\sqrt{\frac{K}{N}}\right), & \text{if } K > 0, \\[2mm] 1, & \text{if } K = 0, \\[2mm] \theta\sqrt{\frac{K}{N}} \cotanh\left(\theta\sqrt{\frac{K}{N}}\right), & \text{if } K < 0, \end{cases}$$

and we repeat the arguments of Theorem 5.14, we get that on infinitesimally strictly convex $\mathsf{CD}^*(K, N)$ spaces, locally Lipschitz c-concave functions φ are in $D(\boldsymbol{\Delta})$ and it holds

$$\boldsymbol{\Delta}\varphi \leq N\,\tilde{\sigma}_{K,N}(|D\varphi|_w)\,\mathfrak{m}.$$

∎

As a first step towards the proof of the Cheeger-Gromoll splitting theorem in this abstract framework, we show that the Laplacian comparison estimates for the distance function imply that the Busemann function associated to a geodesic ray on $\mathsf{CD}(0, N)$ spaces is subharmonic.

Recall that a map $\gamma : [0, \infty) \to X$ is called geodesic ray provided for any $T > 0$ its restriction to $[0, T]$ is a minimal geodesic. We will always assume that geodesic rays are parametrized by unit speed. Given a geodesic ray, the Busemann function b associated to it is defined by

$$b(x) := \lim_{t \to +\infty} b_t(x), \qquad \text{where} \qquad b_t(x) := t - \mathsf{d}(x, \gamma_t).$$

Notice that this is indeed a good definition, because the triangle inequality ensures that for any $x \in X$ the map $t \mapsto b_t(x)$ is non-decreasing (and the inequality $\mathsf{d}(x, \gamma_t) \geq \mathsf{d}(\gamma_0, \gamma_t) - \mathsf{d}(x, \gamma_0)$ yields that b is finite everywhere).

Technically speaking, the proof of the subharmonicity of b is not a direct consequence of the estimate (5.30) and of the stability result of Proposition 4.15, because we don't really know if $\boldsymbol{\Delta}b_{t_n}$ weakly converges to some limit measure for some $t_n \uparrow \infty$ (we miss a bound from below, which is needed to get compactness). Instead, we will prove an a priori regularity result for the Busemann function which ensures that it is in the domain of the Laplacian. This is the content of the following purely metric lemma. It is stated and proved in geodesic spaces, but a simple approximation argument shows that the same holds on length spaces.

Lemma 5.18 (Regularity of the Busemann function) *Let (X, d) be a geodesic space, $\gamma : [0, \infty) \to X$ a geodesic ray and b the Busemann function associated to it. Then $-b$ is c-concave.*

proof We know from (2.6) that $(-b)^{cc} \geq -b$, so to conclude we only need to prove the opposite inequality. For any $x \in X$ and $t \geq 0$ let $\gamma^{t,x} : [0, \mathsf{d}(x, \gamma_t)] \to X$ be a unit speed geodesic connecting x to γ_t. Fix $x \in X$ and notice that

$$(-b)^{cc}(x) = \inf_{y \in X} \sup_{\tilde{x} \in X} \frac{\mathsf{d}^2(x, y)}{2} - \frac{\mathsf{d}^2(\tilde{x}, y)}{2} - b(\tilde{x}) \leq \sup_{\tilde{x} \in X} \frac{1}{2} - \frac{\mathsf{d}^2(\tilde{x}, \gamma_1^{t,x})}{2} - b_t(\tilde{x}), \qquad \forall t \geq 1.$$

Then observe that for any $\tilde{x} \in X$ and $t \geq 1$ it holds

$$\frac{1}{2} - \frac{\mathsf{d}^2(\tilde{x}, \gamma_1^{t,x})}{2} - b_t(\tilde{x}) = \frac{1}{2} - \frac{\mathsf{d}^2(\tilde{x}, \gamma_1^{t,x})}{2} + \mathsf{d}(\tilde{x}, \gamma_t) - t$$

$$\leq \frac{1}{2} - \frac{\mathsf{d}^2(\tilde{x}, \gamma_1^{t,x})}{2} + \mathsf{d}(\tilde{x}, \gamma_1^{t,x}) + \mathsf{d}(\gamma_1^{t,x}, \gamma_t) - t$$

$$= -\frac{1}{2} - \frac{\mathsf{d}^2(\tilde{x}, \gamma_1^{t,x})}{2} + \mathsf{d}(\tilde{x}, \gamma_1^{t,x}) + \mathsf{d}(x, \gamma_t) - t$$

$$\leq -b_t(x).$$

Thus we proved that $(-b)^{cc}(x) \leq -b_t(x)$ for any $t \geq 1$. The conclusion follows by letting $t \to \infty$. $\qquad\square$

In the next proposition we are going to assume both that $(X, \mathsf{d}, \mathfrak{m})$ is infinitesimally strictly convex and that $W^{1,2}(\Omega)$ is uniformly convex for any $\Omega \subset X$ open with $\mathfrak{m}(\partial\Omega) = 0$: these are two unrelated notions, as the former concerns strict convexity of the norm in the 'tangent bundle' while the latter is about uniform convexity of the norm in the 'cotangent bundle' (see Appendix A for further details about this wording). Notice that on infinitesimally Hilbertian spaces these assumptions are fulfilled.

Proposition 5.19 (Busemann functions are subharmonic in $\mathsf{CD}(0, N)$ spaces) *Let $(X, \mathsf{d}, \mathfrak{m})$ be an infinitesimally strictly convex $\mathsf{CD}(0, N)$ space such that $W^{1,2}(\overline{\Omega}, \mathsf{d}, \mathfrak{m})$ is uniformly convex for any $\Omega \subset X$ open such that $\mathfrak{m}(\partial\Omega) = 0$. Let γ be a geodesic ray and b the Busemann function associated to it. Then*

$$b \in D(\boldsymbol{\Delta}) \qquad and \qquad \boldsymbol{\Delta}b \geq 0.$$

proof Being the pointwise limit of 1-Lipschitz functions, b is itself 1-Lipschitz. Hence from Theorem 5.14 and Lemma 5.18 we deduce $b \in D(\boldsymbol{\Delta})$.

We claim that $\mathrm{lip}^+(b) \equiv 1$. Indeed, fix $x \in X$ and, as in the proof of Lemma 5.18, let $\gamma^{t,x} : [0, \mathsf{d}(x, \gamma_t)] \to X$ be a unit speed geodesic connecting x to γ_t. Since X is proper, for every $r > 0$ and sequence $t_n \uparrow \infty$, the sequence $n \mapsto \gamma_r^{x,t_n}$ has a subsequence, not relabeled, which converges to some $x_r \in X$ such that $\mathsf{d}(x, x_r) = r$. By construction, it holds $b_{t_n}(\gamma_r^{x,t_n}) = b_{t_n}(x) + r$, and from the fact that b_t is 1-Lipschitz we have

$$b_{t_n}(x_r) \geq b_{t_n}(\gamma_r^{x,t_n}) - \mathsf{d}(\gamma_r^{x,t_n}, x_r) = b_{t_n}(x) + r - \mathsf{d}(\gamma_r^{x,t_n}, x_r),$$

so that passing to the limit we obtain $\frac{b(x_r) - b(x)}{r} \geq 1$. Letting $r \downarrow 0$ we get $\mathrm{lip}^+(b)(x) \geq 1$, and our claim is proved.

Let $f \in \mathrm{Test}(X)$ be non-negative and $\Omega \subset X$ an open bounded set such that $\mathfrak{m}(\partial\Omega) = 0$ and $\mathrm{supp}(f) \subset \Omega$. Consider the functions $b_t|_{\overline{\Omega}}$, $b|_{\overline{\Omega}}$ as functions on the space $(\overline{\Omega}, \mathsf{d}, \mathfrak{m}|_{\Omega})$. Since $\overline{\Omega}$ is compact, $b_t|_{\overline{\Omega}}$ uniformly converge to $b|_{\overline{\Omega}}$ as $t \to \infty$. Also, we know that $\mathrm{lip}^+(b_t) \equiv 1$ on $X \setminus \{\gamma_t\}$, so that by (5.3) we deduce $|Db_t|_w = 1$ $\mathfrak{m}$-a.e.. Since we previously proved that $\mathrm{lip}^+(b) \equiv 1$, we also have $|Db|_w = 1$ $\mathfrak{m}$-a.e., and therefore, keeping Proposition 2.6 in mind, we get $b_t|_{\overline{\Omega}} \to b|_{\overline{\Omega}}$ in energy in $W^{1,2}(\overline{\Omega}, \mathsf{d}, \mathfrak{m}|_{\overline{\Omega}})$ as $t \to \infty$. The uniform convexity of $W^{1,2}(\overline{\Omega}, \mathsf{d}, \mathfrak{m})$ then easily imply that $b_t|_{\overline{\Omega}} \to b|_{\overline{\Omega}}$ in $W^{1,2}(\overline{\Omega}, \mathsf{d}, \mathfrak{m})$ as $t \to \infty$. In particular, Corollary 3.4 gives

$$\lim_{t \to +\infty} \int Df(\nabla b_t) \, \mathrm{d}\mathfrak{m} = \int Df(\nabla b) \, \mathrm{d}\mathfrak{m}. \tag{5.31}$$

Since Ω is bounded, there is $\bar{t} \geq 0$ such that for $t \geq \bar{t}$ we have $\mathsf{d}(\gamma_t, \Omega) > 1$. Hence, from the trivial inequality $\mathsf{d}(\gamma_t, x) \geq \mathsf{d}(\gamma_t, \Omega)$ valid for any $x \in \Omega$ and (5.30), for $t \geq \bar{t}$ we have

$$b_t \in D(\mathbf{\Delta}, \Omega) \qquad \text{and} \qquad \mathbf{\Delta} b_{t|\Omega} \geq -\frac{N-1}{\mathsf{d}(\gamma_t, \Omega)} \mathfrak{m}. \tag{5.32}$$

Therefore

$$\int f \, \mathrm{d}\mathbf{\Delta} b = -\int Df(\nabla b) \, \mathrm{d}\mathfrak{m} = \lim_{t \to +\infty} -\int Df(\nabla b_t) \, \mathrm{d}\mathfrak{m} = \lim_{t \to +\infty} \int f \, \mathrm{d}\mathbf{\Delta} b_{t|\Omega}$$
$$\geq \lim_{t \to +\infty} -\frac{N-1}{\mathsf{d}(\gamma_t, \Omega)} \int f \, \mathrm{d}\mathfrak{m} = 0,$$

so that from the arbitrariness of f we get the thesis. $\qquad\square$

Notice the slight abuse of notation in the name of Proposition 5.19: we just proved that $\mathbf{\Delta} b \geq 0$, but we didn't prove any sort of maximum principle, so that we don't really know whether a function with non-negative Laplacian stays below, when restricted to some ball B the harmonic function which attains the same value at the boundary of B.

We also remark that it is the lack of the strong maximum principle that is preventing to prove that the Busemann function associated to a *line* is harmonic on $\mathsf{CD}(0, N)$ spaces. This subject will need further investigations in the future[6].

We conclude the chapter with some comments about the splitting theorem in this context. The key steps of its proof in smooth Riemannian manifolds are:

  i) To prove that the Busemann function is superharmonic

  ii) To use the strong maximum principle to get that the Busemann function associated to a line is harmonic

  iii) To use the equality case in the Bochner inequality to deduce that the Busemann function has 0 Hessian

  iv) To recall that the gradient flow of a convex function is a contraction, so that the gradient flow of the Busemann function produces a family of isometries

  v) To conclude that the manifold isometrically factorizes as $\{b = 0\} \times \mathbb{R}$

Out of all these steps, the key one which fails in a Finsler setting is the fifth[7]. The problem is that on the product space $\{b = 0\} \times \mathbb{R}$ one puts the product distance

$$\tilde{\mathsf{d}}\big((x,t),(x',t')\big) := \sqrt{\mathsf{d}^2(x,x') + |t-t'|^2}, \tag{5.33}$$

---

[6]in the context of non-linear potential theory the strong maximum principle for local sub/super-minimizers of the energy has been proved in [10]. In the recent paper [23] it has been shown that on doubling spaces supporting a 2-Poincaré inequality, local sub/super-mininimizers of the energy can be characterized as those functions having non-negative/non-positive distributional Laplacian, in line with the smooth case, thus providing the converse implication of the one given in Proposition 4.16.

[7]this is not the only part of the proof of the splitting theorem which fails in a Finsler setting. Another problem is that in general the gradient flow of a convex function does *not* contract the distance, as deeply and carefully explained in [36]. Yet, this is not the crucial point. For instance, in the case $(\mathbb{R}^2, \|\cdot\|, \mathcal{L}^2)$, where $\|\cdot\|$ is a smooth and uniformly convex norm, the gradient flow of the Busemann function b associated to any geodesic line, provides a family of translations. Thus in particular it provides a one parameter family of isometries. But due to the local structure of the arbitrary norm $\|\cdot\|$ we can't conclude that it holds $\mathbb{R}^2 = \{b = 0\} \times \mathbb{R}$ as metric spaces.

which is a Riemannian product in nature. For example, in the case $(\mathbb{R}^2, \|\cdot\|, \mathcal{L}^2)$ we certainly have that the measure factorizes, i.e. $\mathcal{L}^2 = \mathcal{L}^1 \times \mathcal{L}^1$, but (5.33) in general fails (as far as I know, the validity of the splitting theorem in a Finsler context is an open problem[8] if one allows for product distances different than the one in (5.33)).

Proposition 5.19 addresses (i) in the abstract framework, but given that Finsler geometries are included in the class of $\mathsf{CD}(K, N)$ spaces, we know that the splitting theorem fails in this setting. The hope then would be to find an appropriate additional condition which enforces a Riemannian-like behavior on small scales. The natural candidate seems to be infinitesimal Hilbertianity, discussed in Section 4.3: this was one of the research perspectives behind the definition of spaces with Riemannian Ricci curvature bounded from below given in [6]. But to understand whether the splitting theorem holds under this additional assumption still requires quite a lot of work[9].

A   On the duality between cotangent and tangent spaces

The topic we try to address here is up to what extent one can use the differential calculus developed in Chapter 3 to provide an abstract notion of cotangent and tangent spaces on arbitrary metric measure spaces. For simplicity, we are going to assume that the reference measure $\mathfrak{m}$ is a probability measure.

Definition A.1 (Cotangent and tangent spaces) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) with $\mathfrak{m} \in \mathscr{P}(X)$, and $p, q \in (1, \infty)$ conjugate exponents.*

The cotangent space $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$ is defined as the abstract completion of $\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})/\sim$, where we say that $f \sim g$ provided $\|f - g\|_{\mathrm{S}^p} = 0$. Given $f \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$, we will denote by Df the equivalence class of f in $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$.

The tangent space $\mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ is the dual of $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$.

Remark A.2 If $(X, \mathsf{d}, \mathfrak{m})$ is a Finsler or Riemannian manifold, then the notion of cotangent space we just provided does not really coincide with the family of p-integrable cotangent vector fields on the manifold, but only with those arising as differentials of functions, which is clearly a smaller class. Similarly for the tangent space.

If one wants a notion which produces all cotangent vector fields, he can proceed with a patching argument, following the idea that every smooth cotangent vector field is, locally, close to the differential of a function. More precisely, one starts considering the set V defined as

$$V := \Big\{ \{(A_i, f_i)\}_{i \in \mathbb{N}} \ : \ A_i \cap A_j = \emptyset \text{ for } i \neq j, \ X = \cup_{i \in \mathbb{N}} A_i, \ f_i \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m}), \ \forall i \in \mathbb{N} \Big\}.$$

---

[8]after the work on this paper was finished, I got aware of a very recent result by Ohta, who showed that a diffeomorphic and measure preserving splitting holds on smooth Finsler geometries, and that the gradient flow of the Busemann function provides a 1-parameter family of isometries on Berwald spaces, see [33].

[9]In the recent paper [21] it has been indeed proved that the splitting holds on infinitesimally Hilbertian $\mathsf{CD}(0, N)$ spaces. The argument does not make explicit use of the Hessian, but is rather based on the fact that the Busemann function is a pointwise minimizer for the Bochner inequality.

Then a sum, a multiplication by a scalar and a seminorm on V are defined as

$$\left\{(A_i, f_i)\right\}_{i\in\mathbb{N}} + \left\{(\tilde{A}_j, \tilde{f}_j)\right\}_{j\in\mathbb{N}} := \left\{(A_i \cap \tilde{A}_j, f_i + \tilde{f}_j)\right\}_{i,j\in\mathbb{N}},$$

$$\lambda \cdot \left\{(A_i, f_i)\right\}_{i\in\mathbb{N}} := \left\{(A_i, \lambda f_i)\right\}_{i\in\mathbb{N}},$$

$$\left\|\left\{(A_i, f_i)\right\}_{i\in\mathbb{N}}\right\|_{\mathrm{S}^p}^p := \sum_{i\in\mathbb{N}} \int_{A_i} |Df_i|_w^p \, \mathrm{d}\mathfrak{m}.$$

It is then easy to check that the quotient of V w.r.t. the subspace of elements with 0 norm is a normed vector space. The abstract completion of this space is a Banach space, which in a smooth situation can be identified with the space of p-integrable cotangent vector fields. Its dual can therefore be identified with the space of q-integrable tangent vector fields.

A similar construction can be made on the space $(\mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m}), \mathrm{Dist}_q)$ introduced below.

The biggest drawback of the definition we proposed is that the cotangent and tangent spaces are just vector spaces, and not moduli over L^∞. This in particular prevents discussing the validity of the Leibniz rule $D(fg) = fDg + gDf$, simply because we didn't define the objects fDg and gDf. Some results in this direction are given in [49] via completely different means. We expect the two presentation to be compatible, but it is not clear to us whether this is really the case or not, partly because the setting is not completely the same: Weaver works with a metric space, a notion of null sets and differential of Lipschitz functions while we work on metric measure spaces and differentials of Sobolev functions.

Anyway, for the purpose of the discussion we want to make here, we will be satisfied with Definition A.1. ∎

For $L \in \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ and $g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$, we will denote by $Dg(L) \in \mathbb{R}$ the dualism between the operator L and the equivalence class Dg of g in $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$. Notice that it holds

$$Dg(L) \leq \|Dg\|_{\mathrm{CoTan}^p}\|L\|_{\mathrm{Tan}^q} \leq \frac{\|Dg\|_{\mathrm{CoTan}^p}^p}{p} + \frac{\|L\|_{\mathrm{Tan}^q}^q}{q}, \tag{A.1}$$

for any $g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$, $L \in \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$, hence in analogy with Definition 3.7, we may say that L q-represents ∇g provided

$$Dg(L) \geq \frac{\|Dg\|_{\mathrm{CoTan}^p}^q}{p} + \frac{\|L\|_{\mathrm{Tan}^q}^q}{q}. \tag{A.2}$$

Also, we say that $L \in \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ is *almost represented* by a q-test plan $\boldsymbol{\pi} \in \mathscr{P}(C([0,1], X))$ provided $(\mathrm{e}_0)_\sharp \boldsymbol{\pi} = \mathfrak{m}$, $\|\boldsymbol{\pi}\|_q < \infty$ and for any $f \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ it holds

$$\underline{\lim_{t\downarrow 0}} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \leq Df(L) \leq \overline{\lim_{t\downarrow 0}} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma). \tag{A.3}$$

Notice that as a consequence of Theorem 3.10, if $(X, \mathsf{d}, \mathfrak{m})$ is q-infinitesimally strictly convex, the the $\underline{\lim}$ and $\overline{\lim}$ in (A.3) are actually limits and the equality holds.

Proposition A.3 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) with $\mathfrak{m} \in \mathscr{P}(X)$, $\boldsymbol{\pi} \in \mathscr{P}(C([0,1], X))$ be a q-test plan with $(\mathrm{e}_0)_\sharp \boldsymbol{\pi} = \mathfrak{m}$ and $\|\boldsymbol{\pi}\|_q < \infty$. Then there exists at least one $L \in \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ which is almost represented by $\boldsymbol{\pi}$, and for any such L it holds*

$$\|L\|_{\mathrm{Tan}^q} \leq \|\boldsymbol{\pi}\|_q. \tag{A.4}$$

Also, if $\boldsymbol{\pi} \in \mathscr{P}(C([0,1], X))$ is a q-test plan which q-represents ∇g for some $g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$, and L is almost represented by $\boldsymbol{\pi}$, then L q-represents ∇g as well and $\|L\|_{\mathrm{Tan}^q} = \|\boldsymbol{\pi}\|_q$.

proof Consider the map

$$\mathrm{S}^p(X, \mathsf{d}, \mathfrak{m}) \ni f \qquad \mapsto \qquad T(f) := \overline{\lim_{t \downarrow 0}} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma).$$

It is clearly convex and positively 1-homogeneous. The trivial inequality $T(f) - T(g) \leq T(f - g) \leq \|f - g\|_{\mathrm{S}^p} \|\boldsymbol{\pi}\|_q$ yields that T passes to the quotient and defines a convex, positively 1-homogeneous map, still denoted by T, on $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$, and this map is also Lipschitz, with Lipschitz constant $\|\boldsymbol{\pi}\|_q$. By the Hahn-Banach theorem, there exists a linear map $L : \mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m}) \to \mathbb{R}$ such that

$$Df(L) \leq T(f), \qquad \forall f \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m}).$$

Since T is Lipschitz, L is bounded, hence $L \in \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$. Exchanging f with $-f$ we conclude that L is almost represented by $\boldsymbol{\pi}$. The fact that any L which is almost represented by $\boldsymbol{\pi}$ satisfies (A.4) is obvious.

The second part of the statement is a direct consequence of the definition and of the inequality (A.4). $\qquad \square$

With the very same arguments used in the proof of Theorem 3.10, we get that for $f, g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ and $L \in \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ which q-represents ∇g it holds

$$\int D^+ f(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}\mathfrak{m} \geq Df(L) \geq \int D^- f(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}\mathfrak{m}. \tag{A.5}$$

The compactness properties of $\mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ allow to prove that equalities can actually hold.

Theorem A.4 (Horizontal and vertical derivatives - II) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) with $\mathfrak{m} \in \mathscr{P}(X)$, $p \in (1, \infty)$ and $f, g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$. Then there exist $L^+, L^- \in \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$, depending on both f and g, q-representing ∇g such that*

$$\int D^+ f(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}\mathfrak{m} = Df(L^+), \tag{A.6}$$

$$\int D^- f(\nabla g)|Dg|_w^{p-2} \, \mathrm{d}\mathfrak{m} = Df(L^-). \tag{A.7}$$

proof We prove (A.6) only, as the proof of (A.7) follows along similar lines. Taking (A.5) into account, we need only to find some L^+ for which $\leq$ holds. For $\varepsilon > 0$, use Theorem 3.14 to find a q-test plan $\boldsymbol{\pi}^\varepsilon$ which q-represents $\nabla(g + \varepsilon f)$ and such that $(\mathrm{e}_0)_\sharp \boldsymbol{\pi}^\varepsilon = \mathfrak{m}$, and then Proposition A.3 to find $L^\varepsilon \in \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ almost represented by $\boldsymbol{\pi}^\varepsilon$. By (A.1) and (A.2) we have

$$\begin{aligned} D(g + \varepsilon f)(L^\varepsilon) &\geq \frac{\|D(g + \varepsilon f)\|_{\mathrm{CoTan}^p}^p}{p} + \frac{\|L^\varepsilon\|_{\mathrm{Tan}^q}^q}{q}, \\ Dg(L^\varepsilon) &\leq \frac{\|Dg\|_{\mathrm{CoTan}^p}^p}{p} + \frac{\|L^\varepsilon\|_{\mathrm{Tan}^q}^q}{q}. \end{aligned} \tag{A.8}$$

Subtracting and dividing by $\varepsilon > 0$ we get

$$Df(L^\varepsilon) \geq \int \frac{|D(g + \varepsilon f)|_w^p - |Dg|_w^p}{\varepsilon p} \, \mathrm{d}\mathfrak{m}.$$

As $\varepsilon \downarrow 0$, the right hand side converges to $\int D^+ f(\nabla g)|Dg|_w^{p-2}\,d\mathfrak{m}$. Concerning the left one, notice that the bound

$$\|L^\varepsilon\|_{\mathrm{Tan}^q} \le \|\boldsymbol{\pi}^\varepsilon\|_q \le \|D(g+\varepsilon f)\|_{\mathrm{CoTan}^p}^{\frac{p}{q}} \le \left(\|Dg\|_{\mathrm{CoTan}^p} + \varepsilon\|Df\|_{\mathrm{CoTan}^p}\right)^{\frac{p}{q}}.$$

ensures that the family $\{L^\varepsilon\}_{\varepsilon \in (0,1)}$ is weakly* relatively compact in $\mathrm{Tan}^q(X,\mathsf{d},\mathfrak{m})$. Hence for some sequence $\varepsilon_n \downarrow 0$ we have $L^{\varepsilon_n} \rightharpoonup^* L^+ \in \mathrm{Tan}^q(X,\mathsf{d},\mathfrak{m})$, and therefore

$$Df(L^+) \ge \int D^+ f(\nabla g)|Dg|_w^{p-2}\,d\mathfrak{m}.$$

It remains to prove that L^+ q-represents ∇g. But this is obvious, as it is sufficient to pass to the limit in the first inequality in (A.8) taking into account that $\|L^+\|_{\mathrm{Tan}^q} \le \varliminf_{n\to\infty} \|L^{\varepsilon_n}\|_{\mathrm{Tan}^q}$. $\qquad\square$

Now assume that $(X,\mathsf{d},\mathfrak{m})$ is q-infinitesimally strictly convex. Then for $g \in \mathrm{S}^p(X,\mathsf{d},\mathfrak{m})$ and $\boldsymbol{\pi}$ q-test plan which q-represents ∇g and $(\mathrm{e}_0)_\sharp \boldsymbol{\pi} = \mathfrak{m}$ we have

$$\lim_{t\to 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t}\,d\boldsymbol{\pi}(\gamma) = \int Df(\nabla g)|Dg|_w^{p-2}\,d\mathfrak{m}, \qquad \forall f \in \mathrm{S}^p(X,\mathsf{d},\mathfrak{m}).$$

In other words, the map

$$\mathrm{S}^p(X,\mathsf{d},\mathfrak{m}) \ni f \qquad \mapsto \qquad \lim_{t\to 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t}\,d\boldsymbol{\pi}(\gamma), \qquad\qquad (\mathrm{A.9})$$

is a linear bounded functional on $\mathrm{CoTan}^p(X,\mathsf{d},\mathfrak{m})$, i.e. an element of $\mathrm{Tan}^q(X,\mathsf{d},\mathfrak{m})$. We define the set $\mathrm{Grad}^q(X,\mathsf{d},\mathfrak{m})$ as:

$$\mathrm{Grad}^q(X,\mathsf{d},\mathfrak{m}) := \Big\{ \boldsymbol{\pi} \in \mathscr{P}(C([0,1],X)) \ : \ \boldsymbol{\pi} \text{ is a } q \text{ test plan with } (\mathrm{e}_0)_\sharp \boldsymbol{\pi} = \mathfrak{m}$$
$$\text{which } q \text{ represents } \nabla g \text{ for some } g \in \mathrm{S}^p \Big\}.$$

Then (A.9) provides a map $\iota : \mathrm{Grad}^q(X,\mathsf{d},\mathfrak{m}) \to \mathrm{Tan}^q(X,\mathsf{d},\mathfrak{m})$. Notice that as a consequence of $Df(\iota(\boldsymbol{\pi})) \le \|f\|_{\mathrm{S}^p}\|\boldsymbol{\pi}\|_q$ and of $Dg(\iota(\boldsymbol{\pi})) \ge \|g\|_{\mathrm{S}^p}\|\boldsymbol{\pi}\|_q$ if $\boldsymbol{\pi}$ q-represents ∇g, we get that $\|\boldsymbol{\pi}\|_q = \|\iota(\boldsymbol{\pi})\|_{\mathrm{Tan}^q}$ for any $\boldsymbol{\pi} \in \mathrm{Grad}^q(X,\mathsf{d},\mathfrak{m})$. It is for this reason that we called $\|\boldsymbol{\pi}\|_q$ the q-"norm" of $\boldsymbol{\pi}$.

Remark A.5 We remark that in passing from $\boldsymbol{\pi}$ to $\iota(\boldsymbol{\pi})$ there is a loss of information because plans can be localized (i.e. one can consider the action of a plan on any Borel subset $A \subset X$ with positive mass by looking at the plan $\mathfrak{m}(A)^{-1}\boldsymbol{\pi}\big|_{\mathrm{e}_0^{-1}(A)}$), while the same operation a priori cannot be performed on elements of $\mathrm{Tan}^q(X,\mathsf{d},\mathfrak{m})$.

Yet, it is possible to check that at least under the assumptions of Theorem A.6 below, an adequate localization of elements on $\mathrm{Tan}^q(X,\mathsf{d},\mathfrak{m})$ is possible. We won't discuss this topic here. $\qquad\blacksquare$

Interestingly enough, if $\mathrm{CoTan}^p(X,\mathsf{d},\mathfrak{m})$ is reflexive, the range of ι is dense in $\mathrm{Tan}^q(X,\mathsf{d},\mathfrak{m})$ in the sense expressed by the following theorem. Notice that we are going to assume that $\mathrm{CoTan}^p(X,\mathsf{d},\mathfrak{m})$ is smooth: by this we will mean that its norm is differentiable everywhere outside 0. This assumption certainly implies that $(X,\mathsf{d},\mathfrak{m})$ is q-infinitesimally strictly convex,

as the latter just means that $\mathrm{S}^p/\sim$ is smooth. It is unclear to us whether the converse implication also holds, as we don't know whether any loss of smoothness can happen in the completion process. We also remark that this result goes in the same direction of Theorem 4.3 in [17] obtained by Cheeger and Kleiner, where the span of the set of derivations along curves is shown to be dense in the tangent space under doubling and Poincaré assumptions.

Theorem A.6 (Duality between cotangent and tangent spaces) *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) with $\mathfrak{m} \in \mathscr{P}(X)$ and $p, q \in (1, \infty)$ two conjugate exponents. Assume that $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$ is smooth and reflexive. Then for every $L \in \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ there exists a sequence $(\boldsymbol{\pi}_n) \subset \mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$ such that*

$$\iota(\boldsymbol{\pi}^n) \rightharpoonup L, \qquad and \qquad \|\boldsymbol{\pi}^n\|_q \to \|L\|_{\mathrm{Tan}^q},$$

as $n \to \infty$. In particular, if $\mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ is uniformly convex, then $\iota(\mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m}))$ is dense in $\mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ w.r.t. the strong topology.

proof Consider the maps $\mathcal{N} : \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m}) \to \mathbb{R}$ and $\mathcal{N}^* : \mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m}) \to \mathbb{R}$ given by

$$\mathcal{N}(L) := \frac{\|L\|_{\mathrm{Tan}^q}^q}{q}, \qquad \forall L \in \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m}),$$

$$\mathcal{N}^*(\omega) := \frac{\|\omega\|_{\mathrm{CoTan}^p}^p}{p}, \qquad \forall \omega \in \mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m}).$$

Clearly, $\mathcal{N}$ is the Legendre transform of $\mathcal{N}^*$, and since $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$ is assumed to be reflexive, also the viceversa holds. Since both functions are locally Lipschitz, their subdifferential is always non-empty, and the assumption that $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$ is strictly convex means that $\partial^- \mathcal{N}^*(\omega)$ is single valued for any $\omega \in \mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$.

Fix $L \in \mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ and choose $\omega \in \partial^- \mathcal{N}(L)$, so that $L \in \partial^- \mathcal{N}^*(\omega)$. By definition of $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$, we can find a sequence $(f_n) \subset \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ such that $Df_n \to \omega$ in $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$. For every $n \in \mathbb{N}$, use Theorem 3.14 to find $\boldsymbol{\pi}_n \in \mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$ which q-represents ∇f_n so that it holds

$$Df_n(\iota(\boldsymbol{\pi}_n)) \geq \frac{\|Df_n\|_{\mathrm{CoTan}^p}^p}{p} + \frac{\|\iota(\boldsymbol{\pi}_n)\|_{\mathrm{Tan}^q}^q}{q}. \tag{A.10}$$

Since $\sup_n \|\iota(\boldsymbol{\pi}_n)\|_{\mathrm{Tan}^q} < \infty$, up to pass to a subsequence, not relabeled, we may assume that $\iota(\boldsymbol{\pi}_n)$ weakly converges to some $\tilde{L}$ for which it holds $\|\tilde{L}\|_{\mathrm{Tan}^q} \leq \varliminf_{n \to \infty} \|\iota(\boldsymbol{\pi}_n)\|_{\mathrm{Tan}^q}$. Passing to the limit in (A.10) using the strong convergence of Df_n to ω in $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$ we obtain

$$\omega(\tilde{L}) \geq \frac{\|\omega\|_{\mathrm{CoTan}^p}^p}{p} + \frac{\|\tilde{L}\|_{\mathrm{Tan}^q}^q}{q}, \tag{A.11}$$

i.e. $\tilde{L} \in \partial^- \mathcal{N}^*(\omega)$. Since we know by assumption that $\partial^- \mathcal{N}^*(\omega)$ is a singleton, it must hold $\tilde{L} = L$. To conclude we need to show that $\|\iota(\boldsymbol{\pi}_n)\|_{\mathrm{Tan}^q} \to \|L\|_{\mathrm{Tan}^q}$ as $n \to \infty$. This is a consequence of the fact that the inequality in (A.11) cannot be strict, thus it can't be any loss of norm in passing from (A.10) to (A.11).

For the last part of the statement, just notice that the strict convexity of $\mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ implies the smoothness of $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$, and that its uniform convexity yields that weak convergence plus convergence of norms is equivalent to strong convergence. $\qquad\square$

As the proof shows, if $(S^p/\sim, \|\cdot\|_{S^p})$ is complete, then *every* element of $\mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ can be recovered as derivation along some $\boldsymbol{\pi} \in \mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$.

In any case, the theorem has the following geometric consequence. Consider $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2 \in \mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$, put $L := \iota(\boldsymbol{\pi}_1) + \iota(\boldsymbol{\pi}_2)$ and find a sequence $(\tilde{\boldsymbol{\pi}}_n) \subset \mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$ converging to L in the sense of the above theorem. Then the (derivative at time 0 of the) curves in the support of $\tilde{\boldsymbol{\pi}}_n$ can be seen as approximate sum of the (derivative at time 0 of the) curves in the support of $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2$. At least in the sense of their action on differentials. In case of infinitesimally Hilbertian spaces, there is no need of passing through an approximation: if $\boldsymbol{\pi}_i$ 2-represents ∇g_i, $g_i \in S^2(X, \mathsf{d}, \mathfrak{m})$, $i = 1, 2$, then any plan 2-representing $\nabla(g_1 + g_2)$ can be seen as the 'sum' of $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ in the sense just described. A deeper geometric meaning to this fact could be given in case the open problem mentioned at the end of this appendix has an affirmative answer, see the discussion below.

The set $\mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$ can be endowed with the following natural semi-distance:

$$\mathrm{Dist}_q(\boldsymbol{\pi}^1, \boldsymbol{\pi}^2) := \overline{\lim_{t\downarrow 0}} \frac{1}{t} \sqrt[q]{\int W_q^q\Big((\mathrm{e}_t)_\sharp \boldsymbol{\pi}_x^1, (\mathrm{e}_t)_\sharp \boldsymbol{\pi}_x^2\Big) \, \mathrm{d}\mathfrak{m}(x)}, \tag{A.12}$$

where $\{\boldsymbol{\pi}_x^i\}$ is the disintegration of $\boldsymbol{\pi}^i$ w.r.t. the map e_0, $i = 1, 2$. Notice that $\|\boldsymbol{\pi}\|_q = \mathrm{Dist}_q(\boldsymbol{\pi}, \mathbf{0})$, where the zero plan $\mathbf{0}$ is the only plan in $\mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$ concentrated on constant curves.

Remark A.7 To understand why we are saying that this is the natural distance, start noticing that if we were on a Finsler manifold and two tangent vector fields v^1, v^2 in L^q were given, then certainly the distance between them would have been $\sqrt[q]{\int \|v^1 - v^2\|^q \, \mathrm{d}\mathfrak{m}}$. Now, if our vector fields were represented by smooth families of diffeomorphisms Φ_t^i, $i = 1, 2$, in the sense that $\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t^i|_{t=0} = v^i$, $i = 1, 2$, then their distance read at the level of the Φ_t^i's can be written as

$$\lim_{t\downarrow 0} \frac{1}{t} \sqrt[q]{\int \mathsf{d}^q(\Phi_t^1, \Phi_t^2) \, \mathrm{d}\mathfrak{m}}. \tag{A.13}$$

In an abstract metric measure space we don't have any a priori notion of tangent vector field. But as already discussed various times along the paper, a general idea behind our discussion is to try to define tangent vector fields as 'derivative at time 0 of appropriate curves'. In practice, it seems more convenient to work with weighted family of curves, i.e. with probability measures in $C([0, 1], X)$, rather than with curves themselves, which lead us to the concept of 'plan representing a gradient'. If they are given $\boldsymbol{\pi}^1, \boldsymbol{\pi}^2 \in \mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$ with the property that $\boldsymbol{\pi}_x^i$ is concentrated on a single curve, say γ_x^i, for $\mathfrak{m}$-a.e. x and $i = 1, 2$, then by analogy with (A.13) their distance in the tangent space should be defined as

$$\overline{\lim_{t\downarrow 0}} \frac{1}{t} \sqrt[q]{\int \mathsf{d}^q\big((\gamma_x^1)_t, (\gamma_x^2)_t\big) \, \mathrm{d}\mathfrak{m}},$$

the lim being substituted with a $\overline{\lim}$ because of a lack of an a priori notion of smoothness. However, for a general plan in $\mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$ we don't know whether its disintegration is made of deltas or not, so that we don't really have a single point $(\gamma_x)_t$ which is the position at time t of the particle which originally was in x. What we have is the probability distribution $(\mathrm{e}_t)_\sharp \boldsymbol{\pi}_x$. Thus we have to substitute the expression $\mathsf{d}^q((\gamma_x^1)_t, (\gamma_x^2)_t)$ with its analogous $W_q^q\Big((\mathrm{e}_t)_\sharp \boldsymbol{\pi}_x^1, (\mathrm{e}_t)_\sharp \boldsymbol{\pi}_x^1\Big)$. This operation leads to formula (A.12). $\blacksquare$

If our interpretation of plans in $\mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$ as tangent vector fields is correct, we expect the distance between $\iota(\boldsymbol{\pi}^1)$ and $\iota(\boldsymbol{\pi}^2)$ on the tangent space to be bounded by $\mathrm{Dist}^q(\boldsymbol{\pi}^1, \boldsymbol{\pi}^2)$, because Dist^q is what controls the infinitesimal relative behavior of $\boldsymbol{\pi}^1$ and $\boldsymbol{\pi}^2$. In the following proposition we show that this is actually the case if $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$ is uniformly convex. Recall that given $f : X \to \mathbb{R}$ the function $\overline{|Df|} : X \to [0, \infty]$ is defined as 0 if x is isolated, otherwise as

$$\overline{|Df|}(x) := \inf_{r>0} \sup_{y_1 \neq y_2 \in B_r(x)} \frac{|f(y_1) - f(y_2)|}{\mathsf{d}(y_1, y_2)}.$$

Proposition A.8 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7) with $\mathfrak{m} \in \mathscr{P}(X)$ and $p, q \in (1, \infty)$ two conjugate exponents. Assume that $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$ is uniformly convex. Let $\boldsymbol{\pi}^1, \boldsymbol{\pi}^2 \in \mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$. Then for every $f \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$ it holds*

$$\varlimsup_{t \downarrow 0} \left| \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}^1(\gamma) - \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}^2(\gamma) \right| \leq \|Df\|_{\mathrm{CoTan}^p} \mathrm{Dist}_q(\boldsymbol{\pi}^1, \boldsymbol{\pi}^2).$$
(A.14)

In particular, if $\mathrm{Dist}_q(\boldsymbol{\pi}^1, \boldsymbol{\pi}^2) = 0$ and $\boldsymbol{\pi}^1$ q-represents ∇g for some $g \in \mathrm{S}^p(X, \mathsf{d}, \mathfrak{m})$, then $\boldsymbol{\pi}^2$ q-represents ∇g as well.

proof Let $f : X \to \mathbb{R}$ be a Lipschitz function and for every $x \in X$ define $G_x : [0, \infty) \to \mathbb{R}$ as

$$G_x(r) := \begin{cases} \displaystyle \sup_{y_1 \neq y_2 \in B_r(x)} \frac{|f(y_1) - f(y_2)|}{\mathsf{d}(y_1, y_2)}, & \text{if } r > 0, \text{ and } B_r(x) \neq \{x\}, \\[2em] \overline{|Df|}(x), & \text{if } r = 0, \text{ or } B_r(x) = \{x\}. \end{cases}$$

By definition, $G_x(\cdot)$ is continuous at 0, and it holds $|f(y_1) - f(y_2)| \leq \mathsf{d}(y_1, y_2) G_x(\max\{\mathsf{d}(y_1, x), \mathsf{d}(y_2, x)\})$ for any $x, y_1, y_2 \in X$.

Let $\{\boldsymbol{\pi}_x^i\}$ be the disintegration of $\boldsymbol{\pi}^i$ w.r.t. e_0, $i = 1, 2$ and, for any $t > 0$, let $\boldsymbol{\gamma}_{t,x}$ be an optimal plan from $(\mathrm{e}_0)_\sharp \boldsymbol{\pi}_x^1$ to $(\mathrm{e}_0)_\sharp \boldsymbol{\pi}_x^2$. We have

$$\left| \int f(\gamma_t) - f(\gamma_0) \, \mathrm{d}\boldsymbol{\pi}^1(\gamma) - \int f(\gamma_t) - f(\gamma_0) \, \mathrm{d}\boldsymbol{\pi}^2(\gamma) \right|$$

$$= \left| \int \left(\int f(\gamma_t) \, \mathrm{d}\boldsymbol{\pi}_x^1(\gamma) - \int f(\gamma_t) \, \mathrm{d}\boldsymbol{\pi}_x^2(\gamma) \right) \, \mathrm{d}\mathfrak{m}(x) \right|$$

$$= \left| \iint f(y_1) - f(y_2) \, \mathrm{d}\boldsymbol{\gamma}_{t,x}(y_1, y_2) \, \mathrm{d}\mathfrak{m}(x) \right|$$

$$\leq \iint G_x\big(\max\{\mathsf{d}(y_1, x), \mathsf{d}(y_2, x)\}\big) \mathsf{d}(y_1, y_2) \, \mathrm{d}\boldsymbol{\gamma}_{t,x}(y_1, y_2) \, \mathrm{d}\mathfrak{m}(x)$$

$$\leq \sqrt[p]{\iint \big(G_x\big(\max\{\mathsf{d}(y_1, x), \mathsf{d}(y_2, x)\}\big)\big)^p \, \mathrm{d}\boldsymbol{\gamma}_{t,x}(y_1, y_2) \, \mathrm{d}\mathfrak{m}(x)} \sqrt[q]{\int W_q^q\big((\mathrm{e}_t)_\sharp \boldsymbol{\pi}_x^1, (\mathrm{e}_t)_\sharp \boldsymbol{\pi}_x^2 \big) \, \mathrm{d}\mathfrak{m}(x)}.$$

Dividing by t, letting $t \downarrow 0$ and taking into account that $\boldsymbol{\gamma}_{t,x}$ converges to $\delta_x \times \delta_x$ in duality with $C_b(X^2)$ we get

$$\varlimsup_{t \downarrow 0} \left| \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}^1(\gamma) - \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}^2(\gamma) \right| \leq \big\| \overline{|Df|} \big\|_{L^p} \mathrm{Dist}_q(\boldsymbol{\pi}^1, \boldsymbol{\pi}^2). \quad \text{(A.15)}$$

Pick now $f \in S^p(X, \mathsf{d}, \mathfrak{m}) \cap L^\infty(X, \mathfrak{m}) \subset W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ and use Theorem 2.8 and Corollary 2.9 to find a sequence (f_n) of Lipschitz functions such that $\overline{|Df_n|} \to |Df|_w$ in $L^p(X, \mathfrak{m})$ and $\|f - f_n\|_{S^p} \to 0$ as $n \to \infty$. Then observe that for a general $\boldsymbol{\pi} \in \mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m})$ it holds

$$
\begin{aligned}
&\varlimsup_{t \downarrow 0} \left| \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) - \int \frac{f_n(\gamma_t) - f_n(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \right| \\
&= \varlimsup_{t \downarrow 0} \left| \int \frac{(f - f_n)(\gamma_t) - (f - f_n)(\gamma_0)}{t} \, \mathrm{d}\boldsymbol{\pi}(\gamma) \right| \leq \|f - f_n\|_{S^p} \|\boldsymbol{\pi}\|_q,
\end{aligned}
\tag{A.16}
$$

so that taking (A.15) into account, (A.14) follows.

For the case of general $f \in S^p(X, \mathsf{d}, \mathfrak{m})$, just truncate f defining $f_N := \min\{\max\{f, -N\}, N\} \in S^p(X, \mathsf{d}, \mathfrak{m}) \cap L^\infty(X, \mathfrak{m})$, observe that $\|f - f_N\|_{S^p} \to 0$ as $N \to \infty$ and conclude as in (A.16) $\qquad \square$

Notice that for q-infinitesimally strictly convex spaces, (A.14) reads as

$$
\|\iota(\boldsymbol{\pi}^1) - \iota(\boldsymbol{\pi}^2)\|_{\mathrm{Tan}^q} \leq \mathrm{Dist}_q(\boldsymbol{\pi}^1, \boldsymbol{\pi}^2), \qquad \forall \boldsymbol{\pi}^1, \boldsymbol{\pi}^2 \in \mathrm{Grad}^q,
\tag{A.17}
$$

and that if both $\mathrm{Tan}^q(X, \mathsf{d}, \mathfrak{m})$ and $\mathrm{CoTan}^p(X, \mathsf{d}, \mathfrak{m})$ are uniformly convex, then Theorem A.6 and Proposition A.8 ensure that the map ι can be continuously extended to the abstract completion $\overline{\mathrm{Grad}^q}(X, \mathsf{d}, \mathfrak{m})$ of $(\mathrm{Grad}^q(X, \mathsf{d}, \mathfrak{m}), \mathrm{Dist}_q)$, and this extension provides a 1-Lipschitz map from $\overline{\mathrm{Grad}^q}$ to Tan^q.

It is not clear if this map is a surjection. This is related to the following open problem: does equality hold in (A.17)?. The answer is expected (at least to me) to be affirmative, because 'gradient vector fields should be completely identified by their actions on differential vector fields'. The question is open also on infinitesimally Hilbertian spaces with $p = q = 2$.

Notice that a positive answer would greatly increase the knowledge of the local geometry of the space. For instance, it would tell that for any $g \in S^p(X, \mathsf{d}, \mathfrak{m})$ and $\mathfrak{m}$-a.e. x there is only one curve, up to equivalence, which realizes the gradient of g at x in the sense that $\gamma_0 = x$ and

$$
\lim_{t \to 0} \frac{g(\gamma_t) - g(\gamma_0)}{t} \geq \frac{|Dg|_w^p(x)}{p} + \varlimsup_{t \downarrow 0} \frac{1}{qt} \int_0^t |\dot\gamma_s|^q \, \mathrm{d}s,
$$

where we are saying that two curves γ^1, γ^2 with $\gamma_0^1 = \gamma_0^2 = x$ are equivalent provided they are tangent at 0, i.e.

$$
\varlimsup_{t \downarrow 0} \frac{\mathsf{d}(\gamma_t^1, \gamma_t^2)}{t} = 0.
$$

B Remarks about the definition of the Sobolev classes

The approach we proposed in Section 2.3 to introduce the Sobolev classes is slightly different from the one used in [4], [5]. Here we collect the simple arguments which show that indeed the two approaches are the same.

As in the rest of the paper, (X, d) is a complete and separable metric space, and $\mathfrak{m}$ is a Radon non-negative measure on X. In [5] (generalizing the approach proposed in [4] for the case $p = q = 2$) it has been first given the definition of q-test plan, as we did in Definition 2.3, then it is introduced the concept of p-negligible set of curves as:

Definition B.1 (p-negligible set of curves) *Let $\Gamma \subset C([0,1],X)$ be a Borel set. We say that Γ is p-negligible provided*

$$\boldsymbol{\pi}(\Gamma) = 0, \qquad \text{for any } q-\text{test plan } \boldsymbol{\pi},$$

where $p, q \in (1, \infty)$ are conjugate exponents. A property which holds for any curve, except for those in a p-negligible set is said to hold for p-almost every curve.

Notice that $C([0,1],X) \setminus AC^q([0,1],X)$ is p-negligible.

Then the definition of functions Sobolev along p-a.e. curve and of p-weak upper gradients were proposed as follows:

Definition B.2 (Functions which are Sobolev along p-a.e. curve) *Let $f : X \to \mathbb{R}$ be a Borel function. We say that f is Sobolev along p-a.e. curve provided for p-a.e. γ the function $f \circ \gamma$ coincides a.e. in $[0,1]$ and in $\{0,1\}$ with an absolutely continuous map $f_\gamma : [0,1] \to \mathbb{R}$.*

Definition B.3 (p-weak upper gradients) *Let f be a function Sobolev along p-a.e. curve. A Borel function $G : X \to [0,\infty]$ is a p-weak upper gradient provided*

$$\left| f(\gamma_1) - f(\gamma_0) \right| \leq \int_0^1 G(\gamma_s)|\dot{\gamma}_s|\, \mathrm{d}s, \qquad \text{for } p-\text{a.e. } \gamma. \tag{B.1}$$

Theorem B.4 *Let $(X, \mathsf{d}, \mathfrak{m})$ be as in (2.7), $p, q \in [1, \infty]$ two conjugate exponents and $f : X \to \mathbb{R}$, $G : X \to [0,\infty]$ Borel functions, with $G \in L^p(X, \mathfrak{m})$. Then the following are equivalent.*

   i) $f \in S^p(X, \mathsf{d}, \mathfrak{m})$ and G is a p-weak upper gradient of f in the sense of Definition 2.4.

   ii) f is Sobolev along p-a.e. curve and G is a p-weak upper gradient of f in the sense of Definitions B.2, B.3.

proof

(ii) $\Rightarrow$ **(i)** Fix a q-test plan $\boldsymbol{\pi}$ and notice that by definition of p-negligible set of curves it holds

$$\left| f(\gamma_1) - f(\gamma_0) \right| \leq \int_0^1 G(\gamma_s)|\dot{\gamma}_s|\, \mathrm{d}s, \qquad \text{for } \boldsymbol{\pi}-\text{a.e. } \gamma.$$

Integrating w.r.t. $\boldsymbol{\pi}$ we get

$$\int \left| f(\gamma_1) - f(\gamma_0) \right| \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \iint_0^1 G(\gamma_s)|\dot{\gamma}_s|\, \mathrm{d}s\, \mathrm{d}\boldsymbol{\pi}(\gamma),$$

which is the thesis.

(i) $\Rightarrow$ **(ii)** Fix a q-test plan $\boldsymbol{\pi}$. Arguing as for (2.11) we have that

$$\forall t < s \in [0,1] \quad \text{it holds} \quad |f(\gamma_s) - f(\gamma_t)| \leq \int_t^s G(\gamma_r)|\dot{\gamma}_r|\, \mathrm{d}r, \qquad \boldsymbol{\pi}-a.e.\ \gamma, \tag{B.2}$$

hence from Fubini's theorem we get

$$\text{for } \boldsymbol{\pi}-a.e.\ \gamma \quad \text{it holds} \quad |f(\gamma_s) - f(\gamma_t)| \leq \int_t^s G(\gamma_r)|\dot{\gamma}_r|\, \mathrm{d}r, \qquad \mathcal{L}^2-a.e.\ t < s \in [0,1].$$
$$\tag{B.3}$$

Now notice that the assumptions $G \in L^p(X, \mathfrak{m})$, $\int E_{q,1} \, d\boldsymbol{\pi} < \infty$ and $(e_t)_\sharp \boldsymbol{\pi} \leq C\mathfrak{m}$ for any $t \in [0,1]$ and some $C \in \mathbb{R}$ easily yield that the function $\gamma \mapsto \int_0^1 G(\gamma_r)|\dot{\gamma}_r| \, dr$ is in $L^1(\boldsymbol{\pi})$. In particular, for $\boldsymbol{\pi}$-a.e. γ the map $t \mapsto G(\gamma_t)|\dot{\gamma}_t|$ is in $L^1(0,1)$. Therefore from (B.3) and Lemma B.5 below we deduce that for $\boldsymbol{\pi}$-a.e. γ the function $t \mapsto f(\gamma_t)$ coincides a.e. in $[0,1]$ with an absolutely continuous map $f_{\gamma, \boldsymbol{\pi}}$, whose derivative is bounded - in modulus - by $G(\gamma_t)|\dot{\gamma}_t|$.

By construction we have

$$f_{\gamma, \boldsymbol{\pi}}(t) = f(\gamma_t), \qquad \boldsymbol{\pi} \times \mathcal{L}^1|_{[0,1]} - a.e. \ (\gamma, t), \tag{B.4}$$

and we need to prove that for $\boldsymbol{\pi}$-a.e. γ it holds $f_{\gamma, \boldsymbol{\pi}}(0) = f(\gamma_0)$ and $f_{\gamma, \boldsymbol{\pi}}(1) = f(\gamma_1)$. Using (B.2) with $t = 0$ and (B.4) we can find a sequence $s_n \downarrow 0$ such that $f_{\gamma, \boldsymbol{\pi}}(s_n) = f(\gamma_{s_n})$ for $\boldsymbol{\pi}$-a.e. γ and every $n \in \mathbb{N}$ and

$$|f_{\gamma, \boldsymbol{\pi}}(s_n) - f(\gamma_0)| \leq \int_0^{s_n} G(\gamma_r)|\dot{\gamma}_r| \, dr, \qquad \boldsymbol{\pi} - a.e. \ \gamma, \ \forall n \in \mathbb{N}.$$

Passing to the limit as $n \to \infty$ we get that for $\boldsymbol{\pi}$-a.e. γ it holds $f_{\gamma, \boldsymbol{\pi}}(0) = \lim_n f_{\gamma, \boldsymbol{\pi}}(s_n) = f(\gamma_0)$. A similar argument works for $t = 1$.

It remains to prove that the functions $f_{\gamma, \boldsymbol{\pi}}$ do not depend on $\boldsymbol{\pi}$, but only on γ. But this is obvious, because the map $t \mapsto f(\gamma_t)$ has at most one continuous representative. $\qquad \square$

The following lemma was proved in [5], the proof is reported here for completeness.

Lemma B.5 *Let* $f : (0,1) \to \mathbb{R}$, $q \in (1, \infty)$, $g \in L^q(0,1)$ *nonnegative be satisfying*

$$|f(s) - f(t)| \leq \left| \int_s^t g(r) \, dr \right| \qquad for \ \mathcal{L}^2\text{-}a.e. \ (s,t) \in (0,1)^2.$$

Then $f \in W^{1,q}(0,1)$ *and* $|f'| \leq g$ *a.e. in* $(0,1)$.

proof Let $N \subset (0,1)^2$ be the $\mathcal{L}^2$-negligible subset where the above inequality fails. By Fubini's theorem, also the set $\{(t,h) \in (0,1)^2 : (t, t+h) \in N \cap (0,1)^2\}$ is $\mathcal{L}^2$-negligible. In particular, by Fubini's theorem, for a.e. h we have $(t, t+h) \notin N$ for a.e. $t \in (0,1)$. Let $h_i \downarrow 0$ with this property and use the identities

$$\int_0^1 f(t) \frac{\phi(t+h) - \phi(t)}{h} \, dt = -\int_0^1 \frac{f(t-h) - f(t)}{-h} \phi(t) \, dt$$

with $\phi \in C_c^1(0,1)$ and $h = h_i$ sufficiently small to get

$$\left| \int_0^1 f(t)\phi'(t) \, dt \right| \leq \int_0^1 g(t)|\phi(t)| \, dt.$$

By L^q duality, this gives the $W^{1,q}(0,1)$ regularity and, at the same time, the inequality $|f'| \leq g$ a.e. in $(0,1)$. $\qquad \square$

References

[1] L. AMBROSIO AND S. DI MARINO, *BV functions and equivalence of weak total variations in metric measure spaces*, preprint, avaliable at `http://cvgmt.sns.it/paper/1860/`, (2012).

[2] L. Ambrosio, N. Gigli, A. Mondino, and T. Rajala, *Riemannian ricci curvature lower bounds in metric measure spaces with σ-finite measure.* to appear in Transactions of the AMS, arXiv:1207.4924, 2011.

[3] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005.

[4] ——, *Calculus and heat flows in metric measure spaces with Ricci curvature bounded from below*, to appear in Inv. Math., arXiv:1106.2090, (2011).

[5] ——, *Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces*, to appear in Rev. Mat. Iber., arXiv:1111.3730, (2011).

[6] ——, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, Submitted paper, arXiv:1109.0222, (2011).

[7] L. Ambrosio, N. Gigli, and G. Savaré, *Heat Flow and calculus on metric measure spaces with Ricci curvature bounded below - the compact case*, in Analysis and Numerics of Partial Differential Equations, F. Brezzi, P. Colli Franzone, U. Gianazza, and G. Gilardi, eds., vol. 4 of Springer INdAM Series, Springer Milan, 2013, pp. 63–115.

[8] K. Bacher and K.-T. Sturm, *Localization and Tensorization Properties of the Curvature-Dimension Condition for Metric Measure Spaces*, JFA, 259 (2010), pp. 28–56.

[9] D. Bakry, *Functional inequalities for Markov semigroups*, in Probability measures on groups: recent directions and trends, Tata Inst. Fund. Res., Mumbai, 2006, pp. 91–147.

[10] A. Björn and J. Björn, *Nonlinear potential theory on metric spaces*, vol. 17 of EMS Tracts in Mathematics, European Mathematical Society (EMS), Zürich, 2011.

[11] V. I. Bogachev, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007.

[12] F. Cavalletti and K.-T. Sturm, *Local curvature-dimension condition implies measure-contraction property*, J. Funct. Anal., 262 (2012), pp. 5110–5127.

[13] J. Cheeger, *Differentiability of Lipschitz functions on metric measure spaces*, Geom. Funct. Anal., 9 (1999), pp. 428–517.

[14] J. Cheeger and T. Colding, *On the structure of spaces with Ricci curvature bounded below I*, J. Diff. Geom., 45 (1997), pp. 406–480.

[15] ——, *On the structure of spaces with Ricci curvature bounded below II*, J. Diff. Geom., 54 (2000), pp. 13–35.

[16] ——, *On the structure of spaces with Ricci curvature bounded below III*, J. Diff. Geom., 54 (2000), pp. 37–74.

[17] J. Cheeger and B. Kleiner, *Differentiability of Lipschitz maps from metric measure spaces to Banach spaces with the Radon-Nikodým property*, Geom. Funct. Anal., 19 (2009), pp. 1017–1028.

[18] A. FIGALLI AND N. GIGLI, *Local semiconvexity of Kantorovich potentials on non-compact manifolds*, ESAIM Contrl Optim. Calc. Var., 17 (2011), pp. 648–653.

[19] N. GIGLI, *On the heat flow on metric measure spaces: existence, uniqueness and stability*, Calc. Var. PDE, 39 (2010), pp. 101–120.

[20] ——, *On the relation between the curvature dimension condition and Bochner inequality*, Unpublished paper, available at `http://cvgmt.sns.it/person/226/`, (2012).

[21] ——, *The splitting theorem in non-smooth context.* Preprint, arXiv:1302.5555, 2013.

[22] N. GIGLI, K. KUWADA, AND S.-I. OHTA, *Heat flow on Alexandrov spaces*, Comm. Pure Appl. Math., 66 (2013), pp. 307–331.

[23] N. GIGLI AND A. MONDINO, *A PDE approach to nonlinear potential theory in metric measure spaces.* Accepted at JMPA, arXiv:1209.3796, 2012.

[24] J. HEINONEN, *Nonsmooth calculus*, Bull. Amer. Math. Soc., 44 (2007), pp. 163–232.

[25] P. KOSKELA AND P. MACMANUS, *Quasiconformal mappings and Sobolev spaces*, Studia Math., 131 (1998), pp. 1–17.

[26] K. KUWADA, *Duality on gradient estimates and wasserstein controls*, Journal of Functional Analysis, 258 (2010), pp. 3758–3774.

[27] K. KUWAE, Y. MACHIGASHIRA, AND T. SHIOYA, *Beginning of analysis on Alexandrov spaces*, in Geometry and topology: Aarhus (1998), vol. 258 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2000, pp. 275–284.

[28] S. LISINI, *Characterization of absolutely continuous curves in Wasserstein spaces*, Calc. Var. Partial Differential Equations, 28 (2007), pp. 85–120.

[29] J. LOTT AND C. VILLANI, *Weak curvature bounds and functional inequalities*, J. Funct. Anal., 245 (2007), pp. 311–333.

[30] J. LOTT AND C. VILLANI, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2), 169 (2009), pp. 903–991.

[31] S.-I. OHTA, *Products, cones, and suspensions of spaces with the measure contraction property*, J. Lond. Math. Soc. (2), 76 (2007), pp. 225–236.

[32] ——, *Finsler interpolation inequalities*, Calc. Var. PDE, (2009), pp. 211–249.

[33] ——, *Splitting theorems for Finsler manifolds of nonnegative Ricci curvature.* to appear in J. Reine Angew. Math., arXiv:1203.0079, 2012.

[34] S.-I. OHTA AND K.-T. STURM, *Heat flow on Finsler manifolds*, Comm. Pure Appl. Math., 62 (2009), pp. 1386–1433.

[35] ——, *Bochner-Weitzenböck formula and Li-Yau estimates on Finsler manifolds*, preprint, arXiv:1104.5276, (2011).

[36] ——, *Non-contraction of heat flow on Minkowski spaces*, Arch. Ration. Mech. Anal., 204 (2012), pp. 917–944.

[37] Y. Otsu and T. Shioya, *The Riemannian structure of Alexandrov spaces*, J. Differential Geom., 39 (1994), pp. 629–658.

[38] G. Perelman, *Dc structure on Alexandrov Space*. Unpublished preprint, available online at `http://www.math.psu.edu/petrunin/papers/alexandrov/Cstructure.pdf`.

[39] A. Petrunin, *Harmonic functions on Alexandrov spaces and their applications*, Electron. Res. Announc. Amer. Math. Soc., 9 (2003), pp. 135–141.

[40] T. Rajala, *Interpolated measures with bounded density in metric spaces satisfying the curvature-dimension conditions of Sturm*, J. Funct. Anal., 263 (2012), pp. 896–924.

[41] ⸺, *Local Poincaré inequalities from stable curvature conditions on metric spaces*, Calc. Var. Partial Differential Equations, 44 (2012), pp. 477–494.

[42] 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.

[43] ⸺, *Self-improvement of the Bakry-émery condition and Wasserstein contraction of the heat flow in $RCD(K, \infty)$ metric measure spaces*. Preprint, arXiv:1304.0643, 2012.

[44] N. Shanmugalingam, *Newtonian spaces: an extension of Sobolev spaces to metric measure spaces*, Rev. Mat. Iberoamericana, 16 (2000), pp. 243–279.

[45] Z. Shen, *The non-linear Laplacian for Finsler manifolds*, in The theory of Finslerian Laplacians and applications, vol. 459 of Math. Appl., Kluwer Acad. Publ., Dordrecht, 1998, pp. 187–198.

[46] K.-T. Sturm, *On the geometry of metric measure spaces. I*, Acta Math., 196 (2006), pp. 65–131.

[47] ⸺, *On the geometry of metric measure spaces. II*, Acta Math., 196 (2006), pp. 133–177.

[48] C. Villani, *Optimal transport. Old and new*, vol. 338 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 2009.

[49] N. Weaver, *Lipschitz algebras and derivations. II. Exterior differentiation*, J. Funct. Anal., 178 (2000), pp. 64–112.