Scuola Normale Superiore di Pisa

Classe di Scienze Matematiche, Fisiche e Naturali

PhD Thesis in Mathematics

Recent advances in BV and Sobolev spaces in metric measure spaces

Simone Di Marino

Advisor:
Prof. Luigi Ambrosio

Referees:
Prof. Pekka Koskela
Prof. Nicola Gigli

2011 – 2014

Al nonno Domenico
Alla zia Luisa

Contents

Ringraziamenti **i**

Introduction **iii**

1 Preliminary notions **1**

1.1 Topological spaces and Choquet theorem 1
1.1.1 Polish spaces 1
1.1.2 Souslin, Lusin and analytic sets, Choquet theorem 2
1.2 Absolutely continuous curves 3
1.2.1 Reparameterization 4
1.2.2 Equivalence relation in $\mathrm{AC}([0,1];X)$ 5
1.2.3 Non-parametric curves 6
1.3 Slopes, asymptotic Lipschitz constant and upper gradients 8
1.4 Gradient flows of convex and lower semicontinuous functionals 10
1.5 N-functions and Orlicz spaces 10
1.5.1 Orlicz spaces 11
1.5.2 Properties of the (Φ)-norm 13
1.6 Hopf-Lax formula and Hamilton-Jacobi equation 14
1.7 The space $(\mathcal{M}_+(X), W_\Psi)$ and the superposition principle 18
1.8 Γ-convergence 20
1.9 Doubling metric measure spaces and maximal functions 21

2 Duality between p-Modulus and probability measures **23**

2.1 $(p,\mathfrak{m})$-modulus $\mathrm{Mod}_{p,\mathfrak{m}}$ 24
2.2 Plans with barycenter in $L^q(X,\mathfrak{m})$ and $(p,\mathfrak{m})$-capacity 28
2.3 Equivalence between $C_{p,\mathfrak{m}}$ and $\mathrm{Mod}_{p,\mathfrak{m}}$ 31
2.4 Modulus of families of non-parametric curves 34
2.5 Modulus of families of parametric curves 35
2.6 Test plans and their null sets 39
2.7 Weak upper gradients 42

3 Orlicz-Sobolev Spaces **49**
3.1 Variational definition and relaxed energy $\mathcal{F}_v^\Phi$ 50
3.1.1 Gradient flow of $\mathcal{F}_v^\Phi$ 51
3.2 Beppo Levi definition and Φ-weak energy 53
3.3 Proof of equivalence 55
3.4 Consequences of the equivalence theorem 62
3.4.1 Example of Φ-Sobolev function that is not absolutely continuous along almost every curve 64
3.4.2 Ψ doubling: existence of the gradient 65
3.4.3 The relaxation of the integral functional 70
3.5 Other possible definitions 72
3.5.1 Comparison of $\mathcal{F}_c^\Phi$ with $\mathcal{F}_{BL}^\Phi, \mathcal{F}_v^\Phi$ 74
3.5.2 Comparison of $N^{1,\Phi}$ with $\mathcal{F}_{BL}^\Phi, \mathcal{F}_v^\Phi$ 74

4 The spaces BV and $H^{1,1}$ **77**
4.1 BV functions and total variation on Euclidean spaces 77
4.2 Hopf-Lax formula and Hamilton-Jacobi equation 78
4.3 The ∞-Wasserstein distance 81
4.4 Four notions of BV function 81
4.4.1 BV functions in the variational sense 81
4.4.2 BV functions in the relaxed sense 82
4.4.3 Weak-BV functions 85
4.4.4 BV_{BL} functions 89
4.5 Proof of equivalence 89
4.6 Possible definitions for $W^{1,1}(X, \mathsf{d}, \mathfrak{m})$ 94
4.6.1 Comparison between $H^{1,1}$ and $H^{1,\Phi}$ 95

5 Reflexivity and discrete approximation of the gradient **97**
5.1 Reflexivity of $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$, $1 < q < \infty$ 98
5.2 Lower semicontinuity of the slope of Lipschitz functions 106
5.3 Discrete gradients for general metric spaces 110
5.4 Optimality of the Poincaré assumption for the lower semicontinuity of slope. . . 110

6 The p-Weak Gradient depends on p **113**
6.1 Construction of the weight 115
6.2 The p-modulus on curves is trivial for small p 116
6.3 The Muckenhoupt A_p condition for large p 118
6.4 Characterization of the weak gradient on $\mathbb{R}$ 121

7 A definition via derivations and integration by parts **125**
7.1 Sobolev spaces via derivations 125
7.1.1 Derivations 125
7.1.2 Definition via derivations 127
7.2 Equivalence with other definitions 130
7.2.1 $H_v^{1,p} \subseteq W^{1,p}$ 131
7.2.2 $W^{1,p} \subseteq BL^{1,p}$ 132
7.3 BV space via derivations 134

7.3.1 Equivalence of BV spaces 136
7.4 Sobolev Bundle 137

Ringraziamenti

Non potrei non iniziare ringraziando il mio relatore per questi anni di dottorato, Luigi Ambrosio, non solo per il suo acume scientifico e per avermi portato dentro il mondo degli spazi metrici, ma anche per essere riuscito a sopportare i miei innumerevoli ritardi, i miei tempi e, nonostante questo, ad avere sempre tempo per fare due chiacchiere scientifiche e anche personali.

Doverosi ringraziamenti scientifici vanno anche a tutti quelli con cui ho avuto il piacere di scrivere un articolo o anche solo scambiato due parole su quello che stavo facendo: Samuele, Mario, Dario, Guido, Maria, Luigi, Giuseppe S., Andrea M., Matteo, Andrea P., Gareth, Giuseppe B. (anche per aver avuto il coraggio di avermi come esercitatore), Bozhidar, Berardo, Gian Maria. Poi ringrazio anche Nicola e Pekka per aver avuto il compito (e il piacere, spero) di leggere questa tesi e darmi preziosi consigli per il suo miglioramento.

Un sentito ringraziamento va anche alle varie istituzioni che mi hanno ospitato: l'Hausdorff Center of Mathematics a Bonn dal prof. S. Müller, dove grazie ad Antonio ho respirato un po' di Pisa e forse ho imparato qualcosa di analisi nonlineare dalle innumerevoli ore di chiacchierate con Guido. Poi ringrazio l'università di Jyväskylä e la sua calda accoglienza per merito di Tapio, P. Koskela, E. LeDonne, N. Fusco e Changyu, grazie al quale ho avuto il coraggio di mettere giù il capitolo 3 di questa tesi. Infine un ringraziamento ad Andrea ed Emanuele per i freddi giorni a Lipsia, un po' finlandesi... Maledizione, sono ancora nel tunnel dei giochi da tavola!

Tuttavia questo percorso di studi parte da lontano, e i miei inverni pisani oramai stanno per salire a 9. Non avendo scritto ringraziamenti per le altre due tesi, penso che mi tocca farli qui... Grazie a tutti i compagni olimpionici, vecchi e nuovi, in particolare Francesco, Valerio e Claudio, per tutti i bei momenti passati assieme, e perché sicuramente molta della mia passione per la matematica nasce da quei momenti. Nondimeno chiaramente grazie a Gobbino e Dvornicich per tutto, e Morandin e Pernazza per il resto... Grazie a tutti gli abruzzo-princetoniani, in particolare Stefano, Alessandro e Michela; porto sempre con me un grande bagaglio di esperienza da lì. Grazie poi a tutti i "pisani" con cui in questi anni ho condiviso pezzi di strada: senz'altro Gabriele, e non posso elencare tutti i motivi altrimenti esaurirei lo spazio solo per te. Poi Ugo (o Aldo? o Gabbo? o forse Valerio)... vuoi mettere che sugo? Francesco, per la sua inesauribile voglia di matematica e per la sua semplice determinazione. Valerio, per avermi insegnato a sciallarmela un po' di più. Lorenzo, perché c'è sempre. Aldo, per avermi insegnato un po' di fisica. Alessandro, per quel pizzico di naturale follia e di bicarbonato. Nicola, perché i pipistrelli sono nostri amici. Mario, perché se c'è qualcosa da condividere mangiando, non manca mai. Samuele, per i mille discorsi matematici e non. Alessandra, perché simonedimarinopoco. Bea, per il suo entusiasmo coinvolgente. Dario ed Enzo, per i memorabili pacchi. Diego, per i peli. Marco, per la foto al CERN. Andreacarlomarini, perché è l'uomo più buono del mondo, e Andrea per avercelo fatto scoprire. Tutta la cricca abruzzese (con aggiunte), per aver appoggiato e promulgato le arrosticinate: Paolo, Lorenzo, Alessio, Alessandro C., La Chioma, Lin, etc... Luca, Federica e Monty, per le discussioni, anche accese, sui massimi sistemi e non. Il gruppo-capodanno, perché è sempre bello ritrovarsi insieme in un posto random, nonostante siamo sparsi per il mondo...

Ringrazio anche la mia seconda famiglia a Pescara, la parrocchia di S. Lucia, dove ho imparato cosa vuol dire comunità. Di questo ringrazio Don Ermete e Don Nando e in particolare anche Federica, Pierpaolo e Simone: con voi sono cresciuto spiritualmente e come uomo, e spero che le nostre strade, seppur separate, si incrocino ogni tanto. Anche a Pisa sono riuscito a vivere vita di comunità, grazie al coro Mezzogiorno di Fuoco; un mix di coro parrocchiale, gruppo di amici, giovani in cammino e chi più ne ha più ne metta. Non posso non ringraziare Fabio e Gabriella, che tanto impegno mettono per sostenere e dar linfa a questo coro. Nonostante l'orologio (5 anni or sono), continuo a essere in ritardo...

Ringrazio poi gli amici di sempre: Giando, Anastasio, Stefano, Hakkak. Persone che non importa per quanto tempo non ci si sente o non ci si vede, sembra sempre ieri. Grazie anche alle belle chiacchierate e passeggiate avute con i compagni di scuola acquisiti Olmo, Matteo, Alberto, Silvia.

Dulcis in fundo, grazie a tutta la mia famiglia: al nonno Domenico e alla zia Luisa, che mi hanno sempre spronato allo studio, a mamma e papà, per l'esempio, l'amore, il sacrifico, per esserci sempre senza se e senza ma. Paolo, per avermi sopportato quando ero (ero?) un *homo semper ragionis*. I cugini tutti: Giulia, Dario, Nicolò ed Alessandra, perché in famiglia si sta sempre bene. Aurelia e Diego, per avermi accolto, sempre. I cugini acquisiti, per lo spiedo e le cuginolimpiadi. E poi grazie a te, compagna della mia vita: il tuo sorriso è sempre lì ad incoraggiarmi e rasserenarmi, tu sei sempre lì per sostenermi. Mara, ti amo.

Introduction

This thesis is devoted to the topic which I investigated more in my years of PhD: the theory of Sobolev and BV Spaces in Metric Measure Spaces. The first attempts to define spaces of weakly differentiable functions in $\mathbb{R}^n$, what we now call Sobolev Spaces, go back to the beginning of the twentieth century. The theory then reached a mature stage at the end of the '50. We now know that several equivalent definitions can be given, but we refer only to three of them.

(1) The H definition, namely the definition by *relaxation*: $H^{1,p}(\mathbb{R}^n)$ is defined as the closure of $C_c^\infty(\mathbb{R}^n)$ under the norm $\|u\|_{1,p}^p = \|u\|_p^p + \|\nabla u\|_p^p$. Another equivalent definition of $H^{1,p}$ (which will be more useful in our point of view) is simply the domain of finiteness of the relaxation of the functional

$$F_p(u) = \begin{cases} \int_{\mathbb{R}^n} |\nabla u|^p \,\mathrm{d}x & \text{if } u \in C_c^\infty(\mathbb{R}^n) \\ +\infty & \text{otherwise} \end{cases},$$

in the L^p topology. There exists also a local representation for the relaxation of this functional. Indeed, for every $u \in H^{1,p}$ we can define the relaxed gradient ∇u as the weak limit of ∇u_n where $u_n \to u$ in L^p and $\sup_n \|\nabla u_n\|_p < +\infty$, and this limit is unique.

(2) The W definition, namely the definition via an integration by parts formula: $W^{1,p}(\mathbb{R}^n)$ is the set of functions $u \in L^p$ such that there exists a function $g \in L^p(\mathbb{R}^n, \mathbb{R}^n)$, called weak gradient of u, such that

$$\int_{\mathbb{R}^n} \mathrm{div}(\varphi) u \,\mathrm{d}x = -\int_{\mathbb{R}^n} \langle \varphi, g \rangle \,\mathrm{d}x \qquad \forall \varphi \in C_c^\infty(\mathbb{R}^n; \mathbb{R}^n). \tag{1}$$

(3) The BL definition, namely the definition on curves: $BL^{1,p}(\mathbb{R}^n)$ is the set of function u such that $u(\cdot, x_2, \ldots, x_n) : \mathbb{R} \to \mathbb{R}$ has an absolutely continuous representative for $\mathscr{L}^{n-1}$-a.e. $(x_2, \ldots, x_n)$, and $\frac{\partial u}{\partial x_1} \in L^p(\mathbb{R}^n)$, and a similar property holds for every other direction $x_2, \ldots, x_n$.

These three definitions turn out to be equivalent, but the last definition, due to Beppo Levi [60], wasn't taken in great consideration, because it is not frame indifferent and it doesn't seem very useful either. A major improvement of (3), overcoming the lack of frame indifference, is due to Fuglede: we look at the behaviour of the function u not only along the lines $t \mapsto (t, x_2, \ldots, x_n)$ but along all rectifiable curves. In fact in [37] he proved that a function u is in $H^{1,p}$ if and only if the function $u \circ \gamma$ is absolutely continuous for "almost every" curve γ and $|\frac{\mathrm{d}}{\mathrm{d}t} u \circ \gamma| \leq g|\gamma'|$ for some $g \in L^p$: we can summarize this condition by saying that

$$|u(\gamma_1) - u(\gamma_0)| \leq \int_0^1 g(\gamma_t)|\gamma_t'|\,\mathrm{d}t \qquad \text{for "almost every" curve } \gamma.$$

The minimal function that realizes this property is precisely the modulus of the weak gradient $|\nabla u|$. Of course, it is important to recall the concept of negligibility of sets of curves used by Fuglede, but we will come back to this later.

In the last years, since the seminal work of Cheeger [25], a large attention has been devoted to the field of Sobolev Spaces in metric measure spaces $(X, \mathsf{d}, \mathfrak{m})$, see for example [9], [43], [45], [51], [75]; the mild assumptions we require on this metric measure structure is that (X, d) is a separable and complete metric space, and that $\mathfrak{m}$ is finite on bounded sets.

In [25] a major role is played by functions which have an *upper gradient*: we recall that a nonnegative Borel function $g : X \to [0, \infty]$ is an upper gradient for f if

$$|f(\gamma_1) - f(\gamma_0)| \leq \int_\gamma g\,\mathrm{d}s \qquad \forall \gamma \text{ rectifiable,} \tag{2}$$

where γ is said to be rectifiable if $\gamma \in \mathrm{AC}([0,1]; X)$. The set of upper gradients of f is denoted by $UG(f)$. The basic examples of functions which have an upper gradient are Lipschitz functions: given a Lipschitz function f we have that $\mathrm{lip}_a(f) \in UG(f)$, where $\mathrm{lip}_a(f)$ is the asymptotic Lipschitz constant

$$\mathrm{lip}_a(f) = \limsup_{y,z \to x} \frac{|f(y) - f(z)|}{\mathsf{d}(y,z)}. \tag{3}$$

Cheeger's definition of Sobolev Space in metric measure spaces is based upon the H definition, replacing the role of C_c^∞ functions with functions which have an L^p-integrable upper gradient. Another similar definition, used for example in [4], uses only Lipschitz functions with bounded support as "good" functions. Already in [25] these definitions are seen to be equivalent, but under the assumption that $\mathfrak{m}$ is doubling (namely that there exists a constant $C > 1$ such that $\mathfrak{m}(B(x, 2r)) \leq C\mathfrak{m}(B(x, r))$), and a $(1, p)$-Poincaré inequality holds true, that is, there exist constants $\tau > 1$, $C > 0$ such that

$$\min_{m \in \mathbb{R}} ⨍_{B(x_0,r)} |u(x) - m|\,\mathrm{d}\mathfrak{m} \leq C \left(⨍_{B(x_0,\tau r)} |g|^p \right)^{1/p} \qquad \forall g \in UG(f).$$

These two conditions will be often recalled as doubling measure and p-Poincaré assumption.

Up to now, we have already two different definitions of Sobolev Spaces in general metric spaces:

(1a) $H_c^{1,p}(X, \mathsf{d}, \mathfrak{m})$ is the domain of finiteness of the functional

$$\mathcal{F}_c^p(f) = \inf \left\{ \liminf_{n \to \infty} \int_X |g_n|^p\,\mathrm{d}\mathfrak{m} \ : \ f_n \to f \text{ in } L^p,\ g_n \in UG(f_n) \right\};$$

(1b) $H^{1,p}_v(X,\mathsf{d},\mathfrak{m})$ is the domain of finiteness of the functional

$$\mathcal{F}^p_v(f)=\inf\left\{\liminf_{n\to\infty}\int_X|\mathrm{lip}_a(f_n)|^p\,\mathrm{d}\mathfrak{m}\ :\ f_n\to f\ \text{in}\ L^p,\ f_n\in\mathrm{Lip}_0(X,\mathsf{d})\right\}.$$

Unfortunately there is not uniqueness of the weak limit of $\mathrm{lip}_a(f_n)$ as $f_n\to f$ in L^p and so a good definition for the (modulus of the) gradient is to consider, among all possible weak limits, the one with minimal L^p-norm. This will be called the minimal relaxed gradient.

As for the generalization of BL space, we have to introduce some concept of negligibility of set of curves. The original Fuglede approach on $\mathbb{R}^n$ has been generalized in metric measure spaces by Koskela, MacManus in [57] and subsequently by Shanmugalingam in [75], and it relies on the p-modulus $\mathrm{Mod}_{p,\mathfrak{m}}$. We recall that, given a set Γ of curves, we have

$$\mathrm{Mod}_{p,\mathfrak{m}}(\Gamma)=\inf\left\{\int_X f^p\,\mathrm{d}\mathfrak{m}\ :\ f:X\to[0,\infty]\ \text{Borel},\ \int_\gamma f\,\mathrm{d}s\geq 1\ \text{for all}\ \gamma\in\Gamma\right\}. \tag{4}$$

A property is said to hold for $\mathrm{Mod}_{p,\mathfrak{m}}$-almost every curve if the set of curves on which it fails is $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible. Another relevant notion of negligibility of curves is obtained via the so-called q-plans (introduced in [10] for $q=2$ and then in [9] for a generic $q\in(1,+\infty)$), that are probability measures on $\mathrm{AC}([0,1];X)$, concentrated on AC^q, such that $(\mathrm{e}_t)_\sharp\boldsymbol{\pi}\leq C(\boldsymbol{\pi})\mathfrak{m}$ for some $C(\boldsymbol{\pi})\geq 0$. Then a Borel set $\Gamma\subset C([0,1];X)$ is said p-negligible if $\boldsymbol{\pi}(\Gamma)=0$ for every q-plan $\boldsymbol{\pi}$. Now we are ready to state two more definitions of Sobolev Spaces:

(3a) $N^{1,p}(X,\mathsf{d},\mathfrak{m})$ is the set of function f such that there exists $g\in L^p(X,\mathfrak{m})$ such that

$$|f(\gamma_1)-f(\gamma_0)|\leq\int_0^1 g(\gamma_t)|\dot\gamma_t|\,\mathrm{d}t\qquad\text{for }\mathrm{Mod}_{p,\mathfrak{m}}\text{-almost every curve }\gamma.$$

The minimal such g is called the minimal p-upper gradient.

(3b) $BL^{1,p}(X,\mathsf{d},\mathfrak{m})$ is the set of function f such that there exists $g\in L^p(X,\mathfrak{m})$ such that

$$|f(\gamma_1)-f(\gamma_0)|\leq\int_0^1 g(\gamma_t)|\dot\gamma_t|\,\mathrm{d}t\qquad\text{for }p\text{-almost every curve }\gamma.$$

The minimal such g is called the minimal p-weak upper gradient.

In [9] it is proved that (1a), (1b), (3a), (3b) are equivalent; it is important to underline that we have not only equivalence of spaces, but also equality for the minimal gradients. From now on we will refer to any of this equivalent Sobolev Spaces as $W^{1,p}(X,\mathsf{d},\mathfrak{m})$.

In this thesis we will describe and improve the results present in the articles [4]–[6], [34] about Sobolev and BV spaces written in these years, as well as new unpublished results.

The first Chapter contains some preliminary results needed in the the rest of the thesis; the other chapters are devoted to specific parts of the theory, usually based on one of the articles. We give briefly an outline here and then we introduce every specific Chapter with a little more specifics.

- Chapter 2: in [6], in collaboration with L. Ambrosio and G. Savaré, we look more closely to the relation between $\mathrm{Mod}_{p,\mathfrak{m}}$-negligibility and p-negligibility (also at the level of sets of measures), providing a dual formulation of $\mathrm{Mod}_{p,\mathfrak{m}}$. We obtain also another proof that $N^{1,p}=BL^{1,p}$, exploiting the structural properties of the set where the upper gradient inequality (2) fails.

- Chapter 3: we generalize the equivalence theorem in [9] to the Orlicz-Sobolev case.
- Chapter 4: in [5], in collaboration with L. Ambrosio, we prove the analogous of the equivalences stated before in the context of BV spaces.
- Chapter 5: in [4], in collaboration with L. Ambrosio and M. Colombo, we prove that under the mild assumption that (X, d) is a doubling space, the space $W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ is reflexive, extending the result in [25], where the author proves it under doubling and Poincaré assumptions.
- Chapter 6: in [34], in collaboration with G. Speight, we answer positively to the question "does $|\nabla f|_{w,p}$ depends on p?", showing for every $\alpha > 0$ the existence of a measure μ on $\mathbb{R}$, absolutely continuous with respect to $\mathscr{L}^1$, such that for any Lipschitz function f we have $|\nabla f|_{p,\mu} = 0$ for $p \le 1+\alpha$ while $|\nabla f|_{p,\mu} = |f'|$ for $p > 1+\alpha$.
- Chapter 7: we extend the W definition of Sobolev Space to a very general metric setting, with an integration by parts formula made up with Weaver's derivations. This latter chapter contains also an (abstract) characterization of the weak gradient in Hilbert spaces endowed with a general measure, extending a previous result in [22].

Negligibility of set of curves

The notion of p-modulus $\mathrm{Mod}_p(\Gamma)$ for a family Γ of curves has been introduced by Beurling and Ahlfors in [1] and then it has been deeply studied by Fuglede in [37], as we recalled, also in connection with the theory of Sobolev Spaces in $\mathbb{R}^n$. It is obvious that the definition of p-Modulus (4) (as the notion of length) is parametric-free, because the curves are involved in the definition only through the curvilinear integral $\int_\gamma f$. Furthermore, as in [37], one can even go a step further, realizing that this curvilinear integral can be written as

$$\int_X f \,\mathrm{d}J\gamma,$$

where $J\gamma$ is a positive finite measure in X, the image under γ of the measure $|\dot\gamma|\mathscr{L}^1 \llcorner I$, namely

$$J\gamma(B) = \int_{\gamma^{-1}(B)} |\dot\gamma_t| \,\mathrm{d}t \qquad \forall B \in \mathscr{B}(X) \tag{5}$$

(here $\mathscr{L}^1 \llcorner I$ stands for the Lebesgue measure on I). It follows that one can define in a similar way the notion of p-modulus for families of measures in X.

In more recent times, Koskela-Mac Manus [57] and then Shanmugalingham [75] used the p-modulus to define the notion of p-weak upper gradient for a function f, while, even more recently, Ambrosio, Gigli and Savaré introduced another notion of weak upper gradient, based on suitable classes of probability measures on curves, described more in detail in the final section of this chapter.

Since the axiomatization in [11] is quite different and sensitive to parameterization, it is a surprising fact that the two approaches lead essentially to the same Sobolev space theory (see Remark 5.12 of [11]). We say essentially because, strictly speaking, the axiomatization of [11] is invariant (unlike Fuglede's approach) under modification of f in $\mathfrak{m}$-negligible sets and thus provides only Sobolev regularity and not absolute continuity along almost every curve;

however, choosing properly representatives in the Lebesgue equivalence class, the two Sobolev spaces can be identified.

With the goal of understanding deeper connections between the $\mathrm{Mod}_{p,\mathfrak{m}}$ and the probabilistic approaches, we show in Chapter 2 that the theory of p-modulus has a "dual" point of view, based on suitable probability measures $\boldsymbol{\pi}$ in the space of curves; the main difference with respect to [11] is that, as it should be, the curves here are non-parametric, namely $\boldsymbol{\pi}$ should be rather thought as measures in a quotient space of curves. Actually, this and other technical aspects (also relative to tightness, since much better compactness properties are available at the level of measures) are simplified if we consider p-modulus of families of measures in $\mathcal{M}_+(X)$ (the space of all nonnegative and finite Borel measures on X), rather than p-modulus of families of curves: if we have a family Γ of curves, we can consider the family $\Sigma = J(\Gamma)$ and derive a representation formula for $\mathrm{Mod}_{p,\mathfrak{m}}(\Gamma)$, see Section 2.4. Correspondingly, $\boldsymbol{\pi}$ will be a measure on the Borel subsets of $\mathcal{M}_+(X)$.

Assuming only that (X, d) is complete and separable and $\mathfrak{m}$ is finite, we prove in Theorem 2.3.1 that for all Borel sets $\Sigma \subset \mathcal{M}_+(X)$ (and actually in the more general class of Souslin sets) the following duality formula holds:

$$\big[\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma)\big]^{1/p} = \sup_{\boldsymbol{\eta}} \frac{\boldsymbol{\eta}(\Sigma)}{c_q(\boldsymbol{\eta})} = \sup_{\boldsymbol{\eta}(\Sigma)=1} \frac{1}{c_q(\boldsymbol{\eta})}, \qquad \frac{1}{p} + \frac{1}{q} = 1. \tag{6}$$

Here the supremum in the right hand side runs in the class of Borel probability measures $\boldsymbol{\eta}$ in $\mathcal{M}_+(X)$ with barycenter in $L^q(X, \mathfrak{m})$, so that

$$\text{there exists } g \in L^q(X, \mathfrak{m}) \text{ s.t.} \qquad \int \mu(A)\,\mathrm{d}\boldsymbol{\eta}(\mu) = \int_A g\,\mathrm{d}\mathfrak{m} \qquad \forall A \in \mathscr{B}(X);$$

the constant $c_q(\boldsymbol{\eta})$ is then defined as the $L^q(X, \mathfrak{m})$ norm of the "barycenter" g. A byproduct of our proof is the fact that $\mathrm{Mod}_{p,\mathfrak{m}}$ is a Choquet capacity in $\mathcal{M}_+(X)$, see Theorem 2.3.1. In addition, we can prove in Corollary 2.3.2 existence of maximizers in (6) and obtain out of this necessary and sufficient optimality conditions, both for $\boldsymbol{\eta}$ and for the minimal f involved in the definition of p-modulus. See also Remark 2.1.3 for a simple application of these optimality conditions involving pairs (μ, f) on which the constraint is saturated, namely $\int_X f\,\mathrm{d}\mu = 1$.

In the second part of Chapter 2 we show how the basic duality result of the first part can be read in terms of measures and moduli in spaces of curves. For non-parametric curves this is accomplished in Section 2.4, mapping curves in X to measures in X with the canonical map J in (5); in this case, the condition of having a barycenter in $L^q(X, \mathfrak{m})$ becomes

$$\left| \int \int_0^1 f(\gamma_t)|\dot\gamma_t|\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi}(\gamma) \right| \le C\|f\|_{L^p(X,\mathfrak{m})} \qquad \forall f \in \mathrm{C}_b(X). \tag{7}$$

Section 2.5 is devoted instead to the case of parametric curves, where the relevant map curves-to-measures is

$$M\gamma(B) := \mathscr{L}^1(\gamma^{-1}(B)) \qquad \forall B \in \mathscr{B}(X).$$

In this case the condition of having a *parametric* barycenter in $L^q(X, \mathfrak{m})$ becomes

$$\left| \int \int_0^1 f(\gamma_t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi}(\gamma) \right| \le C\|f\|_{L^p(X,\mathfrak{m})} \qquad \forall f \in \mathrm{C}_b(X). \tag{8}$$

The parametric barycenter can of course be affected by reparameterizations; a key result, stated in Theorem 2.5.5, shows that suitable reparameterizations improve the parametric barycenter from $L^q(X,\mathfrak{m})$ to $L^\infty(X,\mathfrak{m})$. Then, in Section 2.6 we discuss the notion of null set of curves according to [11] and [9] (where (8) is strengthened by requiring $\left|\int f(\gamma_t)\,\mathrm{d}\boldsymbol{\pi}(\gamma)\right| \leq C\|f\|_{L^1(X,\mathfrak{m})}$ for all t, for some C independent of t) and, under suitable invariance and stability assumptions on the set of curves, we compare this notion with the one based on p-modulus. Eventually, in Section 2.7 we use there results to prove that if a Borel function $f : X \to \mathbb{R}$ has a continuous representative along a collection Γ of the set $\mathrm{AC}^\infty([0,1];X)$ of the Lipschitz parametric curves with $\mathrm{Mod}_{p,\mathfrak{m}}\big(M(\mathrm{AC}^\infty([0,1];X)\setminus\Gamma)\big) = 0$, then it is possible to find a distinguished $\mathfrak{m}$-measurable representative $\tilde f$ such that $\mathfrak{m}(\{f \neq \tilde f\}) = 0$ and $\tilde f$ is absolutely continuous along $\mathrm{Mod}_{p,\mathfrak{m}}$-a.e.-nonparametric curve. By using these results to provide a more direct proof of the equivalence of the two above mentioned notions of weak upper gradient, where different notions of null sets of curves are used to quantify exceptions to (2).

Orlicz-Sobolev spaces in metric measure spaces

In Chapter 3 we generalize the equivalence result in [9] to the Orlicz-Sobolev case. Orlicz spaces are a natural generalization of Lebesgue spaces $L^p(X,\mathfrak{m})$, where the role of the function $t \mapsto t^p$ is replaced by an even convex function $\Phi : \mathbb{R} \to [0,\infty]$ such that $\Phi(0) = 0$. Then one can define the norm

$$\|g\|_{(\Phi),\mathfrak{m}} = \left\{ \int_X fg\,\mathrm{d}\mathfrak{m} \ : \ \int_X \Psi(f) \leq 1 \right\},$$

where Ψ is the convex conjugate of Φ. The Orlicz space $L^\Phi(X,\mathfrak{m})$ is simply the set of $\mathfrak{m}$-measurable functions which have finite (Φ)-norm.

The Sobolev-Orlicz space $W^{1,\Phi}(X,\mathsf{d},\mathfrak{m})$ is, roughly speaking, the set of functions $f \in L^1$ such that $|\nabla f| \in L^\Phi$. In the Euclidean case this makes sense since there is an a priori gradient (namely the distributional gradient), while in general metric measure spaces this is no more possible, and thus one can think of several different definitions, as in the $W^{1,p}$ case. In the literature, the dominant approach is the Newtonian space one, based upon the Φ-modulus: this space is called $N^{1,\Phi}(X,\mathsf{d},\mathfrak{m})$, and the definition follows precisely the one we introduced before for the homogeneous case: a function $g \in L^\Phi(X,\mathfrak{m})$ is a Φ-upper gradient for f if the upper gradient inequality (2) holds for $\mathrm{Mod}_{(\Phi),\mathfrak{m}}$-almost every curve γ (see [2], [78] and subsequently [65], [68] for the generalization to Banach and quasi-Banach function spaces).

Here we give different definitions, in the spirit of (1a) and (3b): first we define the space $H^{1,\Phi}_v(X,\mathsf{d},\mathfrak{m})$ as the domain of finiteness of the following functional

$$\mathcal{F}^\Phi_v(f) = \inf\left\{ \liminf_{n\to\infty} \|\mathrm{lip}_a(f_n)\|_{(\Phi),\mathfrak{m}} \ : \ f_n \to f \text{ in } L^1(X,\mathfrak{m}),\ (f_n) \subset \mathrm{Lip}_0(X,\mathsf{d}) \right\}.$$

The definition of $BL^{1,\Phi}(X,\mathsf{d},\mathfrak{m})$ is a little more subtle: $f \in BL^{1,\Phi}$ if there exists a constant $E \geq 0$ such that for every finite Radon measure $\boldsymbol{\pi}$ on $\mathrm{AC}([0,1];X)$ such that $(\mathrm{e}_t)_\sharp\boldsymbol{\pi} \leq C\mathfrak{m}$ for some $C(\boldsymbol{\pi}) \geq 1$, we have

$$\int |f(\gamma_0) - f(\gamma_1)|\,\mathrm{d}\boldsymbol{\pi} \leq E \cdot C(\boldsymbol{\pi}) \cdot \int_0^1 \|\dot\gamma_t\|_{\Psi,\boldsymbol{\pi}}\,\mathrm{d}t. \tag{9}$$

Notice that it is necessary to have the weaker integral form in order to deal with a generic N-function Φ (see also the BV case below for comparison). The main result of this chapter is that $BL^{1,\Phi} = H^{1,\Phi}_v$ and moreover $\mathcal{F}^\Phi_v(f) = \mathcal{F}^\Phi_{BL}(f)$, where $\mathcal{F}^\Phi_{BL}(f)$ is the minimal constant E such that (9) holds. That is, Lipschitz functions are dense in energy in $BL^{1,\Phi}$.

The proof that $BL^{1,\Phi}(X,\mathsf{d},\mathfrak{m})$ includes $H_v^{1,\Phi}(X,\mathsf{d},\mathfrak{m})$ and that $\mathcal{F}_{BL}^{\Phi} \leq \mathcal{F}_v^{\Phi}$ is not too difficult. Notice that proving equivalence of the two definitions amounts to passing from a (quantitative) information on the behavior of the function along random curves to the construction of a Lipschitz approximation. Remarkably, this result does not rely on doubling and Poincaré assumptions on the metric measure structure. As in [9] (based essentially on ideas come from [11], dealing with the case of $W^{1,2}$ Sobolev spaces), the proof is not really constructive: it is obtained with optimal transportation tools and using the theory of gradient flows of convex and lower semicontinuous functionals in Hilbert spaces. Specifically, in our case we shall use the gradient flow in $L^2(X,\mathfrak{m})$ of the functional $f \mapsto \mathcal{F}_v^{\Phi}(f)$. We will not enter in further details of the proof; a summary with the main ideas can be found at the beginning of the chapter.

A consequence of this equivalence theorem is that $f \circ \gamma$ is BV along Φ-almost every curve whenever $f \in BL^{1,\Phi}$, but we can't expect more, as shown by the example in Subsection 3.4.1, where a characteristic function is proved to belong to $H_v^{1,\Phi}$, where $\Phi(t) = (t+1)\log(t+1) - t$.

In Section 3.4.2 the easier case when Ψ is doubling is treated, i.e. when there exists $C > 1$ such that $\Psi(2x) \leq C\Psi(x)$ for all $x \in \mathbb{R}$. In this case we have that $f \circ \gamma$ is $W^{1,1}$ along Φ-almost every curve, and also there exists a well defined gradient $|\nabla f|_{w,\Phi}$; moreover under this assumption, in Theorem 3.5.6 we prove that

$$H_v^{1,\Phi}(X,\mathsf{d},\mathfrak{m}) = BL^{1,\Phi}(X,\mathsf{d},\mathfrak{m}) = N^{1,\Phi}(X,\mathsf{d},\mathfrak{m}),$$

along with the equality between $|\nabla f|_{w,\Phi}$ and the minimal Φ-upper gradient. Then a strict relationship between $W^{1,\Phi}(X,\mathsf{d},\mathfrak{m})$ and the L^1-relaxation of the functional

$$G_\Phi(f) = \begin{cases} \int_X \Phi(\mathrm{lip}_a(f))\,\mathrm{d}\mathfrak{m} & \text{if } f \in \mathrm{Lip}_0(X,\mathsf{d}) \\ +\infty & \text{otherwise} \end{cases}$$

is made clear, proving a representation formula that involves the Φ-weak gradient $|\nabla f|_{w,\Phi}$. All these results in the case in which Ψ is doubling are achieved thanks to a Mazur-type lemma for weak-$*$ convergence in $L^\Phi(X,\mathfrak{m})$, contained in Lemma 3.4.3.

Functions of Bounded Variation in metric measure spaces

In this chapter we provide a positive answer to a problem raised in [9]. Recall that, following the notion of BV function given in [67], a function $f \in L^1(X,\mathfrak{m})$ belongs to $BV_*(X,\mathsf{d},\mathfrak{m})$ if there exist Lipschitz functions with bounded support f_n convergent to f in $L^1(X,\mathfrak{m})$ such that

$$\limsup_{n\to\infty} \int_X \mathrm{lip}_a(f_n)\,\mathrm{d}\mathfrak{m} < \infty.$$

By localizing this construction one can define

$$|Df|_*(A) := \inf\left\{\liminf_{n\to\infty} \int_A \mathrm{lip}_a(f_n)\,\mathrm{d}\mathfrak{m} :\ (f_n) \subset \mathrm{Lip}_{\mathrm{loc}}(A),\ f_h \to f \text{ in } L^1(A)\right\}, \tag{10}$$

for any open set $A \subseteq X$. In [67], it is proved (with minor variants in the definition, namely the convergence is in L^1_{loc} and the asymptotic Lipschitz constant is replaced by the slope) that this set function is the restriction to open sets of a finite Borel measure, called total variation measure and, following basically the same strategy, we will extend this result to our more general setup.

Then we consider a new definition of BV function in the spirit of the theory of weak, rather than relaxed, upper gradients [57], [75] that we already recalled. Without entering in this introduction in too many technical details, we say that $f \in w-BV(X, \mathsf{d}, \mathfrak{m})$ if there exists a finite Borel measure μ with this property: for any probability measure $\boldsymbol{\pi}$ on $\mathrm{Lip}([0,1];X)$ the function $t \mapsto f \circ \gamma_t$ belongs to $BV(0,1)$ for $\boldsymbol{\pi}$-a.e. curve γ_t and

$$\frac{1}{C(\boldsymbol{\pi})\|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi})}} \int \gamma_\sharp |D(f\circ\gamma)| \,\mathrm{d}\boldsymbol{\pi} \le \mu.$$

Here $C(\boldsymbol{\pi})$ is the least constant C such that $(e_t)_\sharp \boldsymbol{\pi} \le C\mathfrak{m}$ for all $t \in [0,1]$, where $e_t(\gamma) := \gamma_t$ are the evaluation maps at time t. The smallest measure μ with this property will be denoted by $|Df|_w$.

We will prove that these two definitions are equivalent, and we have also $|Df|_* = |Df|_w$; this result extends also to intermediate spaces, such as the one considered by Cheeger, and to even weaker definitions, in the spirit of the BL definition in the Orlicz case.

The proof follows the same lines of the equivalence theorem in Chapter 3. We recall that the functional of which we take the gradient flow, namely $f \mapsto |Df|_*(X)$, is also called total variation flow in image processing [16]. We will not enter into details of the proof here but we just mention that some properties of BV functions readily extend to the more general framework considered in this chapter. For instance, the coarea formula

$$|Df|_* = \int_0^\infty |D\chi_{\{f>t\}}|_* \,\mathrm{d}t + \int_{-\infty}^0 |D\chi_{\{f<t\}}|_* \,\mathrm{d}t$$

can be achieved following *verbatim* the proof in [67]. On the other hand, more advanced facts, as the decomposition alone curves in absolutely continuous and singular part of the derivative (see [7, Section 3.11]), seem to be open at this level of generality: for instance, Example 4.5.4 shows that, in contrast to what happens in Euclidean metric measure spaces (here the supremum is understood in the lattice of measures), the measure

$$\sup_{\boldsymbol{\pi}} \frac{1}{C(\boldsymbol{\pi})\|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi})}} \int \gamma_\sharp |D^a(f\circ\gamma)| \,\mathrm{d}\boldsymbol{\pi},$$

which is easily seen to be smaller than the absolutely continuous part of $|Df|_w$, may be strictly smaller.

Reflexivity and discrete approximation of the gradient

In [25], Cheeger investigated the fine properties of Sobolev functions on metric measure spaces, with the main aim of providing generalized versions of Rademacher's theorem and, along with it, a description of the cotangent bundle. Assuming that the Polish metric measure structure $(X, \mathsf{d}, \mathfrak{m})$ is doubling and satisfies a Poincaré inequality (see Definitions 1.9.1 and 5.2.1 for precise formulations of these structural assumptions) he proved that the Sobolev spaces are reflexive and that the q-power of the slope is $L^q(X,\mathfrak{m})$-lower semicontinuous, namely

$$f_h,\, f \in \mathrm{Lip}(X),\ \int_X |f_h - f|^q \,\mathrm{d}\mathfrak{m} \to 0 \quad \Longrightarrow \quad \liminf_{h\to\infty} \int_X |\nabla f_h|^q \,\mathrm{d}\mathfrak{m} \ge \int_X |\nabla f|^q \,\mathrm{d}\mathfrak{m}. \tag{11}$$

Here the slope $|\nabla f|$, also called local Lipschitz constant, is defined by

$$|\nabla f|(x) := \limsup_{y\to x} \frac{|f(y)-f(x)|}{\mathsf{d}(y,x)}.$$

These results come also as a byproduct of a generalized Rademacher's theorem, which can be stated as follows: there exist an integer N, depending on the doubling and Poincaré constants, a Borel partition $\{X_i\}_{i\in I}$ of X and Lipschitz functions f^i_j, $1 \le j \le N(i) \le N$, with the property that for all $f \in \mathrm{Lip}(X)$ it is possible to find Borel coefficients c^i_j, $1 \le j \le N$, uniquely determined $\mathfrak{m}$-a.e. on X_i, satisfying

$$\Big|\nabla(f - \sum_{j=1}^{N(i)} c^i_j(x) f^i_j)\Big|(x) = 0 \qquad \text{for } \mathfrak{m}\text{-a.e. } x \in X_i. \tag{12}$$

It turns out that the family of norms on $\mathbb{R}^{N(i)}$

$$\|(\alpha_1,\ldots,\alpha^i_{N(i)})\|_x := \Big|\nabla \sum_{j=1}^{N(i)} \alpha_j f^i_j\Big|(x)$$

indexed by $x \in X_i$ satisfies, thanks to (12),

$$\|(c^i_1(x),\ldots,c^i_{N(i)}(x))\|_x = |\nabla f|(x) \qquad \text{for } \mathfrak{m}\text{-a.e. } x \in X_i.$$

Therefore, this family of norms provides the norm on the cotangent bundle on X_i. Since $N(i) \le N$, using for instance John's lemma one can find Hilbertian equivalent norms $|\cdot|_x$ with bi-Lipschitz constant depending only on N. This leads to an equivalent (but not canonical) Hilbertian norm and then to reflexivity. In this chapter we aim mostly at lower semicontinuity and reflexivity: we recover the latter (and separability as well) without assuming the validity of the Poincaré inequality and replacing the doubling assumption on $(X,\mathsf{d},\mathfrak{m})$ with a weaker assumption, namely the geometric doubling of $(\mathrm{supp}\,\mathfrak{m},\mathsf{d})$.

In particular we prove that the Sobolev space $W^{1,q}(X,\mathsf{d},\mathfrak{m})$ is reflexive when $1 < q < \infty$, $(\mathrm{supp}\,\mathfrak{m},\mathsf{d})$ is separable and doubling, and $\mathfrak{m}$ is finite on bounded sets. Instead of looking for an equivalent Hilbertian norm (whose existence is presently known only if the metric measure structure is doubling and the Poincaré inequality holds), we rather look for a discrete scheme, involving functionals $\mathcal{F}_\delta(f)$ of the form

$$\mathcal{F}_\delta(f) = \sum_i \frac{1}{\delta^q} \sum_{A^\delta_j \sim A^\delta_i} |f_{\delta,i} - f_{\delta,j}|^q \mathfrak{m}(A^\delta_i).$$

Here A^δ_i is a well chosen decomposition of $\mathrm{supp}\,\mathfrak{m}$ on scale δ, $f_{\delta,i} = \fint_{A^\delta_i} f$ and the sum involves cells A^δ_j close to A^δ_i, in a suitable sense. This strategy is very close to the construction of approximate q-energies on fractal sets and more general spaces, see for instance [52], [76].

It is fairly easy to show that any Γ-limit point $\mathcal{F}_0$ of $\mathcal{F}_\delta$ as $\delta \to 0$ satisfies

$$\mathcal{F}_0(f) \le c(c_D,q) \int_X \mathrm{lip}_a(f)^q \,\mathrm{d}\mathfrak{m} \qquad \text{for all Lipschitz } f \text{ with bounded support}, \tag{13}$$

where c_D is the doubling constant of (X, d) (our proof gives $c(c_D, q) \leq 6^q c_D^3$). More delicate is the proof of lower bounds of $\mathcal{F}_0$, which uses a suitable discrete version of the weak upper gradient property and leads to the inequality

$$\frac{1}{4^q}\int_X |\nabla f|_{w,q}^q \,\mathrm{d}\mathfrak{m} \leq \mathcal{F}_0(f) \qquad \forall f \in W^{1,q}(X, \mathsf{d}, \mathfrak{m}). \tag{14}$$

Combining (13), (14) and the equivalence of weak gradients gives

$$\frac{1}{4^q}\int_X |\nabla f|_{w,q}^q \,\mathrm{d}\mathfrak{m} \leq \mathcal{F}_0(f) \leq c(c_D, q)\int_X |\nabla f|_{w,q}^q \,\mathrm{d}\mathfrak{m} \qquad \forall f \in W^{1,q}(X, \mathsf{d}, \mathfrak{m}).$$

The discrete functionals $\mathcal{F}_\delta(f) + \sum_i |f_{\delta,i}|^q \mathfrak{m}(A_i^\delta)$ describe L^q norms in suitable discrete spaces, hence they satisfy the Clarkson inequalities; these inequalities (which reduce to the parallelogram identity in the case $q = 2$) are retained by the Γ-limit point $\mathcal{F}_0 + \|\cdot\|_q^q$. This leads to an equivalent uniformly convex norm in $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$, and therefore to reflexivity. As a byproduct one obtains density of bounded Lipschitz functions in $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$ and separability. In this connection, notice that the results of [11], [9] provide, even without a doubling assumption, a weaker property (but still sufficient for some applications), the so-called density in energy; on the other hand, under the assumptions of [25] one has even more, namely density of Lipschitz functions in the Lusin sense.
Notice however that $\mathcal{F}_0$, like the auxiliary Hilbertian norms of [25], is not canonical: it might depend on the decomposition A_i^δ and we don't expect the whole family $\mathcal{F}_\delta$ to Γ-converge as $\delta \to 0^+$. We also provide an example showing that reflexivity may fail if the metric doubling assumption is dropped.

In the final part of the chapter we prove also (11), following in large part the scheme of [25] (although we get the result in a more direct way, without an intermediate result in length spaces). In particular we need the Poincaré inequality to establish the bound

$$|\nabla f| \leq C\,|\nabla f|_{w,q} \qquad \text{for any Lipschitz function } f \text{ with bounded support,}$$

which, among other things, prevents $|\nabla f|_{w,q}$ from being trivial.

The p-weak gradient depends on p

Another important issue in the theory is whether the weak gradient depends on p or not (at least for Lipschitz functions). For example it was known [25] that under p-Poincaré and doubling assumptions on the measure, the weak gradient equals $|\nabla u|$ for a Lipschitz function u and so, it doesn't depend on the exponent, at least for $q \geq p$. Another recent result by Gigli and Han [40] in this direction is that in every $RCD(K, \infty)$ spaces an even stronger property holds true: if $f \in W^{1,p}$ has a weak gradient $|\nabla f|_{w,p} \in L^q(X, \mathfrak{m})$ then $f \in W^{1,q}$ and $|\nabla f|_{w,q} = |\nabla f|_{w,p}$ (the result holds even if $f \in BV$ and $|Df| = |\nabla f|_{w,1}\mathfrak{m}$).

In Chapter 6, based on the results of [34], given α we find a result in the opposite direction: we construct a weighted Lebesgue measure on $\mathbb{R}^n$ for which the family of non constant curves has p-modulus zero for $p \leq 1 + \alpha$ but the weight is a Muckenhoupt A_p weight for $p > 1 + \alpha$. In particular, the p-weak gradient is trivial for small p but non trivial for large p. We also give a full description of the p-weak gradient for any locally finite Borel measure on $\mathbb{R}$.

This is the main theorem:

Theorem 1 *Let $n \in \mathbb{N}$ and $\alpha > 0$. Then there exists a Borel function $w\colon \mathbb{R}^n \to \mathbb{R}^+$ such that the measure $\mu := w\mathcal{L}^n$ is doubling and:*

- *For $p \leq 1+\alpha$ we have $\mathrm{Mod}_{p,\mu}(\Gamma_c) = 0$ where Γ_c is the family of non constant absolutely continuous curves in $\mathbb{R}^n$. This implies that the p-weak gradient on $(\mathbb{R}^n, |\cdot|, \mu)$ is identically zero for every function.*

- *For $p > 1+\alpha$ the function w is a Muckenhoupt A_p-weight. This implies that a weak p-Poincaré inequality holds; it follows that the p-weak gradient on $(\mathbb{R}^n, |\cdot|, \mu)$ agrees with the slope for Lipschitz functions.*

The simple structure of curves in $\mathbb{R}$ gives rise to a simple description of the p-weak gradient with respect to each measure. In Theorem 6.4.2 we show that, for any locally finite Borel measure on $\mathbb{R}$ and $p > 1$, the corresponding p-weak gradient of a Lipschitz function $f\colon \mathbb{R} \to \mathbb{R}$ is, at almost every point x, either equal to zero or equal to $|f'(x)|$. Roughly speaking, the points where the p-weak gradient is non zero are those points which have a neighborhood that, when considered as a set containing a single curve, has positive p-modulus.

A definition via derivation and integration by parts

In Chapter 7 we want to give a definition of Sobolev spaces via integration by parts formula, in the spirit of (2) in euclidean spaces; the role of the vector field will be played by derivations.

The derivations were introduced in the seminal papers by Weaver [80], and then in more recent times widely used in the Lipschitz theory of metric spaces, for example in connection with Rademacher theory for metric spaces, but also as a generalization of sections of the tangent space [14], [15], [39], [73], [74]. Here we see that the derivations are also powerful tools in the Sobolev theory, as already point out in [39]. A derivation, in our definition, is simply a linear map $\boldsymbol{b} : \mathrm{Lip}_0(X, \mathsf{d}) \to L^0(X, \mathfrak{m})$ such that the Liebniz rule holds and it has the locality property $|\boldsymbol{b}(f)| \leq g \cdot \mathrm{lip}_a(f)$ for some $g \in L^0(X, \mathfrak{m})$. Now we simply say that $f \in L^p$ is a function in $W^{1,p}$ if there is a $\mathrm{Lip}_b(X)$-linear map L_f such that integration by part holds:

$$\int_X L_f(\boldsymbol{b})\, \mathrm{d}\mathfrak{m} = -\int_X f \cdot \mathrm{div}\, \boldsymbol{b}\, \mathrm{d}\mathfrak{m} \qquad \forall \boldsymbol{b} \in \mathrm{Der}^{q,q},$$

where $\mathrm{Der}^{q,q}$ is the subset of derivation for which $|\boldsymbol{b}|, \mathrm{div}\, \boldsymbol{b} \in L^q(X, \mathfrak{m})$.

We will see that it is well defined a proper "differential" $df : \mathrm{Der}^{q,q} \to L^1$, and so it is possible to provide also a notion of modulus of the gradient $|\nabla f|$ in such a way that $|df(\boldsymbol{b})| \leq |\nabla f| \cdot |\boldsymbol{b}|$; in Section 7.2 we see that this notion coincides with all the other (equivalent) notion of modulus of the gradient given in [9] (namely (1a), (1b), (3a) and (3b)), and in particular there is also identification of the Sobolev spaces.

The easy part is the inclusion of the Sobolev Space obtained via relaxation of the asymptotic Lipschitz constant into the one defined by derivations. The other inclusion uses the fact that q-plans, namely measures on the space of curves with some integrability assumptions, induces derivations thanks to the basic observation that, even in metric spaces, we can always take the derivative of Lipschitz functions along absolutely continuous curves; this observation has already been used in [15], [73], [74] to find correlation between the differential structure of (X, d) and the structure of measures on the set of curves (a peculiar role is played by Alberti representation).

In Section 7.3.1 we extend this equivalence to the BV space, using the results in [5].

In the last section, we eventually apply this new definition in order to find an abstract characterization of the weak gradient for C^1 functions, when X is a Banach space. This characterization has already been obtained in [22] for $\mathbb{R}^n$, while in Theorem 6.4.2 we re-obtain it in the one dimensional case. However here we employ a different strategy and the proof will follow the line of [3], where a similar bundle (the differentiability bundle) is constructed in order to find the directions of "almost everywhere" differentiability of Lipschitz functions given an arbitrary measure $\mathfrak{m}$ in $\mathbb{R}^n$.

It is important to remark that the differentiability bundle is always contained in the Sobolev bundle S_p that we construct; this link is not at all trivial and we believe that this connection has to be inspected deeply.

Other works

Here we give a short summary of the other research made during the PhD studies. We briefly report the results obtained and we refer to the original papers for a complete treatment of the problems and the relevant and related literature.

Equality between Monge and Kantorovich multi marginal problems with Coulomb cost

In [30], in collaboration with M. Colombo, we generalize a previous result of Pratelli [71] to the multimarginal case. Given a probability measure μ on a Polish metric space (X, d), and integer $n \geq 2$ and a lower semicontinuous cost function $c : X^n \to [0, \infty]$, we introduce the following infimum problems:

$$(K) := \inf \Big\{ \int_{X^n} c(x_1, \dots, x_n) \, \mathrm{d}\pi : \pi \in \mathscr{P}((X)^n), \, (e_i)_\sharp \pi = \mu \; \forall i \in \{1, ..., n\} \Big\};$$

$$(M) = \inf \Big\{ \int_{\mathbb{R}^d} c(x, T(x), ..., T^{(n-1)}(x)) \, d\mu(x) : \, T_\sharp \mu = \mu, \, T^{(n)} = Id \Big\}.$$

It is obvious that $(K) \leq (M)$ since given an admissible T in (M) we have that $\pi = (Id, T, T^{(2)}, \dots, T^{(n-1)})_\sharp \mu$ is admissible in (K) and has the same cost. We prove the following:

Theorem 2 *Let μ be a non atomic probability measure and let $c : X^n \to [0, \infty]$ be a l.s.c. cost that is continuous in its finiteness domain and cyclical, namely $c(x_1, x_2, \dots, x_n) = c(x_n, x_1, \dots x_{n-1})$. Then $(M) = (K)$.*

In particular this is true when $X = \mathbb{R}^d$ and c is the Coulomb cost

$$c(x_1, x_2, \dots, x_n) = \sum_{i<j} \frac{1}{|x_i - x_j|}.$$

This was our model cost, whose multimarginal optimal transport problem is studied in a mathematical model for the strong interaction limit in the density functional theory; the mathematical quest in this setting is to prove that the minimum in (K) is attained by a cyclical map admissible in (M). The equality $(M) = (K)$ can be seen as a first validation of this conjecture.

Multimarginal optimal transport maps for 1-dimensional repulsive costs

In [29], in collaboration with M. Colombo and L. De Pascale, we deal with a particular multi marginal optimal transportation problem. Referring to the previous subsection, we prove that in the case $X = \mathbb{R}$ with the cost c with the peculiar structure

$$c(x_1, \ldots, x_n) = \sum_{i<j} f(x_i - x_j), \tag{15}$$

where f is an even nonnegative l.s.c. function, that restricted to $(0, \infty)$ is convex and decreasing, we have that the minimum in (M) is reached, and moreover we find also explicitly its form:

Theorem 3 *Let c be the cost* (15). *Let ρ be a non-atomic probability measure on $\mathbb{R}$ such that $(K) < \infty$. Let $-\infty = d_0 < d_1 < \ldots < d_N = +\infty$ be such that*

$$\rho([d_i, d_{i+1}]) = 1/N \qquad \forall i = 0, \ldots, N-1. \tag{16}$$

Let $T : \mathbb{R} \to \mathbb{R}$ be the unique (up to ρ-null sets) function increasing on each interval $[d_i, d_{i+1}]$, $i = 0, ..., N-1$, and such that

$$T_\sharp(1_{[d_i,d_{i+1}]}\rho) = 1_{[d_{i+1},d_{i+2}]}\rho \quad \forall i = 0, \ldots, N-2, \quad \textit{and} \quad T_\sharp(1_{[d_{N-1},d_N]}\rho) = 1_{[d_0,d_1]}\rho. \tag{17}$$

Then T is an admissible map for (M) and

$$(K) = \int_{\mathbb{R}} c(x, T(x), T^{(2)}(x), \ldots, T^{(N-1)}(x))\, d\rho. \tag{18}$$

Moreover the only symmetric optimal transport plan is the symmetrization of the plan induced by the map T.

We recall that a symmetric transport plan is a transport plan π such that $(\sigma)_\sharp \boldsymbol{\pi} = \boldsymbol{\pi}$ for every permutation of the coordinates σ.

Lower semicontinuity for non-coercive polyconvex integrals in the limit case

In [32], in collaboration with G. De Philippis and M. Focardi, we deal with the problem of lower semicontinuity for integrals $F : W^{1,m-1}(\mathbb{R}^n;\mathbb{R}^m)$ of the form

$$F(u) = \int_\Omega f(x, u(x), \mathcal{M}^l(\nabla u))\, \mathrm{d}x,$$

where $l := \min\{m, n\}$ and $\mathcal{M}^l(\mathbb{A})$ denotes the vector whose components are all the minors of order un to l of the matrix $\mathbb{A} \in \mathbb{R}^{m\times n}$.

In this paper, we investigate the lower semicontinuity properties of energies with densities f satisfying

(Hp) $f = f(x, u, \xi) : \Omega \times \mathbb{R}^n \times \mathbb{R}^\sigma \to [0, \infty)$ is in $C^0(\Omega \times \mathbb{R}^n \times \mathbb{R}^\sigma)$ and $f(x, u, \cdot)$ is convex for all $(x, u) \in \Omega \times \mathbb{R}^n$

along sequences

(Seq) $(u_j)_j \subset W^{1,\ell}(\Omega, \mathbb{R}^m)$ satisfying

$$u_j \rightharpoonup u \quad \text{in } W^{1,\ell-1}. \tag{19}$$

Our main results are:

Theorem 4 *Let $2 \le m \le n$, let f satisfy* (Hp), *and suppose in addition that*

$$f(\cdot, \cdot, \xi) \quad \textit{is locally Lipschitz continuous for all } \xi \in \mathbb{R}^\sigma. \tag{20}$$

Then, for every sequence $(u_j)_j \subset W^{1,m}(\Omega, \mathbb{R}^m)$ satisfying (Seq) *we have*

$$F(u) \le \liminf_j F(u_j).$$

Theorem 5 *Let $2 \le m = n$, and let f enjoy* (Hp).

Then, for every sequence $(u_j)_j \subset W^{1,n}(\Omega, \mathbb{R}^n)$ satisfying (Seq) *we have*

$$F(u) \le \liminf_j F(u_j). \tag{21}$$

Theorem 6 *Let $2 \le m = n + 1$, let $f : \mathbb{R}^\sigma \to [0, \infty)$ be convex, and*

$$F(u) = \int_\Omega f\big(\mathscr{M}^n(\nabla u(x))\big)\, dx\,.$$

Then, for every sequence $(u_j)_j \subset W^{1,n}(\Omega, \mathbb{R}^{n+1})$ satisfying (Seq) *we have*

$$F(u) \le \liminf_j F(u_j).$$

CHAPTER 1

Preliminary notions

In this section we introduce some notation and recall a few basic facts about capacities and Choquet theorem, absolutely continuous functions and Lipschitz functions, gradient flows of convex functionals, Orlicz spaces and optimal transportation, see also [8], [79], [72] as general references.

Furthermore we will recall Hopf-Lax formula and Hamilton-Jacobi equation in metric spaces, a tool that will be useful in Chapter 3.

1.1 Topological spaces and Choquet theorem

In a topological Hausdorff space (E, τ), we denote by $\mathscr{P}(E)$ the collection of all subsets of E, by $\mathscr{F}(E)$ (resp. $\mathscr{K}(E)$) the collection of all closed (resp. compact) sets of E, by $\mathscr{B}(E)$ the σ-algebra of Borel sets of E. We denote by $\mathrm{C}_b(E)$ the space of bounded continuous functions on (E, τ), by $\mathcal{M}_+(E)$, the set of σ-additive measures $\mu : \mathscr{B}(E) \to [0, \infty)$, by $\mathcal{P}(E)$ the subclass of probability measures. For a set $F \subset E$ and $\mu \in \mathcal{M}_+(E)$ we shall respectively denote by $\chi_F : E \to \{0, 1\}$ the characteristic function of F and by $\mu \llcorner F$ the measure $\chi_F \mu$, if F is μ-measurable. For a Borel map $L : E \to F$ we shall denote by $L_\sharp : \mathcal{M}_+(E) \to \mathcal{M}_+(F)$ the induced push-forward operator between Borel measures, namely

$$L_\sharp \mu(B) := \mu\big(L^{-1}(B)\big) \qquad \forall \mu \in \mathcal{M}_+(E),\ B \in \mathscr{B}(F).$$

We shall denote by $\mathbb{N} = \{0, 1, \ldots\}$ the natural numbers, by $\mathscr{L}^d$ the Lebesgue measure on the d-dimensional Euclidean space $\mathbb{R}^d$.

1.1.1 Polish spaces

Recall that (E, τ) is said to be Polish if there exists a distance ρ in E which induces the topology τ such that (E, ρ) is complete and separable. Notice that the inclusion of $\mathcal{M}_+(E)$ in $(\mathrm{C}_b(E))^*$ may be strict, because we are not making compactness or local compactness assumptions on (E, τ). Nevertheless, if (E, τ) is Polish we can always endow $\mathcal{M}_+(E)$ with a Polish topology w-$\mathrm{C}_b(E)$ whose convergent sequences are precisely the weakly convergent ones, i.e. sequences convergent in the duality with $\mathrm{C}_b(E)$. Obviously this Polish topology is unique. A possible choice, which can be easily adapted from the corresponding Kantorovich-Rubinstein distance on $\mathcal{P}(E)$ (see e.g. [20, §8.3] or [8, Section 7.1]) is to consider the duality with bounded and Lipschitz functions

$$\rho_{KR}(\mu,\nu) := \sup\Big\{\Big|\int_E f\,\mathrm{d}\mu - \int_E f\,\mathrm{d}\nu\Big| : f \in \mathrm{Lip}_b(E),\ \sup_E |f| \le 1,\\ |f(x)-f(y)| \le \rho(x,y) \quad \forall\, x,y \in E\Big\}.$$

1.1.2 Souslin, Lusin and analytic sets, Choquet theorem

Denote by $\mathbb{N}^\infty$ the collection of all infinite sequences of natural numbers and by $\mathbb{N}_0^\infty$ the collection of all finite sequences $(n_0,\ldots,n_i)$, with $i \ge 0$ and n_i natural numbers. Let $\mathscr{A} \subset \mathscr{P}(E)$ containing the empty set (typical examples are, in topological spaces (E,τ), the classes $\mathscr{F}(E)$, $\mathscr{K}(E)$, $\mathscr{B}(E)$). We call table of sets in $\mathscr{A}$ a map C associating to each finite sequence $(n_0,\ldots,n_i) \in \mathbb{N}_0^\infty$ a set $C_{(n_0,\ldots,n_i)} \in \mathscr{A}$.

Definition 1.1.1 ($\mathscr{A}$-analytic sets) *$S \subset E$ is said to be $\mathscr{A}$-analytic if there exists a table C of sets in $\mathscr{A}$ such that*

$$A = \bigcup_{(n)\in\mathbb{N}^\infty} \bigcap_{i=0}^{\infty} C_{(n_0,\ldots,n_i)}.$$

Recall that, in a topological space (E,τ), $\mathscr{B}(E)$-analytic sets are *universally measurable* [20, Theorem 1.10.5]: this means that they are σ-measurable for any $\sigma \in \mathcal{M}_+(E)$.

Definition 1.1.2 (Souslin and Lusin sets) *Let (E,τ) be an Hausdorff topological space. $S \in \mathscr{P}(E)$ is said to be a Souslin (resp. Lusin) set if it is the image of a Polish space under a continuous (resp. continuous and injective) map.*

Even though the Souslin and Lusin properties for subsets of a topological space are intrinsic, i.e. they depend only on the induced topology, we will often use the diction "S Suslin subset of E" and similar to emphasize the ambient space; the Borel property, instead, is not intrinsic, since $S \in \mathscr{B}(S)$ if we endow S with the induced topology. Besides the obvious stability with respect to trasformations through continuous (resp. continuous and injective) maps, the class of Souslin (resp. Lusin) sets enjoys nice properties, detailed below.

Proposition 1.1.3 *The following properties hold:*

(i) In a Hausdorff topological space (E,τ), Souslin sets are $\mathscr{F}(E)$-analytic;

(ii) if (E,τ) is a Souslin space (in particular if it is a Polish or a Lusin space), the notions of Souslin and $\mathscr{F}(E)$-analytic sets concide and in this case Lusin sets are Borel and Borel sets are Souslin;

(iii) if E, F are Souslin spaces and $f : E \to F$ is a Borel injective map, then f^{-1} is Borel;

(iv) if E, F are Souslin spaces and $f : E \to F$ is a Borel map, then f maps Souslin sets to Souslin sets.

Proof. We quote [20] for all these statements: (i) is proved in Theorem 6.6.8; in connection with (ii), the equivalence between Souslin and $\mathscr{F}(E)$-analytic sets is proved in Theorem 6.7.2, the fact that Borel sets are Souslin in Corollary 6.6.7 and the fact that Lusin sets are Borel in Theorem 6.8.6; finally, (iii) and (iv) are proved in Theorem 6.7.3. $\square$

Since in Polish spaces (E,τ) we have at the same time tightness of finite Borel measures and coincidence of Souslin and $\mathscr{F}(E)$-analytic sets, the measurability of $\mathscr{B}(E)$-analytic sets yields in particular that

$$\sigma(B) = \sup\{\sigma(K)\ :\ K \in \mathscr{K}(E),\ K \subset B\} \quad \text{for all } B \subset E \text{ Souslin},\ \sigma \in \mathcal{M}_+(E). \tag{1.1.1}$$

We will need a property analogous to (1.1.1) for capacities [33], whose definition is recalled below.

Definition 1.1.4 (Capacity) *A set function $\mathfrak{I} : \mathscr{P}(E) \to [0,\infty]$ is said to be a* capacity *if:*

- *$\mathfrak{I}$ is nondecreasing and, whenever $(A_n) \subset \mathscr{P}(E)$ is nondecreasing, the following holds*

$$\lim_{n\to\infty} \mathfrak{I}(A_n) = \mathfrak{I}\left(\bigcup_{n=0}^{\infty} A_n\right);$$

- *if $(K_n) \subset \mathscr{K}(E)$ is nonincreasing, the following holds:*

$$\lim_{n\to\infty} \mathfrak{I}(K_n) = \mathfrak{I}\left(\bigcap_{n=0}^{\infty} K_n\right).$$

A set $B \subset E$ is said to be $\mathfrak{I}$-capacitable if $\mathfrak{I}(B) = \sup_{K\in\mathscr{K}(E)} \mathfrak{I}(K)$.

Theorem 1.1.5 (Choquet) ([33, Thm 28.III]) *Every $\mathscr{K}(E)$-analytic set is capacitable.*

1.2 Absolutely continuous curves

If (X,d) is a metric space and $I \subset \mathbb{R}$ is an interval, we denote by $\mathrm{C}(I;X)$ the class of continuous maps (often called parametric curves) from I to X. We will use the notation γ_t for the value of the map at time t and $\mathrm{e}_t : \mathrm{C}(I;X) \to X$ for the evaluation map at time t; occasionally, in order to avoid double subscripts, we will also use the notation $\gamma(t)$. The subclass $\mathrm{AC}(I;X)$ is defined by the property

$$\mathsf{d}(\gamma_s,\gamma_t) \le \int_s^t g(r)\,\mathrm{d}r \qquad s,\, t \in I,\ s \le t$$

for some (nonnegative) $g \in L^1(I)$. The least, up to $\mathscr{L}^1$-negligible sets, function g with property is the so-called metric derivative (or metric speed)

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

see [13] for its existence. The classes $\mathrm{AC}^p(I;X)$, $1 \le p \le \infty$ are defined analogously, requiring that $|\dot\gamma| \in L^p(I)$. The p-energy of a curve is then defined as

$$\mathcal{E}_p(\gamma) := \begin{cases} \int_I |\dot\gamma_t|^p \,\mathrm{d}t & \text{if } \gamma \in \mathrm{AC}^p(I;X), \\ +\infty & \text{otherwise,} \end{cases} \tag{1.2.1}$$

and $\mathcal{E}_1(\gamma) = \ell(\gamma)$, the length of γ, when $p = 1$. Notice that $\mathrm{AC}^1 = \mathrm{AC}$ and that $\mathrm{AC}^\infty(I;X)$ coincides with the class of d-Lipschitz functions.

If (X, d) is complete the interval I can be taken closed with no loss of generality, because absolutely continuous functions extend continuously to the closure of the interval. In addition, if (X, d) is complete and separable then $\mathrm{C}(I;X)$ is a Polish space, and $\mathrm{AC}^p(I;X)$, $1 \le p \le \infty$ are Borel subsets of $\mathrm{C}(I;X)$ (see for instance [11]). We will use the short notation $\mathcal{M}_+(\mathrm{AC}^p(I;X))$ to denote finite Borel measures in $\mathrm{C}(I;X)$ concentrated on $\mathrm{AC}^p(I;X)$. The integration of a Borel function g along a curve $\gamma \in \mathrm{AC}(I;X)$ is well defined by the formula

$$\int_\gamma g = \int_I g(\gamma_t)|\dot\gamma_t| \,\mathrm{d}t$$

1.2.1 Reparameterization

We collect in the next proposition a few properties which are well-known in a smooth setting, but still valid in general metric spaces. We introduce the notation

$$\mathrm{AC}^\infty_c([0,1];X) := \Big\{\sigma \in \mathrm{AC}^\infty([0,1];X) : |\dot\sigma| = \ell(\sigma) > 0 \quad \mathscr{L}^1\text{-a.e. on } (0,1)\Big\} \tag{1.2.2}$$

for the subset of $\mathrm{AC}([0,1];X)$ consisting of all nonconstant curves with constant speed. It is easy to check that $\mathrm{AC}^\infty_c([0,1];X)$ is a Borel subset of $\mathrm{C}([0,1];X)$, since it can also be characterized by

$$\gamma \in \mathrm{AC}^\infty_c([0,1];X) \quad \Longleftrightarrow \quad 0 < \mathrm{Lip}(\gamma) \le \ell(\gamma), \tag{1.2.3}$$

and the maps $\gamma \mapsto \mathrm{Lip}(\gamma)$ and $\gamma \mapsto \ell(\gamma)$ are lower semicontinuous.

Proposition 1.2.1 (Constant speed reparameterization) *For any $\gamma \in \mathrm{AC}([0,1];X)$ with $\ell(\gamma) > 0$, setting*

$$\mathsf{s}(t) := \frac{1}{\ell(\gamma)} \int_0^t |\dot\gamma_r| \,\mathrm{d}r, \tag{1.2.4}$$

there exists a unique $\eta \in \mathrm{AC}^\infty_c([0,1];X)$ such that $\gamma = \eta \circ \mathsf{s}$. Furthermore, $\eta = \gamma \circ \mathsf{s}^{-1}$ where s^{-1} is any right inverse of s. We shall denote by

$$\mathsf{k} : \Big\{\gamma \in \mathrm{AC}([0,1];X) : \ell(\gamma) > 0\Big\} \to \mathrm{AC}^\infty_c([0,1];X) \qquad \gamma \mapsto \eta = \gamma \circ \mathsf{s}^{-1} \tag{1.2.5}$$

the corresponding map.

Proof. We prove existence only, the proof of uniqueness being analogous. Les us now define a right inverse, denoted by s^{-1}, of s (i.e. $\mathsf{s} \circ \mathsf{s}^{-1}$ is equal to the identity): we define in the obvious way s^{-1} at points $y \in [0,1]$ such that $\mathsf{s}^{-1}(y)$ is a singleton; since, by construction, γ is constant in all (maximal) intervals $[c,d]$ where s is constant, at points y such that $\{y\} = \mathsf{s}([c,d])$ we define $\mathsf{s}^{-1}(y)$ by choosing any element of $[c,d]$, so that $\gamma \circ \mathsf{s}^{-1} \circ \mathsf{s} = \gamma$ (even though it could be that $\mathsf{s}^{-1} \circ \mathsf{s}$ is not the identity). Therefore, if we define $\eta = \gamma \circ \mathsf{s}^{-1}$, we obtain that $\gamma = \eta \circ \mathsf{s}$ and that η is independent of the chosen right inverse.

In order to prove that $\eta \in \mathrm{AC}^\infty_c([0,1];X)$ we define $\ell_k := \ell(\gamma) + 1/k$ and we approximate uniformly in $[0,1]$ the map s by the maps $\mathsf{s}_k(t) := \ell_k^{-1} \int_0^t (k^{-1} + |\dot\gamma_r|)\,dr$, whose inverses $\mathsf{s}_k^{-1} : [0,1] \to I$ are Lipschitz. By Helly's theorem and passing to the limit as $k \to \infty$ in $\mathsf{s}_k \circ \mathsf{s}_k^{-1}(y) = y$, we can assume that a subsequence $\mathsf{s}^{-1}_{k(p)}$ pointwise converges to a right inverse s^{-1} as $p \to \infty$; the curves $\eta^p := \gamma \circ \mathsf{s}^{-1}_{k(p)}$ are absolutely continuous, pointwise converge to $\eta := \gamma \circ \mathsf{s}^{-1}$ and

$$|\eta^p(t)'| = \frac{|\gamma'(\mathsf{s}^{-1}_{k(p)}(t))|}{\mathsf{s}'_{k(p)}(\mathsf{s}^{-1}_{k(p)}(t))} \le \ell_{k(p)} \qquad \text{for } \mathscr{L}^1\text{-a.e. in } t \in (0,1).$$

It follows that η is absolutely continuous and that $|\dot\eta| \le \ell(\gamma)$ $\mathscr{L}^1$-a.e. in $(0,1)$. If the strict inequality occurs in a set of positive Lebesgue measure, the inequality $\ell(\eta) < \ell(\gamma)$ provides a contradiction. □

1.2.2 Equivalence relation in $\mathrm{AC}([0,1];X)$

We can identify curves $\gamma \in \mathrm{AC}([0,1],X)$, $\tilde\gamma \in \mathrm{AC}([0,1];X)$ if there exists $\varphi : [0,1] \to [0,1]$ increasing with $\varphi \in \mathrm{AC}([0,1];[0,1])$, $\varphi^{-1} \in \mathrm{AC}([0,1];[0,1])$ such that $\gamma = \tilde\gamma \circ \varphi$. In this case we write $\gamma \sim \tilde\gamma$. Thanks to the following lemma, the absolute continuity of φ^{-1} is equivalent to $\varphi' > 0$ $\mathscr{L}^1$-a.e. in $(0,1)$.

Lemma 1.2.2 (Absolute continuity criterion) *Let I, $\tilde I$ be compact intervals in $\mathbb{R}$ and let $\varphi : I \to \tilde I$ be an absolutely continuous homeomorphism with $\varphi' > 0$ $\mathscr{L}^1$-a.e. in I. Then $\varphi^{-1} : \tilde I \to I$ is absolutely continuous.*

Proof. Let $\psi = \varphi^{-1}$; it is a continuous function of bounded variation whose distributional derivative we shall denote by μ. Since $\mu([a,b]) = \psi(b) - \psi(a)$ for all $0 \le a \le b \le 1$, we need to show that $\mu \ll \mathscr{L}^1$. It is a general property of continuous BV functions (see for instance [7, Proposition 3.92]) that $\mu(\psi^{-1}(B)) = 0$ for all Borel and $\mathscr{L}^1$-negligible sets $B \subset \mathbb{R}$. Choosing $B = \psi(E)$, where E is a $\mathscr{L}^1$-negligible set where the singular part μ^s of μ is concentrated, the area formula gives

$$\int_B \varphi'(s)\,\mathrm{d}s = \mathscr{L}^1(E) = 0,$$

so that the positivity of φ' gives $\mathscr{L}^1(B) = 0$. It follows that $\mu^s = 0$. □

Definition 1.2.3 (The map J) *For any $\gamma \in \mathrm{AC}([0,1];X)$ we denote by $J\gamma \in \mathcal{M}_+(X)$ the push forward under γ of the measure $|\dot\gamma|\mathscr{L}^1 \llcorner [0,1]$, namely the measure that represents the integration along the curve γ:*

$$\int_X g\,\mathrm{d}J\gamma = \int_0^1 g(\gamma_t)|\dot\gamma_t|\,\mathrm{d}t \qquad \text{for all } g : X \to [0,\infty] \text{ Borel.} \tag{1.2.6}$$

In particular we have that $J\gamma = J\eta$ whenever $\gamma \sim \eta$, and that $J\gamma = J\mathsf{k}\gamma$.

Although this will not play a role in the sequel, for completeness we provide an intrinsic description of the measure $J\gamma$. We denote by $\mathscr{H}^1$ the 1-dimensional Hausdorff measure of a subset B of X, namely $\mathscr{H}^1(B) = \lim_{\delta\downarrow 0} \mathscr{H}^1_\delta(B)$, where

$$\mathscr{H}^1_\delta(B) := \inf\left\{\sum_{i=0}^{\infty} \mathrm{diam}(B_i) \ :\ B \subset \bigcup_{i=0}^{\infty} B_i,\ \ \mathrm{diam}(B_i) < \delta\right\}$$

(with the convention $\mathrm{diam}(\emptyset) = 0$).

Proposition 1.2.4 (Area formula) *If $\gamma \in \mathrm{AC}([0,1];X)$, then for all $g : X \to [0,\infty]$ Borel the area formula holds:*

$$\int_0^1 g(\gamma_t)|\dot\gamma_t|\,\mathrm{d}t = \int_X g(x)N(\gamma,x)\,\mathrm{d}\mathscr{H}^1(x), \tag{1.2.7}$$

where $N(\gamma,x) := \mathrm{card}(\gamma^{-1}(x))$ is the multiplicity function of γ. Equivalently,

$$J\gamma = N(\gamma,\cdot)\mathscr{H}^1. \tag{1.2.8}$$

Proof. For an elementary proof of (1.2.7), see for instance [13, Theorem 3.4.6]. □

1.2.3 Non-parametric curves

We can now introduce the class of non-parametric curves; notice that we are conventionally excluding from this class the constant curves. We first introduce the notation

$$\mathrm{AC}_0([0,1];X) := \left\{\gamma \in \mathrm{AC}([0,1];X) :\ |\dot\gamma| > 0\ \mathscr{L}^1\text{-a.e. on } (0,1)\right\}.$$

It is not difficult to show that $\mathrm{AC}_0([0,1];X)$ is a Borel subset of $\mathrm{C}([0,1];X)$. In addition, Lemma 1.2.2 shows that for any $\gamma \in \mathrm{AC}_0([0,1];X)$ the curve $\mathsf{k}\gamma \in \mathrm{AC}^\infty_c([0,1];X)$ is equivalent to γ.

Definition 1.2.5 (The class $\mathscr{C}(X)$ of non-parametric curves) *The class $\mathscr{C}(X)$ is defined as*

$$\mathscr{C}(X) := \mathrm{AC}_0([0,1];X)/\sim\,, \tag{1.2.9}$$

endowed with the quotient topology $\tau_{\mathscr{C}}$ and the canonical projection $\pi_{\mathscr{C}(X)}$.

We shall denote the typical element of $\mathscr{C}(X)$ either by $\underline{\gamma}$ or by $[\gamma]$, to mark a distinction with the notation used for parametric curves. We will use the notation $\underline{\gamma}_{\mathrm{ini}}$ and $\underline{\gamma}_{\mathrm{fin}}$ the initial and final point of the curve $\underline{\gamma} \in \mathscr{C}(X)$, respectively.

Definition 1.2.6 (Canonical maps) *We denote:*

(a) by $\mathsf{i} := \pi_{\mathscr{C}} \circ \mathsf{k} : \left\{\gamma \in \mathrm{AC}([0,1];X) :\ \ell(\gamma) > 0\right\} \to \mathscr{C}(X)$ the projection provided by Proposition 1.2.1, which coincides with the canonical projection $\pi_{\mathscr{C}(X)}$ on the quotient when restricted to $\mathrm{AC}_0([0,1];X)$;

(b) by $\mathsf{j} := \mathsf{k}\circ\pi_{\mathscr{C}}^{-1} : \mathscr{C}(X) \to \mathrm{AC}^\infty_c([0,1];X)$ the canonical representation of a non-parametric curve by a parametrization in $[0,1]$ with constant velocity.

(c) by $\tilde{J} : \mathscr{C}(X) \to \mathcal{M}_+(X) \setminus \{0\}$ *the quotient of the map* J *in* (1.2.6)*, defined by*

$$\tilde{J}[\gamma] := J\gamma. \tag{1.2.10}$$

We notice that $\mathrm{AC}_0/\sim \neq \mathrm{AC}/\sim$; in particular in the latter there are equivalence classes without representatives in $\mathrm{AC}_c^\infty([0,1];X)$, for example when we consider the equivalence class of a curve that stops for positive time on a single point.

Remark 1.2.7 *Thanks to* (1.2.6) *we have that* $\int_\gamma g$ *is well defined for* $\gamma \in \mathscr{C}(X)$*; in particular, we have that* $\int_\gamma g = \int_X g \,\mathrm{d}\tilde{J}\gamma$.

Lemma 1.2.8 (Measurable structure of $\mathscr{C}(X)$) *If* (X,d) *is complete and separable, the space* $(\mathscr{C}(X), \tau_{\mathscr{C}})$ *is a Lusin Hausdorff space and the restriction of the map* i *to* $\mathrm{AC}_c^\infty([0,1];X)$ *is a Borel isomorphism. In particular, a collection of curves* $\Gamma \subset \mathscr{C}(X)$ *is Borel if and only if* $\mathsf{j}(\Gamma)$ *is Borel in* $\mathrm{C}([0,1];X)$*. Analogously,* $\Gamma \subset \mathscr{C}(X)$ *is Souslin if and only if* $\mathsf{j}(\Gamma)$ *is Souslin in* $\mathrm{C}([0,1];X)$.

Proof. Let us first show that $(\mathscr{C}(X), \tau_{\mathscr{C}})$ is Hausdorff. We argue by contradiction and we suppose that there exist curves $\mathsf{i}(\sigma_i) \in \mathscr{C}(X)$ with $\sigma_i \in \mathrm{AC}_c^\infty([0,1];X)$, $i = 1, 2$, and a sequence of parametrizations $\mathsf{s}_i^n \in \mathrm{AC}([0,1];[0,1])$ with $(\mathsf{s}_i^n)' > 0$ $\mathscr{L}^1$-a.e. in $(0,1)$, such that

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \mathsf{d}(\sigma_1(\mathsf{s}_1^n(t)), \sigma_2(\mathsf{s}_2^n(t))) = 0.$$

Denoting by $\mathsf{r}_1^n(t) := \mathsf{s}_1^n \circ (\mathsf{s}_2^n)^{-1}$ and $\mathsf{r}_2^n(t) := \mathsf{s}_2^n \circ (\mathsf{s}_1^n)^{-1}$, we get

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \mathsf{d}(\sigma_1(t), \sigma_2(\mathsf{r}_2^n(t))) = 0, \qquad \lim_{n\to\infty} \sup_{t\in[0,1]} \mathsf{d}(\sigma_1(\mathsf{r}_1^n(t)), \sigma_2(t)) = 0.$$

The lower semicontinuity of the length with respect to uniform convergence yields $\ell := \ell(\sigma_1) = \ell(\sigma_2)$ and therefore for every $0 \le t' < t'' \le 1$

$$\ell \liminf_{n\to\infty} \big(r_2^n(t'') - r_2^n(t')\big) = \lim_{n\to\infty} \int_{t'}^{t''} |(\sigma_2 \circ \mathsf{r}_2^n)'| \,\mathrm{d}t \ge \int_{t'}^{t''} |\sigma_1'| \,\mathrm{d}t = \ell(t'' - t').$$

Choosing first $t' = t$ and $t'' = 1$ and then $t' = 0$ and $t'' = t$ we conclude that $\lim_n r_2^n(t) = t$ for every $t \in [0,1]$ and therefore $\sigma_1 = \sigma_2$.

Notice that $\mathrm{AC}_c^\infty([0,1];X)$ is a Lusin space, since $\mathrm{AC}_c^\infty([0,1];X)$ is a Borel subset of $\mathrm{C}([0,1];X)$. The restriction of i to $\mathrm{AC}_c^\infty([0,1];X)$ is thus a continuous and injective map from the Lusin space $\mathrm{AC}_c^\infty([0,1];X)$ to the Hausdorff space $(\mathscr{C}(X), \tau_{\mathscr{C}})$ (notice that the topology $\tau_{\mathscr{C}}$ is a priori weaker than the one induced by the restriction of i to $\mathrm{AC}_c^\infty([0,1];X)$). It follows by definition that $\mathscr{C}(X)$ is Lusin. Now, Proposition 1.1.3(iii) yields that the restriction of i is a Borel isomorphism. □

Lemma 1.2.9 (Borel regularity of J and $\tilde{J}$) *The map* $J : \mathrm{AC}([0,1];X) \to \mathcal{M}_+(X)$ *is Borel, where* $\mathrm{AC}([0,1];X)$ *is endowed with the* $\mathrm{C}([0,1];X)$ *topology. In particular, if* (X,d) *is complete and separable, the map* $\tilde{J} : \mathscr{C}(X) \to \mathcal{M}_+(X) \setminus \{0\}$ *is Borel and* $\tilde{J}(\Gamma)$ *is Souslin in* $\mathcal{M}_+(X)$ *whenever* Γ *is Souslin in* $\mathscr{C}(X)$.

Proof. It is easy to check, using the formula $J\gamma = \gamma_\sharp(|\dot\gamma|\mathscr{L}^1 \llcorner [0,1])$, that

$$J\gamma = \lim_{n\to\infty} \sum_{i=0}^{n-1} \mathsf{d}(\gamma_{(i+1)/n}, \gamma_{i/n})\delta_{\gamma_{i/n}} \qquad \text{weakly in } \mathcal{M}_+(X)$$

for all $\gamma \in \mathrm{AC}([0,1];X)$ (the simple details are left to the reader). Since the approximating maps are continuous, we conclude that J is Borel. The Borel regularity of $\tilde J$ follows by Lemma 1.2.8 and the identity $\tilde J = J \circ \mathsf{j}$. Since $\tilde J$ is Borel, we can apply Proposition 1.1.3(iv) to obtain that $\tilde J$ maps Souslin sets to Souslin sets. □

1.3 Slopes, asymptotic Lipschitz constant and upper gradients

Let (X, d) be a metric space; given $f : X \to \mathbb{R}$ and $E \subset X$, we denote by $\mathrm{Lip}(f, E)$ the Lipschitz constant of the function f on E, namely

$$\mathrm{Lip}(f, E) := \sup_{x,\, y\in E,\, x\neq y} \frac{|f(x) - f(y)|}{\mathsf{d}(x, y)}.$$

The Lipschitz constant of f will be denoted by $\mathrm{Lip}(f) := \mathrm{Lip}(f, X)$. Given $f : X \to \mathbb{R}$, we define *asymptotic Lipschitz constant* by

$$\mathrm{lip}_a(f, x) := \lim_{r\to 0} \mathrm{Lip}(f, B_r(x)),$$

and *slope* (also called local Lipschitz constant) by

$$|\nabla f|(x) := \varlimsup_{y\to x} \frac{|f(y) - f(x)|}{\mathsf{d}(y, x)}.$$

We will often drop the x dependence, denoting $\mathrm{lip}_a(f)$ for the asymptotic Lipschitz constant, and $|\nabla f|$ for the slope. For f, $g : X \to \mathbb{R}$ Lipschitz it clearly holds

$$\mathrm{lip}_a(\alpha f + \beta g) \le |\alpha|\mathrm{lip}_a f + |\beta|\mathrm{lip}_a g \qquad \forall \alpha, \beta \in \mathbb{R}, \tag{1.3.1a}$$

$$\mathrm{lip}_a(fg) \le |f|\mathrm{lip}_a g + |g|\mathrm{lip}_a f, \tag{1.3.1b}$$

and the same is true also for the slope. We recall the basic relation between the asymptotic Lipschitz constant and the slope in the next proposition.

Proposition 1.3.1 *Let $f : X \to \mathbb{R}$ be a Lipschitz function. Then*

$$\mathrm{Lip}(f) \ge \mathrm{lip}_a(f, x) \ge |\nabla f|^*(x), \tag{1.3.2}$$

where $|\nabla f|^*$ is the upper semicontinuous envelope of the slope of f. In length spaces the second inequality is an equality.

Proof. The first inequality in (1.3.2) is trivial, while the second one follows by the fact that $\mathrm{lip}_a(f,\cdot)$ is upper semicontinuous and larger than $|\nabla f|$. Since $|\nabla f|$ is an upper gradient of f, we have the inequality

$$|f(y) - f(z)| \le \int_0^{\ell(\gamma)} |\nabla f|(\gamma_t)\, dt$$

for any curve γ with constant speed joining y to z. If (X, d) is a length space we can minimize w.r.t. γ to get

$$\mathrm{Lip}\big(f, B(x, r)\big) \le \sup_{B(x,3r)} |\nabla f| \le \sup_{B(x,3r)} |\nabla f|^*.$$

As $r \downarrow 0$ the inequality $\mathrm{Lip}_a(f, x) \le |\nabla f|^*(x)$ follows. □

We will need also this refined Liebniz formula:

Lemma 1.3.2 *Let $A \subset X$ be an open set, and let $f, g, \varphi \in \mathrm{Lip}_{\mathrm{loc}}(A)$ such that $0 \le \varphi \le 1$; then denoting $w = \varphi f + (1-\varphi) g$ we have*

$$\mathrm{lip}_a(w) \le \varphi \cdot \mathrm{lip}_a(f) + (1-\varphi) \cdot \mathrm{lip}_a(g) + \mathrm{lip}_a(\varphi)|f-g| \tag{1.3.3}$$

Proof. First let us not that for every $x, y \in X$ we have

$$\begin{aligned} w(x) - w(y) = \varphi(x)[f(x) - f(y)] + (1-\varphi(x))[g(x) - g(y)] \\ + [\varphi(x) - \varphi(y)] \cdot (f(y) - g(y)); \end{aligned}$$

taking the modulus on the lest hand side and dividing for $\mathsf{d}(x,y)$ we obtain

$$\begin{aligned} \frac{|w(x) - w(y)|}{\mathsf{d}(x,y)} \le \varphi(x) \frac{|f(x) - f(y)|}{\mathsf{d}(x,y)} + (1-\varphi(x)) \frac{|g(x) - g(y)|}{\mathsf{d}(x,y)} \\ + \frac{|\varphi(x) - \varphi(y)|}{\mathsf{d}(x,y)} \cdot |f(y) - g(y)|. \end{aligned} \tag{1.3.4}$$

Now taking the supremum in $x, y \in B_r(z)$ on the left hand side we obtain

$$\mathrm{Lip}(w, B_r) \le \varphi \, \mathrm{Lip}(f, B_r) + (1-\varphi) \, \mathrm{Lip}(g, B_r) + \mathrm{Lip}(\varphi, B_r) \sup_{B_r} |f - g|;$$

letting $r \to 0$ we get (1.3.3). $\square$

Given a real valued function f on X, we denote by $UG(f)$ the set of upper gradients of f (see also [25], [49]), namely the class of Borel functions $g : X \to [0, \infty]$ such that

$$\left| \int_{\partial\gamma} f \right| = \int_\gamma g \qquad \forall \, \gamma \in \mathscr{C}(X), \tag{1.3.5}$$

where $\int_{\partial\gamma} f = f(\gamma_{fin}) - f(\gamma_{ini})$. With a slight abuse of notation we will write $g \in UG(f)$ with $f \in L^1(X, \mathfrak{m})$, but it should be noticed that a priori the concept of upper gradient is not invariant in the equivalence class of an L^1 function, even though Borel representatives are chosen. It is easy to see that $\mathrm{lip}_a(f)$ and $|\nabla f|$ belong to $UG(f)$ whenever f is a locally Lipschitz function.

We shall also need the following calculus lemma.

Lemma 1.3.3 *Let $f : (0,1) \to \mathbb{R}$, $q \in [1, \infty]$, $g \in L^q(0,1)$ nonnegative be satisfying*

$$|f(s) - f(t)| \le \int_s^t g(r) \, \mathrm{d}r \qquad \text{for } \mathscr{L}^2\text{-a.e. } (s,t) \in (0,1)^2.$$

Then $f \in W^{1,q}(0,1)$ and $|f'| \le g$ a.e. in $(0,1)$.

Proof. Let $N \subset (0,1)^2$ be the $\mathscr{L}^2$-negligible subset where the above inequality fails. Choosing $s \in (0,1)$, whose existence is ensured by Fubini's theorem, such that $(s,t) \notin N$ for a.e. $t \in (0,1)$, we obtain that $f \in L^\infty(0,1)$. Since the set $N_1 = \{(t,h) \in (0,1)^2 : \ (t, t+h) \in N \cap (0,1)^2\}$ is $\mathscr{L}^2$-negligible as well, we can apply Fubini's theorem to obtain that for a.e.

h it holds $(t,h) \notin (0,1)^2 \setminus N_1$ for a.e. $t \in (0,1)$. Let $h_i \downarrow 0$ with this property and use the identities

$$\int_0^1 f(t)\frac{\varphi(t-h)-\varphi(t)}{h}\,\mathrm{d}t = \int_0^1 \frac{f(t+h)-f(t)}{h}\varphi(t)\,\mathrm{d}t$$

with $\varphi \in C_c^1(0,1)$ and $h = h_i$ sufficiently small to get

$$\left|\int_0^1 f(t)\varphi'(t)\,\mathrm{d}t\right| \le \int_0^1 g(t)|\varphi(t)|\,\mathrm{d}t.$$

It follows that the distributional derivative of f is a signed measure η with finite total variation which satisfies

$$-\int_0^1 f\varphi'\,\mathrm{d}t = \int_0^1 \varphi\,\mathrm{d}\eta, \qquad \left|\int_0^1 \varphi\,\mathrm{d}\eta\right| \le \int_0^1 g|\varphi|\,\mathrm{d}t \quad \text{for every } \varphi \in C_c^1(0,1).$$

Therefore η is absolutely continuous with respect to the Lebesgue measure with $|\eta| \le g\mathscr{L}^1$. This gives the $W^{1,1}(0,1)$ regularity and, at the same time, the inequality $|f'| \le g$ a.e. in $(0,1)$. The case $q > 1$ immediately follows by applying this inequality when $g \in L^q(0,1)$. □

1.4 Gradient flows of convex and lower semicontinuous functionals

Let H be an Hilbert space, $\mathscr{F} : H \to \mathbb{R} \cup \{+\infty\}$ convex and lower semicontinuous and $D(\mathscr{F}) = \{\mathscr{F} < \infty\}$ its finiteness domain. Recall that a gradient flow $x : (0,\infty) \to H$ of $\mathscr{F}$ is a locally absolutely continuous map with values in $D(\mathscr{F})$ satisfying

$$-\frac{\mathrm{d}}{\mathrm{d}t}x_t \in \partial^-\mathscr{F}(x_t) \qquad \text{for a.e. } t \in (0,\infty).$$

Here $\partial^-\mathscr{F}(x) \subseteq H^*$ is the subdifferential of $\mathscr{F}$, defined at any $x \in D(\mathscr{F})$ by

$$\partial^-\mathscr{F}(x) := \left\{p \in H^* :\ \mathscr{F}(y) \ge \mathscr{F}(x) + \langle p, y-x\rangle \ \forall y \in H\right\}.$$

We shall use the fact that for all $x_0 \in \overline{D(\mathscr{F})}$ there exists a unique gradient flow x_t of $\mathscr{F}$ starting from x_0, i.e. $x_t \to x_0$ as $t \downarrow 0$, and that $t \mapsto \mathscr{F}(x_t)$ is nonincreasing and locally absolutely continuous in $(0,\infty)$. In addition, this unique solution exhibits a regularizing effect, namely $-\frac{\mathrm{d}}{\mathrm{d}t}x_t$ is for a.e. $t \in (0,\infty)$ the element of minimal norm in $\partial^-\mathscr{F}(x_t)$.

1.5 N-functions and Orlicz spaces

We refer to [72] for the general theory; here we will give only a brief overview of the results we will need. A function $\Phi : \mathbb{R} \to [0,\infty)$ is called an N-function (nice Young function) if

(a) Φ is even and convex;

(b) $\Phi(x) = 0$ iff $x = 0$;

(c) $\lim\limits_{x\to 0} \frac{\Phi(x)}{x} = 0$ and $\lim\limits_{x\to\infty} \frac{\Phi(x)}{x} = +\infty$.

Every N-function has a left (right) derivative $\varphi = \Phi'_-$ ($\varphi_+ = \Phi'_+$) that is strictly increasing and lower (upper) semicontinuous, $0 < \varphi(t) < +\infty$ for $0 < t < \infty$, and we have $\varphi(0) = 0$ and $\lim_{t\to+\infty} \varphi = +\infty$; we denote by $\partial^-\Phi = [\varphi, \varphi_+]$ the subdifferential of Φ.

Let ψ be the left inverse of φ, namely $\psi(t) = \inf\{t > 0 \,:\, \varphi(t) > s\}$. Then Ψ, Φ given by

$$\Phi(x) = \int_0^{|x|} \varphi(t)\,\mathrm{d}t, \qquad\qquad \Psi(x) = \int_0^{|x|} \psi(s)\,\mathrm{d}s$$

are called complementary N-functions and they satisfy Young inequality

$$\Phi(x) + \Psi(y) \geq xy \qquad \forall x, y \in \mathbb{R}, \tag{1.5.1}$$

with equality iff $x \in \partial^-\Phi(y)$ or, equivalently, $y \in \partial^-\Psi(x)$; in particular

$$\Phi(x) + \Psi(\varphi(x)) = x\varphi(x) \qquad \forall x \geq 0. \tag{1.5.2}$$

Another important property of Ψ is that it is the least function satisfying (1.5.1), and so it is also the convex conjugate of Φ:

$$\Psi(y) = \Phi^*(y) := \sup_{x\in\mathbb{R}}\big\{xy - \Phi(x)\big\}.$$

Basic examples of complementary N-functions are $\Phi_p(x) = x^p/p$ and $\Psi_p(y) = y^q/q$, whose relative Orlicz space, defined below, is L^p.

1.5.1 Orlicz spaces

Definition 1.5.1 *Let us define the vector space*

$$\mathcal{L}^\Phi(X, \mathfrak{m}) = \Big\{f \ \mathfrak{m}\text{-measurable such that } \int_X \Phi(cf)\,\mathrm{d}\mathfrak{m} < \infty \text{ for some } c > 0\Big\},$$

along with his two norms: the Luxemburg norm and the dual norm

$$\|f\|_{\Phi,\mathfrak{m}} = \inf\Big\{t > 0 \,:\, \int_X \Phi\Big(\frac{f(x)}{t}\Big)\,\mathrm{d}\mathfrak{m} \leq 1\Big\},$$

$$\|f\|_{(\Phi),\mathfrak{m}} = \sup\Big\{\int_X fg\,\mathrm{d}\mathfrak{m} \,:\, g \in \mathcal{L}^\Psi(X,\mathfrak{m}),\ \|g\|_{\Psi,\mathfrak{m}} \leq 1\Big\}.$$

Then we define the Orlicz space L^Φ as $\mathcal{L}^\Phi/\!\sim$ where $f \sim g$ if $\|f - g\|_{\Phi,\mathfrak{m}} = 0$ or, equivalently, if $f = g$ $\mathfrak{m}$-almost everywhere. When there is no ambiguity for the measure we will drop the dependence on $\mathfrak{m}$ of the space and of the norms.

Definition 1.5.2 *Let us define the vector space*

$$\mathcal{M}^\Phi(X, \mathfrak{m}) = \Big\{f \text{ measurable such that } \int_X \Phi(cf)\,\mathrm{d}\mathfrak{m} < \infty \text{ for every } c > 0\Big\};$$

it is readily seen that $M^\Phi = \mathcal{M}^\Phi/\!\sim$ is a closed subspace of L^Φ.

Classical results in Orlicz spaces are that the two norms satisfy the triangle inequality, and that they are comparable, namely

$$\|f\|_\Phi \le \|f\|_{(\Phi)} \le 2\|f\|_\Phi \qquad \forall f \in L^\Phi(X,\mathfrak{m}).$$

Furthermore, as it is clear by the definition, a sharp Hölder inequality holds true: whenever $f \in L^\Phi(X,\mathfrak{m})$ and $g \in L^\Psi(X,\mathfrak{m})$ we have $fg \in L^1(X,\mathfrak{m})$, more precisely

$$\int_X fg\,\mathrm{d}\mathfrak{m} \le \|f\|_\Phi \cdot \|g\|_{(\Psi)} \qquad \text{and} \qquad \int_X fg\,\mathrm{d}\mathfrak{m} \le \|f\|_{(\Phi)} \cdot \|g\|_\Psi.$$

Lemma 1.5.3 (Dominated convergence in L^Φ) *Let $(f_n) \subset L^\Phi$ such that $f_n \to f$ $\mathfrak{m}$-a.e. and $|f_n - f| \le g$ for some $g \in M^\Phi$; then $f_n \to f$ strongly in L^Φ.*

Proof. Let us fix $m \in \mathbb{N}$; then we can consider $h_n = \Phi(m|f - f_n|)$ and $h = \Phi(mg)$. By the assumption we know that $h_n \le h$, $h \in L^1$ and $h_n \to 0$ as $n \to \infty$. By the standard dominated convergence theorem, this guarantees that $\int_X h_n \to 0$, in particular there is an integer n_0 such that for all $n \ge n_0$ we have

$$\int_X \Phi(m|f - f_n|)\,\mathrm{d}\mathfrak{m} = \int_X h_n\,\mathrm{d}\mathfrak{m} \le 1,$$

so that $\|f - f_n\|_\Phi \le m^{-1}$ for $n \ge n_0$. Since m was arbitrary we can conclude. $\square$

A simple application of dominated convergence and classical approximation results give that $\mathrm{Lip}_0(X,\mathsf{d}) \cap M^\Phi$ is dense in M^Φ, that thus is also separable; for any function $f \in M^\Phi$ we get also that the norm is absolutely continuous, meaning that

$$\lim_{\mathfrak{m}(A)\to 0} \|f\chi_A\|_{(\Phi),\mathfrak{m}} = 0. \tag{1.5.3}$$

Now we list an important definition for an N-function:

- Φ is doubling if there exist K such that $\Phi(2x) \le K\Phi(x)$ for all $x \ge 0$; in the case $\mathfrak{m}$ finite we say that Φ is doubling if the inequality for x large enough. We will say that Φ satisfies (D);
- Φ satisfies the double doubling condition if both Φ and Ψ are doubling. In this case we will say that Φ (or equivalently Ψ) satisfies (DD).

It is easy to see $M^\Phi = L^\Phi$ if and only if Φ is doubling. Another important property is the characterization of the dual spaces:

Theorem 1.5.4 ([72], Sec. 1.2, Theorem 13) *Let $\mathfrak{m}$ be a finite measure on X, and let Φ be an N-function. Then $(M^\Phi, \|\cdot\|_{\Phi,\mathfrak{m}})^* = (L^\Psi, \|\cdot\|_{(\Psi),\mathfrak{m}})$ (and the same is true with the norms reversed). In particular if both Ψ and Φ are doubling then L^Φ is reflexive.*

We remark that the double doubling condition is almost necessary to have reflexivity, in the sense that as soon as $\mathfrak{m}$ has also a diffuse part then the reflexivity implies that Φ and Ψ are doubling (see [72], Sec. 1.2).

1.5.2 Properties of the (Φ)-norm

In Chapter 3, we will use a couple of properties of the dual norm; the first one is dealing with the continuity of this norm with respect to the reference measure, while the second one is dealing with the continuity of a character for an N-function Φ. Here we are always assuming that $\mathfrak{m}$ is a σ-finite measure on X.

Lemma 1.5.5 (Representation formula) *For every function g we have*

$$\|g\|_{(\Phi),\mathfrak{m}} = \inf_{k>0} \left\{ \frac{1}{k} \Big(1 + \int_X \Phi(kg)\, \mathrm{d}\mathfrak{m} \Big) \right\}. \tag{1.5.4}$$

In particular $\|g\|_{(\Phi),\mu} \le \max\{1, C\} \|g\|_{(\Phi),\mathfrak{m}}$ whenever $\mu \le C\mathfrak{m}$.

Proof. The representation formula is proved in Sec. 1.2 of [72] (see Equation (24)). The implication is then obvious. □

Definition 1.5.6 (Character of Φ) *We define the character of Φ to be the concave function $A_\Phi : [0, +\infty) \to [0, +\infty)$ defined by*

$$A_\Phi(c) = \inf_{g \in L^\Phi} \left\{ \frac{1 + c \int_X \Phi(g)\, \mathrm{d}\mathfrak{m}}{\|g\|_{(\Phi),\mathfrak{m}}} \right\}.$$

This function is continuous in $(0, 1]$ and has the property that $A_\Phi(1) = 1$ if the measure is finite.

Proof. The concavity and the continuity simply follows from the fact that A_Φ is an infimum of linear positive functions, and $A_\Phi(0) = 0$. Now we want to prove that $A_\Phi(1) = 1$. It is clear that $\|f\|_\Psi \le 1$ if and only if $\int_X \Psi(f) \le 1$ so we can write

$$\frac{1 + \int_X \Phi(g)}{\|g\|_{(\Phi)}} \ge \frac{\int_X \Psi(f) + \int_X \Phi(g)}{\|g\|_{(\Phi)}} \ge \frac{\int_X fg}{\|g\|_{(\Phi)}} \qquad \forall f \in L^\Psi \text{ s.t. } \|f\|_\Psi \le 1$$

taking the supremum over all f and recalling the definition of $\|g\|_{(\Phi)}$ we get precisely

$$\frac{1 + \int_X \Phi(g)}{\|g\|_{(\Phi)}} \ge 1. \tag{1.5.5}$$

Now we need only to show that there exists a function g that realizes equality in (1.5.5). It is sufficient to take $g = k\chi_B$, where B is a set with finite positive measure. Then a simple computation shows that letting $m = \mathfrak{m}(B)$, we have $\|g\|_{(\Phi)} = k \cdot m \cdot \Psi^{-1}(m^{-1})$. In particular we are looking to some k such that

$$\frac{1 + m\Phi(k)}{km\Psi^{-1}(m^{-1})} = 1;$$

If we let $\Psi^{-1}(m^{-1}) = x$ then we can rewrite this equation as

$$\Psi(x) + \Phi(k) = xk,$$

and so it is sufficient to take $k \in \partial^- \Psi(x)$, that is always nonempty. □

1.6 Hopf-Lax formula and Hamilton-Jacobi equation

Aim of this section is to study the properties of the Hopf-Lax formula in a metric space (X, d) and its relations with the Hamilton-Jacobi equation. Notice that there is no reference measure $\mathfrak{m}$ here and that not even completeness is needed for the results of this section. We fix an N-function Ψ and denote by Φ its convex conjugate; we will assume also that Ψ is of class $C^1(\mathbb{R})$ and strictly convex. In the sequel we will follow [9], despite we notice that in [59] and [42] the same results are presented, also in more generality (they don't assume f to be Lipschitz), but they still use similar methods. We notice also that we need Φ to be C^1 in order to achieve Proposition 1.6.4; we don't know whether if (1.6.10) remains true at least in the $\mathfrak{m} \times \mathscr{L}^1$-almost everywhere sense, for every Φ convex. We could avoid the strictly convexity assumption by modifying some propositions along the proofs (in particular it is not true anymore that $D^+(x,t) \le D^-(x,s)$ for $t < s$), but we prefer to keep the exposition simpler.

Let $f : X \to \mathbb{R}$ be a Lipschitz function. For $t > 0$ define

$$F(t,x,y) := f(y) + t\Psi\Big(\frac{\mathsf{d}(x,y)}{t}\Big), \tag{1.6.1}$$

and the function $Q_t f : X \to \mathbb{R}$ by

$$Q_t f(x) := \inf_{y \in X} F(t,x,y). \tag{1.6.2}$$

Notice that $Q_t f(x) \le f(x)$; on the other hand, if L denotes the Lipschitz constant of f, Young's inequality $t\big(\Psi(\mathsf{d}/t) + \Phi(L)\big) \ge L\mathsf{d}$ gives

$$F(t,x,y) \ge f(x) - L\mathsf{d}(x,y) + t\Psi\Big(\frac{\mathsf{d}(x,y)}{t}\Big) \ge f(x) - t\Phi(L),$$

so that $Q_t f(x) \uparrow f(x)$ as $t \downarrow 0$.

Also, we introduce the functions D^+, $D^- : X \times (0,\infty) \to \mathbb{R}$ as

$$\begin{aligned} D^+(x,t) &:= \sup \limsup_{n\to\infty} \mathsf{d}(x,y_n), \\ D^-(x,t) &:= \inf \liminf_{n\to\infty} \mathsf{d}(x,y_n), \end{aligned} \tag{1.6.3}$$

where, in both cases, the sequences (y_n) vary among all minimizing sequences for $F(t,x,\cdot)$. We also set $Q_0 f = f$ and $D^\pm(x,0) = 0$. Arguing as in [8, Lemma 3.1.2] it is easy to check that the map $X \times [0,\infty) \ni (x,t) \mapsto Q_t f(x)$ is continuous. Furthermore, the fact that f is Lipschitz easily yields

$$-LD^\pm + t\Psi\Big(\frac{D^\pm}{t}\Big) \le 0 \qquad \Longrightarrow \qquad \Psi\Big(\frac{D^\pm}{t}\Big) \le L\frac{D^\pm}{t}. \tag{1.6.4}$$

Thanks to the superlinearity of Ψ in the definition of an N-function, we can found $\lambda = \lambda(L)$ such that $\Psi(x) > Lx$ for all $x \ge \lambda$, and so we get

$$D^-(x,t) \le D^+(x,t) \le t\lambda(L) \tag{1.6.5}$$

Proposition 1.6.1 (Monotonicity of $D^{\pm}$) *For all $x \in X$ it holds*

$$D^+(x,t) \le D^-(x,s) \qquad 0 \le t < s. \tag{1.6.6}$$

As a consequence, $D^+(x,\cdot)$ and $D^-(x,\cdot)$ are both nondecreasing, and they coincide with at most countably many exceptions in $[0,\infty)$.

Proof. Fix $x \in X$. For $t = 0$ there is nothing to prove. Now pick $0 < t < s$ and for every $\varepsilon \in (0,1)$ choose $x_{t,\varepsilon}$ and $x_{s,\varepsilon}$ minimizers up to ε of $F(t,x,\cdot)$ and $F(s,x,\cdot)$ respectively, namely such that $F(t,x,x_{t,\varepsilon}) - \varepsilon \le F(t,x,w)$ and $F(s,x,x_{s,\varepsilon}) - \varepsilon \le F(s,x,w)$ for every $w \in X$. Let us assume that $\mathsf{d}(x,x_{t,\varepsilon}) \ge (1-\varepsilon)D^+(x,t)$ and $\mathsf{d}(x,x_{s,\varepsilon}) \le D^-(x,s)+\varepsilon$. The minimality up to ε of $x_{t,\varepsilon}$, $x_{s,\varepsilon}$ gives

$$\begin{aligned} f(x_{t,\varepsilon}) + t\Psi\Big(\frac{\mathsf{d}(x_{t,\varepsilon},x)}{t}\Big) &\le f(x_{s,\varepsilon}) + t\Psi\Big(\frac{\mathsf{d}(x_{s,\varepsilon},x)}{t}\Big) + \varepsilon \\ f(x_{s,\varepsilon}) + s\Psi\Big(\frac{\mathsf{d}(x_{s,\varepsilon},x)}{s}\Big) &\le f(x_{t,\varepsilon}) + s\Psi\Big(\frac{\mathsf{d}(x_{t,\varepsilon},x)}{s}\Big) + \varepsilon. \end{aligned}$$

Adding up we deduce

$$t\Psi\Big(\frac{\mathsf{d}(x_{t,\varepsilon},x)}{t}\Big) - t\Psi\Big(\frac{\mathsf{d}(x_{s,\varepsilon},x)}{t}\Big) \le s\Psi\Big(\frac{\mathsf{d}(x_{t,\varepsilon},x)}{s}\Big) - s\Psi\Big(\frac{\mathsf{d}(x_{t,\varepsilon},x)}{s}\Big) + \varepsilon.$$

Now, letting $\varepsilon \to 0$ we have $\mathsf{d}(x_{t,\varepsilon},x) \to D^+(x,t)$ and $\mathsf{d}(x_{s,\varepsilon},x) \to D^-(x,s)$ and so we can deduce

$$t\Psi\Big(\frac{D^+(x,t)}{t}\Big) - t\Psi\Big(\frac{D^-(x,s)}{t}\Big) \le s\Psi\Big(\frac{D^+(x,t)}{s}\Big) - s\Psi\Big(\frac{D^-(x,s)}{s}\Big). \tag{1.6.7}$$

Let us suppose that $D^+(x,t) > D^-(x,s)$, then dividing by $D^+(x,s) - D^-(x,t)$, and denoting by $\Delta\Psi(w_1,w_2) = \frac{\Psi(w_1)-\Psi(w_2)}{w_1-w_2}$ the difference quotient of Ψ, we can write (1.6.7) as

$$\Delta\Psi\Big(\frac{D^+(x,t)}{t}, \frac{D^-(x,s)}{t}\Big) \le \Delta\Psi\Big(\frac{D^+(x,t)}{s}, \frac{D^-(x,s)}{s}\Big).$$

This is in contradiction with the strict convexity of Ψ since $\Delta\Psi$ is strictly increasing separately in each variable, and $\frac{1}{t} > \frac{1}{s}$.

In the end we obtained (1.6.6). Combining this with the inequality $D^- \le D^+$ we immediately obtain that both functions are nonincreasing. At a point of right continuity of $D^-(x,\cdot)$ we get

$$D^+(x,t) \le \inf_{s>t} D^-(x,s) = D^-(x,t).$$

This implies that the two functions coincide out of a countable set. □

Next, we examine the semicontinuity properties of $D^{\pm}$. These properties imply that points (x,t) where the equality $D^+(x,t) = D^-(x,t)$ occurs are continuity points for both D^+ and D^-.

Proposition 1.6.2 (Semicontinuity of $D^{\pm}$) *D^+ is upper semicontinuous and D^- is lower semicontinuous in $X \times [0,\infty)$.*

Proof. We prove lower semicontinuity of D^-, the proof of upper semicontinuity of D^+ being similar. Let (x_i, t_i) be any sequence converging to (x, t) such that the limit of $D^-(x_i, t_i)$ exists and assume that $t > 0$ (the case $t = 0$ is trivial). For every i, let (y_i^n) be a minimizing sequence of $F(t_i, x_i, \cdot)$ for which $\lim_n \mathsf{d}(y_i^n, x_i) = D^-(x_i, t_i)$, so that

$$\lim_{n\to\infty} f(y_i^n) + t_i \Psi\Big(\frac{\mathsf{d}(y_i^n, x_i)}{t_i}\Big) = Q_{t_i} f(x_i).$$

Using the continuity of Q_t we get

$$\begin{aligned} Q_t f(x) &= \lim_{i\to\infty} \lim_{n\to\infty} f(y_i^n) + t_i \Psi\Big(\frac{\mathsf{d}(y_i^n, x_i)}{t_i}\Big) \\ &\geq \limsup_{i\to\infty} \limsup_{n\to\infty} f(y_i^n) + t\Psi\Big(\frac{\mathsf{d}(y_i^n, x)}{t}\Big) \geq Q_t f(x), \end{aligned}$$

where the first inequality follows from the boundedness of y_i^n and the estimate

$$\Psi\Big(\frac{\mathsf{d}(y_i^n, x_i)}{t_i}\Big) - \Psi\Big(\frac{\mathsf{d}(y_i^n, x)}{t}\Big) \leq \Big(\frac{\mathsf{d}(y_i^n, x_i)}{t_i} - \frac{\mathsf{d}(y_i^n, x)}{t}\Big) \cdot \varphi\Big(\frac{\mathsf{d}(y_i^n, x_i)}{t_i} \vee \frac{\mathsf{d}(y_i^n, x)}{t_i}\Big),$$

which in turn can be proved thanks to the inequality $\Psi(a) - \Psi(b) \leq |a - b|(\varphi(a) \vee \varphi(b))$. Analogously

$$\lim_{i\to\infty} D^-(x_i, t_i) = \lim_{i\to\infty} \lim_{n\to\infty} \mathsf{d}(y_i^n, x_i) \geq \limsup_{i\to\infty} \limsup_{n\to\infty} \mathsf{d}(y_i^n, x).$$

Therefore by a diagonal argument we can find a minimizing sequence $(y_i^{n(i)})$ for $F(t, x, \cdot)$ with $\limsup_i \mathsf{d}(y_i^{n(i)}, x) \leq \lim_i D^-(x_i, t_i)$, which gives the result. $\square$

Proposition 1.6.3 (Time derivative of $Q_t f$) *The map $t \mapsto Q_t f$ is Lipschitz from $[0, \infty)$ to the extended metric space of continuous functions $C(X)$, endowed with the distance*

$$\|f - g\|_\infty = \sup_{x\in X} |f(x) - g(x)|.$$

Moreover, for all $x \in X$, it satisfies:

$$\frac{\mathrm{d}}{\mathrm{d}t} Q_t f(x) = -\Phi\left(\psi\Big(\frac{D^\pm(x, t)}{t}\Big)\right), \tag{1.6.8}$$

for any $t > 0$, with at most countably many exceptions; we recall that $\psi = \Psi'$.

Proof. Let $t < s$ and for every $\varepsilon \in (0, 1)$ choose $x_{t,\varepsilon}$ and $x_{s,\varepsilon}$ minimizers up to ε of $F(t, x, \cdot)$ and $F(s, x, \cdot)$ respectively, namely such that $F(t, x, x_{t,\varepsilon}) - \varepsilon \leq F(t, x, w)$ and $F(s, x, x_{s,\varepsilon}) - \varepsilon \leq F(s, x, w)$ for every $w \in X$. Let us assume that $\mathsf{d}(x, x_{t,\varepsilon}) \geq D^+(x, t) - \varepsilon$ and $\mathsf{d}(x, x_{s,\varepsilon}) \leq D^-(x, s) + \varepsilon$. We have

$$\begin{aligned} Q_s f(x) - Q_t f(x) &\leq F(s, x, x_{t,\varepsilon}) - F(t, x, x_{t,\varepsilon}) + \varepsilon \\ &= t\Psi\Big(\frac{\mathsf{d}(x, x_{t,\varepsilon})}{t}\Big) - s\Psi\Big(\frac{\mathsf{d}(x, x_{t,\varepsilon})}{s}\Big) + \varepsilon, \\ Q_s f(x) - Q_t f(x) &\geq F(s, x, x_{s,\varepsilon}) - F(t, x, x_{s,\varepsilon}) - \varepsilon \\ &= t\Psi\Big(\frac{\mathsf{d}(x, x_{s,\varepsilon})}{t}\Big) - s\Psi\Big(\frac{\mathsf{d}(x, x_{s,\varepsilon})}{s}\Big) - \varepsilon. \end{aligned}$$

For ε small enough, dividing by $s-t$, using the definition of $x_{t,\varepsilon}$ and $x_{s,\varepsilon}$ and using the inequality $\Psi(\frac{\mathsf{d}}{t})-\frac{\mathsf{d}}{t}\psi(\frac{\mathsf{d}}{t})\le\frac{t\Psi(\mathsf{d}/t)-s\Psi(\mathsf{d}/s)}{t-s}\le\Psi(\frac{\mathsf{d}}{s})-\frac{\mathsf{d}}{s}\psi(\frac{\mathsf{d}}{s})$ (note that $t\mapsto t\Psi(\mathsf{d}/t)$ is convex) and using (1.5.2) we obtain

$$\frac{Q_sf(x)-Q_tf(x)}{s-t}\le-\Phi\left(\psi\Big(\frac{D^+(x,t)-\varepsilon}{s}\Big)\right)+\frac{\varepsilon}{s-t},$$
$$\frac{Q_sf(x)-Q_tf(x)}{s-t}\ge-\Phi\left(\psi\Big(\frac{D^-(x,s)+\varepsilon}{t}\Big)\right)-\frac{\varepsilon}{s-t},$$

which gives as $\varepsilon\to 0$ that $t\mapsto Q_tf(x)$ is Lipschitz in $[\delta,T]$ for any $0<\delta<T$ uniformly with respect to $x\in X$. Also, taking Proposition 1.6.1 into account, we get (1.6.8). Now notice that from (1.6.5) we get that $|\frac{\mathrm{d}}{\mathrm{d}t}Q_tf(x)|\le\Phi(\psi(\lambda(\mathrm{Lip}(f))))$ for any $x\in X$ and a.e. $t>0$, which, together with the pointwise convergence of Q_tf to f as $t\downarrow 0$, yields that $t\mapsto Q_tf\in C(X)$ is Lipschitz in $[0,\infty)$. □

We will bound from above the asymptotic Lipschitz constant of Q_tf at x with $\psi(D^+(x,t)/t)$.

Proposition 1.6.4 (Bound on the asymptotic Lipschitz constant of Q_tf) *For $(x,t)\in X\times(0,\infty)$ it holds:*

$$\mathrm{lip}_a(Q_tf,x)\le\psi\Big(\frac{D^+(x,t)}{t}\Big).\tag{1.6.9}$$

In particular $\mathrm{lip}_a(Q_tf)\le\psi(\lambda(\mathrm{Lip}(f)))$, where λ is defined in (1.6.5); if in addition (X,d) is a geodesic metric space then $\mathrm{Lip}(Q_tf)\le\mathrm{Lip}(f)$.

Proof. Fix y, $z\in X$ and $t\in(0,\infty)$. For every $\varepsilon>0$ let $y_\varepsilon\in X$ be such that $F(t,y,y_\varepsilon)-\varepsilon\le F(t,y,w)$ for every $w\in X$ and $|\mathsf{d}(y,y_\varepsilon)-D^+(y,t)|\le\varepsilon$. Since it holds

$$\begin{aligned}Q_tf(z)-Q_tf(y)&\le F(t,z,y_\varepsilon)-F(t,y,y_\varepsilon)+\varepsilon\\&=f(y_\varepsilon)+t\Psi\Big(\frac{\mathsf{d}(z,y_\varepsilon)}{t}\Big)-f(y_\varepsilon)-t\Psi\Big(\frac{\mathsf{d}(y,y_\varepsilon)}{t}\Big)+\varepsilon\\&\le t\Psi\Big(\frac{\mathsf{d}(z,y)+\mathsf{d}(y,y_\varepsilon)}{t}\Big)-t\Psi\Big(\frac{\mathsf{d}(y,y_\varepsilon)}{t}\Big)+\varepsilon\\&\le\mathsf{d}(z,y)\psi\Big(\frac{\mathsf{d}(z,y)+D^+(y,t)+\varepsilon}{t}\Big)+\varepsilon,\end{aligned}$$

so that letting $\varepsilon\to 0$, dividing by $\mathsf{d}(z,y)$ and inverting the roles of y and z gives

$$\mathrm{Lip}\big(Q_tf,B(x,r)\big)\le\psi\Big(\frac{2r+\sup_{y\in B(x,r)}D^+(y,t)}{t}\Big).$$

Letting $r\downarrow 0$ and using the upper semicontinuity of D^+ we get (1.6.9); notice that in this limit it is crucial the continuity of ψ (i.e. the fact that $\Psi\in C^1$).

Finally, the bound on the Lipschitz constant of Q_tf follows directly from (1.6.5) and (1.6.9). For the finer estimate in the geodesic case we refer to [42], [59]. □

Theorem 1.6.5 (Subsolution of HJ) *For every $x\in X$ it holds*

$$\frac{\mathrm{d}}{\mathrm{d}t}Q_tf(x)+\Phi\big(\mathrm{lip}_a(Q_tf,x)\big)\le 0\tag{1.6.10}$$

for every $t\in(0,\infty)$, with at most countably many exceptions.

Proof. The claim is a direct consequence of Propositions 1.6.3 and 1.6.4. □

Notice that (1.6.10) is a stronger formulation of the HJ subsolution property

$$\frac{\mathrm{d}}{\mathrm{d}t} Q_t f(x) + \Phi\big(|\nabla Q_t f|(x)\big) \leq 0, \tag{1.6.11}$$

with the asymptotic Lipschitz constant $\mathrm{lip}_a(Q_t f, \cdot)$ in place of $|\nabla Q_t f|$.

1.7 The space $(\mathcal{M}_+(X), W_\Psi)$ and the superposition principle

Let (X, d) be a complete and separable metric space and let $\mathcal{M}_+(X)$ denote the set of positive and finite Borel measures on X. Given a lower semicontinuous cost $c : X \times X \to [0, \infty]$, we can consider the classical Kantorovich transport problem on X between measures with same mass, defining

$$\mathscr{C}_c(\mu, \nu) := \min\Big\{ \int_{X \times X} c(x, y)\, \mathrm{d}\gamma \mid {\pi^1}_\sharp \gamma = \mu,\ {\pi^2}_\sharp \gamma = \nu \Big\},$$

where π^1 and π^2 are respectively the projections on the first and second factors. We shall denote by $\Gamma(\mu, \nu)$ the collection of admissible plans γ in the Kantorovich minimization problem. In the case of $c_p = \mathsf{d}^p$, $1 \leq p < \infty$, we get the classical Wasserstein distances $W_p = (\mathscr{C}_{c_p})^{1/p}$; they can equivalently be written as

$$W_p(\mu, \nu) = \min\{\|\mathsf{d}\|_{L^p(\gamma)} \mid \gamma \in \Gamma(\mu, \nu)\}$$

and so it is somewhat natural to look at the L^Ψ case, when Ψ is a Young function:

$$W_\Psi(\mu, \nu) := \min\left\{\|\mathsf{d}\|_{L^\Psi(\gamma)} \mid \gamma \in \Gamma(\mu, \nu)\right\}.$$

We can recover in this way also the distance W_∞, setting $\Psi(x) = 0$ if $|x| \leq 1$ and $\Psi(x) = +\infty$ otherwise. We want to consider also this general L^Ψ case as a transport problem, in order to have a dual formulation that will be used later on. Notice that this Orlicz-Wasserstein distance was already introduced in [77] and subsequently developed in [55]; they looked for properties of W_Ψ more related to classical optimal transport, while here we focus on the duality. The key point is to consider scaled costs: we introduce the "test" distances

$$W_\Psi^{(s)}(\mu, \nu) = \min\Big\{ \int_{X \times X} s\Psi\Big(\frac{\mathsf{d}(x, y)}{s}\Big) \mathrm{d}\gamma \mid \gamma \in \Gamma(\mu, \nu) \Big\},$$

called this way because

$$W_\Psi^{(s)}(\mu, \nu) \leq s \quad \Longleftrightarrow \quad W_\Psi(\mu, \nu) \leq s \qquad\qquad \text{for all } s > 0. \tag{1.7.1}$$

These "test" distances are given by transport problems with lower semicontinuous costs $c_s(x, y) = s\Psi(\mathsf{d}(x, y)/s)$, so they have a dual formulation [8, Theorem 6.1.1][1]:

$$W_\Psi^{(s)}(\mu, \nu) = \sup_{\psi \in \mathrm{Lip}_b(X)} \int_X \xi^c \, \mathrm{d}\mu + \int_X \xi \, \mathrm{d}\nu,$$

[1] in the reference they prove it that functions in $C_b^0(X, \mathsf{d})$ are sufficient, but if once we fix $\varphi \in C_b^0(X)$ then we can take ψ to be the c_h-transform of φ for some h (following their notation); but then ψ is h-Lipschitz and bounded. Then again we can substitute φ with the c_k-transform of ψ for some k, and so also φ is Lipschitz and bounded.

where ψ^c denotes the so called c-transform of f, defined as:

$$\xi^c(x) = \inf_{y\in X} \Big\{ c_s(x,y) - \xi(y) \Big\}$$

(namely the largest function $g(x)$ satisfying $g(x) + \psi(y) \le c_s(x,y)$ for all (x,y)). By the definition of $Q_s\varphi$ given in the previous section we get:

$$\xi^c(y) = \inf_{y\in X} \Big\{ s\Psi\left(\frac{\mathsf{d}(x,y)}{s}\right) - \xi(y) \Big\} = Q_s(-\xi)(x).$$

Now, setting $\xi = -\varphi$ in the dual formulation, and using this characterization of the c-transform, we get

$$W_\Psi^{(s)}(\mu,\nu) = \sup_{\varphi\in \mathrm{Lip}_b(X)} \int_X Q_s\varphi \, \mathrm{d}\mu - \int_X \varphi \, \mathrm{d}\nu. \tag{1.7.2}$$

The last step we need is to pass to $\varphi \in \mathrm{Lip}_0(X,\mathsf{d})$, $\varphi \ge 0$:

Lemma 1.7.1 *Fix $s > 0$. For every $\mu,\nu \in \mathcal{M}_+(X)$ with the same mass we have*

$$W_\Psi^{(s)}(\mu,\nu) = \sup\left\{ \int_X Q_s\varphi \, \mathrm{d}\mu - \int_X \varphi \, \mathrm{d}\nu \ : \ \varphi \in \mathrm{Lip}_0(X,\mathsf{d}), \ \varphi \ge 0 \right\}. \tag{1.7.3}$$

Proof. In order to prove the equivalence, given (1.7.2), it is easy to see that, up to translation, one can choose $\varphi \ge 0$; therefore it is enough to show that for every $\varphi \in \mathrm{Lip}_b(X)$ nonnegative there holds

$$\liminf_{r\to\infty} \left\{ \int_X Q_s[\chi_r\varphi] \, \mathrm{d}\nu - \int_X \chi_r\varphi \, \mathrm{d}\mu \right\} \ge \int_X Q_s\varphi \, d\nu - \int_X \varphi \, \mathrm{d}\mu, \tag{1.7.4}$$

where χ_r is a Lipschitz cutoff function which is nonnegative, identically equal to 1 in $B(x_0,r)$ and identically equal to 0 outside $B(x_0,r+1)$ for some $x_0 \in X$ fixed. Since $\chi_r\varphi \le \varphi$ it follows that $\int_X \chi_r\varphi \, \mathrm{d}\mu \le \int_X \varphi \, \mathrm{d}\mu$, so that by Fatou's lemma suffices to show that $\liminf_{r\to\infty} Q_s[\chi_r\varphi] \ge Q_s\varphi$. Let $x \in X$ be fixed and let $x_r \in X$ be satisfying

$$\chi_r(x_r)\varphi(x_r) + s\Psi\Big(\frac{\mathsf{d}(x,x_r)}{s}\Big) \le \frac{1}{r} + Q_s[\chi_r\varphi](x).$$

Since $\mathsf{d}(x_r,x)$ is obviously bounded as $r \to \infty$, the same is true for $\mathsf{d}(x_r,x_0)$, so that $\chi_r(x_r) = 1$ for r large enough and $Q_s\varphi(x) \le r^{-1} + Q_s[\chi_r\varphi](x)$ for r large enough.

□

We will need also the following result, proved in [62]: it shows how to associate to an absolutely continuous curve μ_t w.r.t. W_Ψ a plan $\boldsymbol{\pi} \in \mathcal{P}(C([0,1],X))$ representing the curve itself (see also [8, Theorem 8.2.1] for the Euclidean case and [61] for the general L^p case). This is not possible for any Young function (for example it fails for $\Psi(x) = x$); we need the following conditions to hold:

$$\lim_{x\to 0} \frac{\Psi(x)}{x} = \lim_{x\to\infty} \frac{x}{\Psi(x)} = 0 \tag{1.7.5}$$

In particular the superposition principle holds for every N-function.

Proposition 1.7.2 (Superposition principle) *Let (X,d) be a complete and separable metric space, Ψ a Young function satisfying (1.7.5), and let $\mu_t \in AC\big([0,T];(\mathcal{P}(X),W_\Psi)\big)$. Then there exists $\boldsymbol{\pi} \in \mathcal{P}(C([0,1],X))$, concentrated on $AC([0,1],X)$, such that $(\mathrm{e}_t)_\sharp\boldsymbol{\pi} = \mu_t$ for any $t \in [0,T]$ and*

$$\|\|\dot\gamma_t\|\|_{L^\Psi(\boldsymbol{\pi})} = |\dot\mu_t| \qquad \text{for a.e. } t \in [0,T]. \tag{1.7.6}$$

1.8 Γ-convergence

Definition 1.8.1 *Let (X, d) be a metric space and let $F_h : X \to [-\infty, +\infty]$. We say that F_h Γ-converge to $F : X \to [-\infty, +\infty]$ if:*

(a) For every sequence $(u_h) \subset X$ convergent to $u \in X$ we have

$$F(u) \leq \liminf_{h\to\infty} F_h(u_h);$$

(b) For all $u \in X$ there exists a sequence $(u_n) \subset X$ such that

$$F(u) \geq \limsup_{h\to\infty} F_h(u_h).$$

Sequences satisfying the second property are called "recovery sequences"; whenever Γ-convergence occurs, they obviously satisfy $\lim_h F_h(u_h) = F(u)$.

The following compactness property of Γ-convergence (see for instance [31, Theorem 8.5]) is well-known.

Proposition 1.8.2 *If (X, d) is separable, any sequence of functionals $F_h : X \to [-\infty, +\infty]$ admits a Γ-convergent subsequence.*

We quickly sketch the proof, for the reader's convenience. If $\{U_i\}_{i\in\mathbb{N}}$ is a countable basis of open sets of (X, d), we may extract a subsequence $h(k)$ such that $\alpha_i := \lim_k \inf_{U_i} F_{h(k)}$ exists in $\overline{\mathbb{R}}$ for all $i \in \mathbb{N}$. Then, it is easily seen that

$$F(x) := \sup_{U_i \ni x} \alpha_i \qquad x \in X$$

is the Γ-limit of $F_{h(k)}$.

We will also need an elementary stability property of uniformly convex (and quadratic as well) functionals under Γ-convergence. Recall that a positively 1-homogeneous function $\mathcal{N}$ on a vector space V is uniformly convex with modulus ω if there exists a function $\omega : [0, \infty) \to [0, \infty)$ with $\omega > 0$ on $(0, \infty)$ such that

$$\mathcal{N}(u) = \mathcal{N}(v) = 1 \qquad \Longrightarrow \qquad \mathcal{N}\left(\frac{u+v}{2}\right) \leq 1 - \omega(\mathcal{N}(u-v))$$

for all u, $v \in V$.

Lemma 1.8.3 *Let V be a normed space with the induced metric structure and let $\omega : [0, \infty) \to [0, \infty)$ be continuous, nondecreasing, positive on $(0, \infty)$. Let $\mathcal{N}_h$ be uniformly convex positively 1-homogeneous functions on V with the same modulus ω, Γ-convergent to some function $\mathcal{N}$. Then $\mathcal{N}$ is positively 1-homogeneous and uniformly convex with modulus ω.*

Proof. The verification of 1-homogeneity of $\mathcal{N}$ is trivial. Let u, $v \in V$ which satisfy $\mathcal{N}(u) = \mathcal{N}(v) = 1$. Let (u_h) and (v_h) be recovery sequences for u and v respectively, so that both $\mathcal{N}_h(u_h)$ and $\mathcal{N}_h(v_h)$ converge to 1. Hence, $u'_h = u_h/\mathcal{N}_h(u_h)$ and $v'_h = v_h/\mathcal{N}_h(v_h)$ still converge to u and v respectively. By assumption

$$\mathcal{N}_h\left(\frac{u'_h + v'_h}{2}\right) + \omega(\mathcal{N}_h(u'_h - v'_h)) \leq 1.$$

Thanks to property (a) of Γ-convergence, the monotonicity and the continuity of ω and the superadditivity of $\liminf$ we get

$$\begin{aligned}\mathcal{N}\left(\frac{u+v}{2}\right)+\omega\left(\mathcal{N}(u-v)\right) &\le \liminf_{h\to\infty}\mathcal{N}_h\left(\frac{u_h'+v_h'}{2}\right)+\omega\left(\liminf_{h\to\infty}\mathcal{N}_h(u_h'-v_h')\right)\\ &\le \liminf_{h\to\infty}\left(\mathcal{N}_h\left(\frac{u_h'+v_h'}{2}\right)+\omega(\mathcal{N}_h(u_h'-v_h'))\right)\le 1.\end{aligned}$$

□

1.9 Doubling metric measure spaces and maximal functions

From now on, $B(x,r)$ will denote the open ball centered in x of radius r and $\bar{B}(x,r)$ will denote the closed ball:

$$B(x,r)=\{y\in X\ :\ \mathsf{d}(x,y)<r\}\quad,\qquad \bar{B}(x,r)=\{y\in X\ :\ \mathsf{d}(x,y)\le r\}.$$

If not specified, with the term ball we mean the open one.

Recall that a metric space (X,d) is doubling if there exists a natural number c_D such that every ball of radius r can be covered by at most c_D balls of halved radius $r/2$.

Definition 1.9.1 (Doubling m.m. spaces) *The metric measure space* $(X,\mathsf{d},\mathfrak{m})$ *is doubling if there exists* $\tilde{c}_D\ge 0$ *such that*

$$\mathfrak{m}(B(x,2r))\le\tilde{c}_D\mathfrak{m}(B(x,r))\qquad\forall x\in\operatorname{supp}\mathfrak{m},\ r>0. \tag{1.9.1}$$

This condition is easily seen to be equivalent to the existence of two real positive numbers α, $\beta>0$ which depend only on $\tilde{c}_D$ such that

$$\mathfrak{m}(B(x,r_1))\le\beta\left(\frac{r_1}{r_2}\right)^{\alpha}\mathfrak{m}(B(y,r_2))\quad\text{whenever } B(y,r_2)\subset B(x,r_1),\ r_2\le r_1,\ y\in\operatorname{supp}\mathfrak{m}. \tag{1.9.2}$$

Indeed, $B(x,r_1)\subset B(y,2r_1)$, hence $\mathfrak{m}(B(x,r_1))\le\tilde{c}_D^k\mathfrak{m}(B(y,r_2))$, where k is the smallest integer such that $2r_1\le 2^k r_2$. Since $k\le 2+\ln_2(r_1/r_2)$, we obtain (1.9.2) with $\alpha=\ln_2\tilde{c}_D$ and $\beta=\tilde{c}_D^2$.

Condition (1.9.2) is stronger than the metric doubling property, in the sense that $(\operatorname{supp}\mathfrak{m},\mathsf{d})$ is doubling whenever $(X,\mathsf{d},\mathfrak{m})$ is. Indeed, given a ball $B(x,r)$ with $x\in\operatorname{supp}\mathfrak{m}$, let us choose recursively points $x_i\in B(x,r)\cap\operatorname{supp}\mathfrak{m}$ with $\mathsf{d}(x_i,x_j)\ge r/2$, and assume that this is possible for $i=1,\ldots,N$. Then, the balls $B(x_i,r/4)$ are disjoint and

$$\mathfrak{m}\big(B\big(x_i,\frac{r}{4}\big)\big)\ge\tilde{c}_D^{-3}\mathfrak{m}(B(x_i,2r))\ge\tilde{c}_D^{-3}\mathfrak{m}(B(x,r)),$$

so that $N\le\tilde{c}_D^3$; in particular we can find a maximal finite set $\{x_i\}$ with this property, and from the maximality it follows that for every $x'\in B(x,r)\cap\operatorname{supp}\mathfrak{m}$ we have $\mathsf{d}(x_i,x')<r/2$ and so

$$B(x,r)\cap\operatorname{supp}\mathfrak{m}\subset\bigcup_i B(x_i,r/2).$$

It follows that $(\operatorname{supp}\mathfrak{m},\mathsf{d})$ is doubling, with doubling constant $c_D\le\tilde{c}_D^3$. Conversely (but we shall not need this fact) any complete doubling metric space supports a nontrivial doubling measure (see [28], [64]).

Definition 1.9.2 (Local maximal function) *Given $q \in [1,\infty)$, $\varepsilon > 0$ and a Borel function $f : X \to \mathbb{R}$ such that $|f|^q$ is $\mathfrak{m}$-integrable on bounded sets, we define the ε-maximal function*

$$M_q^\varepsilon f(x) := \left(\sup_{0<r\le\varepsilon} ⨍_{B(x,r)} |f|^q \, d\mathfrak{m} \right)^{1/q} \qquad x \in \operatorname{supp}\mathfrak{m}.$$

The function $M_q^\varepsilon f(x)$ is nondecreasing w.r.t. ε, moreover $M_q^\varepsilon f(x) \to |f|(x)$ at any Lebesgue point x of $|f|^q$, namely a point $x \in \operatorname{supp}\mathfrak{m}$ satisfying

$$\lim_{r\downarrow 0} \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} |f(y)|^q \, d\mathfrak{m}(y) = |f(x)|^q. \tag{1.9.3}$$

We recall that, in doubling metric measure spaces (see for instance [47]), under the previous assumptions on f we have that $\mathfrak{m}$-a.e. point is a Lebesgue point of $|f|^q$ (the proof is based on the so-called Vitali covering lemma). By applying this property to $|f-s|^q$ with $s \in \mathbb{Q}$ one even obtains

$$\lim_{r\downarrow 0} \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|^q \, d\mathfrak{m}(y) = 0 \tag{1.9.4}$$

for every $x \in \operatorname{supp}\mathfrak{m}$ that is a Lebesgue point of $|f-s|^q$ for every $s \in \mathbb{Q}$. In particular it is clear that (1.9.4) is satisfied for $\mathfrak{m}$-a.e. $x \in \operatorname{supp}\mathfrak{m}$; we call such points q-Lebesgue points of f. We shall need a further enforcement of the q-Lebesgue point property:

Lemma 1.9.3 *Let $(X, \mathsf{d}, \mathfrak{m})$ be a doubling metric measure space and let $f : X \to \mathbb{R}$ be a Borel function such that $|f|^q$ is $\mathfrak{m}$-integrable on bounded sets. Then, at any point x where (1.9.4) is satisfied, it holds*

$$\lim_{n\to\infty} \frac{1}{\mathfrak{m}(E_n)} \int_{E_n} |f(y) - f(x)|^q \, d\mathfrak{m}(y) = 0 \tag{1.9.5}$$

whenever $E_n \subset X$ are Borel sets satisfying $B(y_n, \tau r_n) \subset E_n \subset B(x, r_n)$ with $y_n \in \operatorname{supp}\mathfrak{m}$ and $r_n \to 0$, for some $\tau \in (0,1]$ independent of n. In particular $⨍_{E_n} f \, d\mathfrak{m} \to f(x)$.

Proof. Since $\mathfrak{m}$ is doubling we can use (1.9.2) to obtain

$$\begin{aligned}
\frac{1}{\mathfrak{m}(E_n)} \int_{E_n} |f(y) - f(x)|^q \, d\mathfrak{m}(y) &\le \frac{1}{\mathfrak{m}(B(y_n,\tau r_n))} \int_{E_n} |f(y) - f(x)|^q \, d\mathfrak{m}(y) \\
&\le \frac{1}{\mathfrak{m}(B(y_n,\tau r_n))} \int_{B(x,r_n)} |f(y) - f(x)|^q \, d\mathfrak{m}(y) \\
&\le \frac{\mathfrak{m}(B(x,r_n))}{\mathfrak{m}(B(y_n,\tau r_n))} ⨍_{B(x,r_n)} |f(y) - f(x)|^q \, d\mathfrak{m}(y) \\
&\le \beta\tau^{-\alpha} ⨍_{B(x,r_n)} |f(y) - f(x)|^q \, d\mathfrak{m}(y).
\end{aligned}$$

Since (1.9.4) is true by hypothesis, the last term goes to 0, and we proved (1.9.5). Finally, by Jensen's inequality,

$$\left| ⨍_{E_n} f \, d\mathfrak{m} - f(x) \right|^q \le ⨍_{E_n} |f - f(x)|^q \, d\mathfrak{m} \to 0.$$

□

CHAPTER 2

Duality between p-Modulus and probability measures

For the reader's convenience we collect in the next table and figure the main notation used, mostly in the second part of the chapter; most of them have been already introduced in the preliminaries, but we prefer to give also a reference here, with all the relations between them.

Main notation

$\mathcal{L}^p_+(X,\mathfrak{m})$	Borel nonnegative functions $f : X \to [0,\infty]$ with $\int_X f^p \, \mathrm{d}\mathfrak{m} < \infty$
$L^p(X,\mathfrak{m})$	Lebesgue space of p-summable $\mathfrak{m}$-measurable functions
$\ell(\gamma)$	Length of a parametric curve γ
$\mathrm{AC}^q([0,1];X)$	Space of parametric curves $\gamma : [0,1] \to X$ with q-integrable metric speed
$\mathrm{AC}_0([0,1];X)$	Space of parametric curves with positive speed $\mathscr{L}^1$-a.e. in $(0,1)$
$\mathrm{AC}^\infty_c([0,1];X)$	Space of parametric curves with positive and constant speed
k	Embedding of $\big\{\gamma \in \mathrm{AC}([0,1];X) :\ \ell(\gamma) > 0\big\}$ into $\mathrm{AC}^\infty_c([0,1];X)$
$\mathscr{C}(X)$	Space of non-parametric and nonconstant curves, see Definition 1.2.5
i	Embedding of $\big\{\gamma \in \mathrm{AC}([0,1];X) :\ \ell(\gamma) > 0\big\}$ in $\mathscr{C}(X)$
j	Embedding of $\mathscr{C}(X)$ into $\mathrm{AC}^\infty_c([0,1];X)$
$\mathcal{M}_+(X)$	Space of nonnegative finite Radon measures on X
J	Embedding of $\big\{\gamma \in \mathrm{AC}([0,1];X) :\ \ell(\gamma) > 0\big\}$ in $\mathcal{M}_+(X)$, see Definition 1.2.3
$\tilde{J}$	Embedding of $\mathscr{C}(X)$ in $\mathcal{M}_+(X)$; quotient map of J, see (1.2.10)
M	Embedding of $C([0,1];X)$ in $\mathcal{M}_+(X)$ via push forward of Lebesgue measure, see (2.5.1)

$$
\begin{tikzcd}
\mathrm{AC}^\infty_c([0,1];X) \arrow[r, shift left, "\pi_{\mathscr{C}}"] & \mathscr{C}(X) \arrow[l, shift left, "\mathsf{j}"] \arrow[ddr, "\tilde{J}"] & \\
\{\gamma \in \mathrm{AC}([0,1];X)\} \ : \ \ell(\gamma) > 0\} \arrow[u, "\mathsf{k}"] \arrow[ur, "\mathsf{i}"'] \arrow[drr, "J"] \arrow[d, hook] & & \\
\mathrm{C}([0,1];X) \arrow[rr, "M"] & & \mathcal{M}_+(X)
\end{tikzcd}
$$

In this Chapter we will prove a duality result for the p-modulus $\mathrm{Mod}_{p,\mathfrak{m}}$. It can be stated as follows: a set of measures Σ is not $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible if there exist a probability measure $\boldsymbol{\eta}$ concentratedon Σ and a function $f \in L^q(X,\mathfrak{m})$, called barycenter of $\boldsymbol{\eta}$, such that

$$\iint_X g \,\mathrm{d}\mu \,\mathrm{d}\boldsymbol{\eta} = \int fg \,\mathrm{d}\mathfrak{m} \qquad \forall g \in C_b(X,\mathsf{d}).$$

Quantitatively we have also $\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma) = \max \|f\|_q^{-1}$, where we take the maximum among all such barycenters. As a byproduct we obtain that $\mathrm{Mod}_{p,\mathfrak{m}}$ is a capacity.

Then we specialize this kind of measures on set of curves, and then we compare them with the so-called q-plans, that are used in [9] to define Sobolev function in abstract metric spaces. In particular the notion of negligibility is compared and in the last section we show that the definition of Sobolev function given in [9] (using the q-plans) coincides with the definition given in [57], [75] (using the p-modulus). This result is not new since in [9] the authors show the equivalence also with other definitions, but the method is new, relying on a fine analysis of the structure of the set of curves where the upper gradient property fails.

2.1 $(p,\mathfrak{m})$-modulus $\mathrm{Mod}_{p,\mathfrak{m}}$

In this section (X,τ) is a topological space and $\mathfrak{m}$ is a fixed Borel and nonnegative reference measure, not necessarily finite or σ-finite.

Given a power $p \in [1,\infty)$, we set

$$\mathscr{L}^p_+(X,\mathfrak{m}) := \left\{ f : X \to [0,\infty] \ : \ f \text{ Borel}, \int_X f^p \,\mathrm{d}\mathfrak{m} < \infty \right\}. \tag{2.1.1}$$

We stress that, unlike $L^p(X,\mathfrak{m})$, this space is not quotiented under any equivalence relation; however we will keep using the notation

$$\|f\|_p := \left(\int_X |f|^p \,\mathrm{d}\mathfrak{m} \right)^{1/p}$$

as a seminorm on $\mathscr{L}^p_+(X,\mathfrak{m})$ and a norm in $L^p(X,\mathfrak{m})$.

Given $\Sigma \subseteq \mathcal{M}_+$ we define (with the usual convention $\inf \emptyset = \infty$)

$$\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma) := \inf \left\{ \int_X f^p \,\mathrm{d}\mathfrak{m} \ : \ f \in \mathscr{L}^p_+(X,\mathfrak{m}), \ \int_X f \,\mathrm{d}\mu \geq 1 \quad \text{for all } \mu \in \Sigma \right\}, \tag{2.1.2}$$

$$\mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma) := \inf \left\{ \int_X f^p \,\mathrm{d}\mathfrak{m} \ : \ f \in \mathrm{C}_b(X), \ \int_X f \,\mathrm{d}\mu \geq 1 \quad \text{for all } \mu \in \Sigma \right\}. \tag{2.1.3}$$

Equivalently, if $0 < \mathrm{Mod}_{p,\mathfrak{m}}(\Sigma) \le \infty$, we can say that $\mathrm{Mod}_p(\Sigma)^{-1}$ is the least number $\xi \in [0,\infty)$ such that the following is true

$$\left(\inf_{\mu\in\Sigma}\int_X f\,\mathrm{d}\mu\right)^p \le \xi\int_X f^p\,\mathrm{d}\mathfrak{m} \quad \text{for all } f \in \mathcal{L}^p_+(X,\mathfrak{m}), \tag{2.1.4}$$

and similarly there is also an equivalent definition for $\mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma)^{-1}$.
Notice that the infimum in (2.1.3) is unchanged if we restrict the minimization to nonnegative functions $f \in \mathrm{C}_b(X)$. As a consequence, since the finiteness of $\mathfrak{m}$ provides the inclusion of this class of functions in $\mathcal{L}^p_+(X,\mathfrak{m})$, we get $\mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma) \ge \mathrm{Mod}_{p,\mathfrak{m}}(\Sigma)$ whenever $\mathfrak{m}$ is finite. Also, whenever Σ contains the null measure, we have $\mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma) \ge \mathrm{Mod}_{p,\mathfrak{m}}(\Sigma) = \infty$.

Definition 2.1.1 ($\mathrm{Mod}_{p,\mathfrak{m}}$-negligible sets) *A set $\Sigma \subseteq \mathcal{M}_+(X)$ is said to be $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible if $\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma) = 0$.*

A property P on $\mathcal{M}_+(X)$ is said to be hold $\mathrm{Mod}_{p,\mathfrak{m}}$-a.e. if the set

$$\{\mu \in \mathcal{M}_+(X) :\ P(\mu) \text{ fails}\}$$

is $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible. With this terminology, we can also write

$$\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma) = \inf\left\{\int_X f^p\,\mathrm{d}\mathfrak{m} \ :\ \int_X f\,\mathrm{d}\mu \ge 1 \quad \text{for } \mathrm{Mod}_{p,\mathfrak{m}}\text{-a.e. } \mu \in \Sigma\right\}. \tag{2.1.5}$$

We list now some classical properties that will be useful in the sequel, most of them are well known and simple to prove, but we provide complete proofs for the reader's convenience.

Proposition 2.1.2 *The set functions $A \subseteq \mathcal{M}_+(X) \mapsto \mathrm{Mod}_{p,\mathfrak{m}}(A)$, $A \subseteq \mathcal{M}_+(X) \mapsto \mathrm{Mod}_{p,\mathfrak{m},c}(A)$ satisfy the following properties:*

(i) *both are monotone and their $1/p$-th power is subadditive;*

(ii) *if $g \in \mathcal{L}^p_+(X,\mathfrak{m})$ then $\int_X g\,\mathrm{d}\mu < \infty$ for $\mathrm{Mod}_{p,\mathfrak{m}}$-almost every μ; conversely, if $\mathrm{Mod}_{p,\mathfrak{m}}(A) = 0$ then there exists $g \in \mathcal{L}^p_+(X,\mathfrak{m})$ such that $\int_X g\,\mathrm{d}\mu = \infty$ for every $\mu \in A$.*

(iii) *if $(f_n) \subset \mathcal{L}^p_+(X,\mathfrak{m})$ converges in $L^p(X,\mathfrak{m})$ seminorm to $f \in \mathcal{L}^p_+(X,\mathfrak{m})$, there exists a subsequence $(f_{n(k)})$ such that*

$$\int_X f_{n(k)}\,\mathrm{d}\mu \to \int_X f\,\mathrm{d}\mu \qquad \mathrm{Mod}_{p,\mathfrak{m}}\text{-a.e. in } \mathcal{M}_+(X); \tag{2.1.6}$$

(iv) *if $p > 1$, for every $\Sigma \subseteq \mathcal{M}_+(X)$ with $\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma) < \infty$ there exists $f \in \mathcal{L}^p_+(X,\mathfrak{m})$, unique up to $\mathfrak{m}$-negligible sets, such that $\int_X f\,\mathrm{d}\mu \ge 1$ $\mathrm{Mod}_{p,\mathfrak{m}}$-a.e. on Σ and $\|f\|_p^p = \mathrm{Mod}_{p,\mathfrak{m}}(\Sigma)$;*

(v) *if $p > 1$ and A_n are nondecreasing subsets of $\mathcal{M}_+(X)$ then $\mathrm{Mod}_{p,\mathfrak{m}}(A_n) \uparrow \mathrm{Mod}_{p,\mathfrak{m}}(\cup_n A_n)$;*

(vi) *if K_n are nonincreasing compact subsets of $\mathcal{M}_+(X)$ then $\mathrm{Mod}_{p,\mathfrak{m},c}(K_n) \downarrow \mathrm{Mod}_{p,\mathfrak{m},c}(\cap_n K_n)$.*

(vii) *Let $A \subseteq \mathcal{M}_+(X)$, $F : A \to (0,\infty)$ be a Borel map, and $B = \big\{F(\mu)\mu :\ \mu \in A\big\}$. If $\mathrm{Mod}_{p,\mathfrak{m}}(A) = 0$ then $\mathrm{Mod}_{p,\mathfrak{m}}(B) = 0$ as well.*

Proof. (i) Monotonicity is trivial. For the subadditivity, if we take $\int_X f \,\mathrm{d}\mu \geq 1$ on A and $\int_X g \,\mathrm{d}\mu \geq 1$ on B, then $\int_X (f+g) \,\mathrm{d}\mu \geq 1$ on $A \cup B$, hence $\mathrm{Mod}_{p,\mathfrak{m}}(A \cup B)^{1/p} \leq \|f+g\|_p \leq \|f\|_p + \|g\|_p$. Minimizing over f and g we get the subadditivity.

(ii) Let us consider the set where the property fails:

$$\Sigma_g = \left\{ \mu \in \mathcal{M}_+(X) \ : \ \int_X g \,\mathrm{d}\mu = \infty \right\}.$$

Then it is clear that $\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma_g) \leq \|g\|_p^p$ but $\Sigma_g = \Sigma_{\lambda g}$ for every $\lambda > 0$ and so we get that Σ_g is $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible. Conversely, if $\mathrm{Mod}_{p,\mathfrak{m}}(A) = 0$ for every $n \in \mathbb{N}$ we can find $g_n \in \mathcal{L}^p_+(X, \mathfrak{m})$ with $\int_X g_n \,\mathrm{d}\mu \geq 1$ for every $\mu \in A$ and $\int_X g_n^p \leq 2^{-np}$. Thus $g := \sum_n g_n$ satisfies the required properties.

(iii) Let $f_{n(k)}$ be a subsequence such that $\|f - f_{n(k)}\|_p \leq 2^{-k}$ so that if we set

$$g(x) = \sum_{k=1}^{\infty} |f(x) - f_{n(k)}(x)|$$

we have that $g \in \mathcal{L}^p_+(X, \mathfrak{m})$ and $\|g\|_p \leq 1$; in particular we have, for (ii) above, that $\int_X g \,\mathrm{d}\mu$ is finite for $\mathrm{Mod}_{p,\mathfrak{m}}$-almost every μ. For those μ we get

$$\sum_{k=1}^{\infty} \int_X |f - f_{n(k)}| \,\mathrm{d}\mu < \infty$$

and thus we get (2.1.6).

(iv) Since we can use (2.1.5) to compute $\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma)$, we obtain from (ii) and (iii) that the class of admissible functions f is a convex and closed subset of the Lebesgue space L^p. Hence, uniqueness follows by the strict convexity of the L^p norm.

(v) By the monotonicity, it is clear that $\mathrm{Mod}_{p,\mathfrak{m}}(A_n)$ is an increasing sequence and that $\mathrm{Mod}_{p,\mathfrak{m}}(\cup_n A_n) \geq \lim \mathrm{Mod}_{p,\mathfrak{m}}(A_n) =: C$. If $C = \infty$ there is nothing to prove, otherwise, we need to show that $\mathrm{Mod}_{p,\mathfrak{m}}(\cup_n A_n) \leq C$; let $(f_n) \subset \mathcal{L}^p_+(X, \mathfrak{m})$ be a sequence of functions such that $\int_X f_n \,\mathrm{d}\mu \geq 1$ on A_n and $\|f_n\|_p^p \leq \mathrm{Mod}_{p,\mathfrak{m}}(A_n) + \frac{1}{n}$. In particular we get that $\limsup_n \|f_n\|_p^p = C < \infty$ and so, possibly extracting a subsequence, we can assume that f_n weakly converge to some $f \in \mathcal{L}^p_+(X, \mathfrak{m})$. By Mazur lemma we can find convex combinations

$$\hat{f}_n = \sum_{k=n}^{\infty} \lambda_{k,n} f_k$$

such that $\hat{f}_n$ converge strongly to f in $L^p(X, \mathfrak{m})$; furthermore we have that $\int_X f_k \,\mathrm{d}\mu \geq 1$ on A_n if $k \geq n$ and so

$$\int_X \hat{f}_n \,\mathrm{d}\mu = \sum_{k=n}^{\infty} \lambda_{k,n} \int_X f_k \,\mathrm{d}\mu \geq 1 \qquad \text{on } A_n.$$

By (iii) in this proposition we obtain a subsequence $n(k)$ and a $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible set $\Sigma \subseteq \mathcal{M}_+(X)$ such that $\int_X \hat{f}_{n(k)} \,\mathrm{d}\mu \to \int_X f \,\mathrm{d}\mu$ outside Σ; in particular $\int_X f \,\mathrm{d}\mu \geq 1$ on $\cup_n A_n \setminus \Sigma$. Then, by the very definition of $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible set, for every $\varepsilon > 0$ we can find $g_\varepsilon \in \mathcal{L}^p_+(X, \mathfrak{m})$ such that $\|g_\varepsilon\|_p^p \leq \varepsilon$ and $\int_X g_\varepsilon \,\mathrm{d}\mu \geq 1$ on Σ, so that we have $\int_X (f + g_\varepsilon) \,\mathrm{d}\mu \geq 1$ on $\cup_n A_n$ and

$$\mathrm{Mod}_{p,\mathfrak{m}}(\cup_n A_n)^{1/p} \leq \|g_\varepsilon + f\|_p \leq \|g_\varepsilon\|_p + \|f\|_p \leq \varepsilon^{1/p} + \liminf \|f_n\|_p \leq \varepsilon^{1/p} + C^{1/p}.$$

Letting $\varepsilon \to 0$ and taking the p-th power the inequality $\mathrm{Mod}_{p,\mathfrak{m}}(A) \le \sup_n \mathrm{Mod}_{p,\mathfrak{m}}(A_n)$ follows.

(vi) As before, by the monotonicity we get $\mathrm{Mod}_{p,\mathfrak{m},c}(K) \le \mathrm{Mod}_{p,\mathfrak{m},c}(K_n)$ and so calling C the limit of $\mathrm{Mod}_{p,\mathfrak{m},c}(K_n)$ as n goes to infinity, we only have to prove $\mathrm{Mod}_{p,\mathfrak{m},c}(K) \ge C$. First, we deal with the case $\mathrm{Mod}_{p,\mathfrak{m},c}(K) > 0$: using the equivalent definition, let $\varphi_\varepsilon \in \mathrm{C}_b(X)$ be such that $\|\varphi_\varepsilon\|_p = 1$ and

$$\inf_{\mu\in K} \int_X \varphi_\varepsilon \,\mathrm{d}\mu \ge \frac{1}{\mathrm{Mod}_{p,\mathfrak{m},c}(K)^{1/p}} - \varepsilon.$$

By the compactness of K and of K_n, it is clear that the infimum above is a minimum and that $\min\limits_{K_n} \int_X \varphi_\varepsilon \,d\mu \to \min\limits_{K} \int_X \varphi_\varepsilon \,\mathrm{d}\mu$, so that

$$\frac{1}{C^{1/p}} = \lim_{n\to\infty} \frac{1}{\mathrm{Mod}_{p,\mathfrak{m},c}(K_n)^{1/p}} \ge \lim_{n\to\infty} \min_{\mu\in K_n} \int_X \varphi_\varepsilon \,\mathrm{d}\mu \ge \frac{1}{\mathrm{Mod}_{p,\mathfrak{m},c}(K)^{1/p}} - \varepsilon.$$

The case $\mathrm{Mod}_{p,\mathfrak{m},c}(K) = 0$ is the same, taking $\varphi_M \in \mathrm{C}_b(X)$ such that $\|\varphi_M\|_p = 1$ and $\int_X \varphi_M \,d\mu \ge M$ on K and then letting $M \to \infty$.

(vii) Since $\mathrm{Mod}_{p,\mathfrak{m}}(A) = 0$, by (ii) we find $g \in \mathcal{L}^p_+(X,\mathfrak{m})$ such that $\int_X g \,\mathrm{d}\mu = \infty$ for every $\mu \in A$: this yields $\int_X g \,\mathrm{d}\big(F(\mu)\mu\big) = \infty$ for every $\mu \in A$, showing that $\mathrm{Mod}_{p,\mathfrak{m}}(B) = 0$. □

Remark 2.1.3 In connection with Proposition 2.1.2(iv), in general the constraint $\int_X f \,d\mu \ge 1$ is not saturated by the optimal f, namely the strict inequality can occur for a subset Σ_0 with positive $(p,\mathfrak{m})$-modulus. For instance, if $X = [0,1]$ and $\mathfrak{m}$ is the Lebesgue measure, then

$$\mathrm{Mod}_{p,\mathfrak{m}}\big(\{\mathscr{L}^1 \llcorner [0,\tfrac{1}{2}], \mathscr{L}^1 \llcorner [\tfrac{1}{2},1], \mathscr{L}^1 \llcorner [0,1]\}\big) = 2^p \quad \text{and} \quad f \equiv 2,$$

but $\int_X f \,\mathrm{d}\mathfrak{m} = 2$. However, we will prove using the duality formula $\mathrm{Mod}_{p,\mathfrak{m}} = C^p_{p,\mathfrak{m}}$ that one can always find a subset $\Sigma' \subseteq \Sigma$ (in the example above $\Sigma \setminus \Sigma' = \{\mathscr{L}^1 \llcorner [0,1]\}$) with the same $(p,\mathfrak{m})$-modulus satisfying $\int_X f \,d\mu = 1$ for all $\mu \in \Sigma'$, see the comment made after Corollary 2.3.2.

On the other hand, if the measures in Σ are non-atomic, using just the definition of p-modulus, one can find instead a family Σ' of *smaller* measures with the same modulus as Σ on which the constraint is saturated: suffices to find, for any $\mu \in \Sigma$, a smaller measure μ' (a subcurve, in the case of measures associated to curves) satisfying $\int_X f \,d\mu' = 1$. In the previous example the two constructions lead to the same result, but the two procedures are conceptually quite different.

Another important property is the tightness of $\mathrm{Mod}_{p,\mathfrak{m}}$ in $\mathcal{M}_+(X)$: it will play a crucial role in the proof of Theorem 2.3.1 to prove the inner regularity of $\mathrm{Mod}_{p,\mathfrak{m}}$ for arbitrary Souslin sets.

Lemma 2.1.4 (Tightness of $\mathrm{Mod}_{p,\mathfrak{m}}$**)** *If (X,τ) is Polish and $\mathfrak{m} \in \mathcal{M}_+(X)$, for every $\varepsilon > 0$ there exists $E_\varepsilon \subseteq \mathcal{M}_+(X)$ compact such that $\mathrm{Mod}_{p,\mathfrak{m}}(E^c_\varepsilon) \le \varepsilon$.*

Proof. Since (X,τ) is Polish, by Ulam theorem we can find an nondecreasing family of sets $K_n \in \mathscr{K}(X)$ such that

$$\mathfrak{m}(K^c_n) \to 0.$$

We claim the existence of $\delta_n \downarrow 0$ such that, defining

$$E_k = \{\mu \in \mathcal{M}_+(X) \ : \ \mu(X) \le k \text{ and } \mu(K_n^c) \le \delta_n \ \ \forall n \ge k\},$$

then E_k is compact and $\mathrm{Mod}_{p,\mathfrak{m}}(E_k^c) \to 0$ as k goes to infinity. First of all it is easy to see that the family $\{E_k\}$ is compact by Prokhorov theorem, because it is clearly tight.

To evaluate $\mathrm{Mod}_{p,\mathfrak{m}}(E_k^c)$ we have to build some functions. Let $m_n = \mathfrak{m}(K_n^c)$, assume with no loss of generality that $m_n > 0$ for all n, set $a_n = (\sqrt{m_n} + \sqrt{m_{n+1}})^{-1/p}$ and note that this latter sequence is nondecreasing and diverging to $+\infty$; let us now define the functions

$$f_k(x) := \begin{cases} 0 & \text{if } x \in K_k, \\ a_n & \text{if } x \in K_{n+1} \setminus K_n \text{ and } n \ge k, \\ +\infty & \text{otherwise.} \end{cases}$$

Now we claim that if we put $\delta_n = a_n^{-1}$ in the definition of the E_k's we will have $\mathrm{Mod}_{p,\mathfrak{m}}(E_k^c) \to 0$: in fact, if $\mu \in E_k^c$ then we have either $\mu(X) > k$ or $\mu(K_n^c) > \delta_n$ for some $n \ge k$. In either case the integral of the function $f_k + \frac{1}{k}$ with respect to μ is greater or equal to 1:

- if $\mu(X) > k$ then
$$\int_X \left(f_k + \frac{1}{k}\right) \mathrm{d}\mu \ge \int_X \frac{1}{k} \,\mathrm{d}\mu \ge 1;$$
- if $\mu(K_n^c) > \delta_n$ for some $n \ge k$ we have that
$$\int_X \left(f_k + \frac{1}{k}\right) \mathrm{d}\mu \ge \int_{K_n^c} f_k \,\mathrm{d}\mu \ge \int_{K_n^c} a_n \,\mathrm{d}\mu > \delta_n a_n = 1.$$

So we have that $\mathrm{Mod}_{p,\mathfrak{m}}(E_k^c) \le \|f_k + \frac{1}{k}\|_p^p \le (\|f_k\|_p + \|1/k\|_p)^p$. But

$$\int_X f_k^p \,\mathrm{d}\mathfrak{m} = \sum_{n=k}^{\infty} \int_{K_{n+1}\setminus K_n} a_n^p \,\mathrm{d}\mathfrak{m} = \sum_{n=k}^{\infty} \frac{m_n - m_{n+1}}{\sqrt{m_n} + \sqrt{m_{n+1}}} = \sqrt{m_k},$$

and so we have $\mathrm{Mod}_{p,\mathfrak{m}}(E_k^c) \le \Big((m_k)^{1/(2p)} + (\mathfrak{m}(X))^{1/p}/k\Big)^p \to 0.$ □

2.2 Plans with barycenter in $L^q(X, \mathfrak{m})$ and $(p, \mathfrak{m})$-capacity

In this section (X, τ) is Polish and $\mathfrak{m} \in \mathcal{M}_+(X)$ is a fixed reference measure. We will endow $\mathcal{M}_+(X)$ with the Polish structure making the maps $\mu \mapsto \int_X f \, d\mu$, $f \in \mathrm{C}_b(X)$, continuous, as described in Section 1.1.

Definition 2.2.1 (Plans with barycenter in $L^q(X, \mathfrak{m})$) *Let $q \in (1, \infty]$, $p = q'$. We say that a Borel probability measure $\boldsymbol{\eta}$ on $\mathcal{M}_+(X)$ is a plan with barycenter in $L^q(X, \mathfrak{m})$ if there exists $c \in [0, \infty)$ such that*

$$\iint_X f \,\mathrm{d}\mu \,\mathrm{d}\boldsymbol{\eta}(\mu) \le c \|f\|_p \qquad \forall f \in \mathcal{L}_+^p(X, \mathfrak{m}). \tag{2.2.1}$$

If $\boldsymbol{\eta}$ is a plan with barycenter in $L^q(X, \mathfrak{m})$, we call $c_q(\boldsymbol{\eta})$ the minimal c in (2.2.1).

Notice that $c_q(\boldsymbol{\eta}) = 0$ iff $\boldsymbol{\eta}$ is the Dirac mass at the null measure in $\mathcal{M}_+(X)$. We also used implicitly in (2.2.1) (and in the sequel it will be used without further mention) the fact that $\mu \mapsto \int_X f \,\mathrm{d}\mu$ is Borel whenever $f \in \mathcal{L}^p_+(X,\mathfrak{m})$. The proof can be achieved by a standard monotone class argument.

An equivalent definition of the class plans with barycenter in $L^q(X,\mathfrak{m})$, which explains also the terminology we adopted, is based on the requirement that the barycenter Borel measure

$$\underline{\mu} := \int \mu \,\mathrm{d}\boldsymbol{\eta}(\mu) \tag{2.2.2}$$

is absolutely continuous w.r.t. $\mathfrak{m}$ and with a density ρ in $L^q(X,\mathfrak{m})$. Moreover,

$$c_q(\boldsymbol{\eta}) = \|\rho\|_q. \tag{2.2.3}$$

Indeed, choosing $f = \chi_A$ in (2.2.1) gives $\underline{\mu}(A) \le (\mathfrak{m}(A))^{1/p}$, hence the Radon-Nikodym theorem provides the representation $\underline{\mu} = \rho\mathfrak{m}$ for some $\rho \in L^1(X,\mathfrak{m})$. Then, (2.2.1) once more gives

$$\int_X \rho f \,\mathrm{d}\mathfrak{m} \le c\|f\|_p \qquad \forall f \in L^p(X,\mathfrak{m})$$

and the duality of Lebesgue spaces gives $\rho \in L^q(X,\mathfrak{m})$ and $\|\rho\|_q \le c$. Conversely, if $\underline{\mu}$ has a density in $L^q(X,\mathfrak{m})$, we obtain by Hölder's inequality that (2.2.1) holds with $c = \|\rho\|_q$.

Obviously, (2.2.1) still holds with $c = c_q(\boldsymbol{\eta})$ for all $f \in \mathrm{C}_b(X)$, not necessarily nonnegative, when $\boldsymbol{\eta}$ is a plan with barycenter in $L^q(X,\mathfrak{m})$. Actually the next proposition shows that we need only to check the inequality (2.2.1) for $f \in \mathrm{C}_b(X)$ nonnegative.

Proposition 2.2.2 *Let $\boldsymbol{\eta}$ be a probability measure on $\mathcal{M}_+(X)$ such that*

$$\int\int_X f \,\mathrm{d}\mu \,\mathrm{d}\boldsymbol{\eta}(\mu) \le c\|f\|_p \qquad \text{for all } f \in \mathrm{C}_b(X) \text{ nonnegative} \tag{2.2.4}$$

for some $c \ge 0$. Then (2.2.4) holds, with the same constant c, also for every $f \in \mathcal{L}^p_+(X,\mathfrak{m})$.

Proof. It suffices to remark that (2.2.4) gives

$$\int_X f \,\mathrm{d}\underline{\mu} \le c\|f\|_p \qquad \forall f \in \mathrm{C}_b(X),$$

with $\underline{\mu}$ defined in (2.2.2). Again the duality of Lebesgue spaces provides $\rho \in L^q(X,\mathfrak{m})$ with $\|\rho\|_q \le c$ satisfying $\int_X f\rho \,\mathrm{d}\mathfrak{m} = \int_X f \,\mathrm{d}\underline{\mu}$ for all $f \in \mathrm{C}_b(X)$, hence $\underline{\mu} = \rho\mathfrak{m}$. $\square$

There is a simple duality inequality, involving the minimization in (2.1.2) and a maximization among all $\boldsymbol{\eta}$'s with barycenter in $L^q(X,\mathfrak{m})$. To see it, let's take $f \in \mathcal{L}^p_+(X,\mathfrak{m})$ such that $\int f \,\mathrm{d}\mu \ge 1$ on $\Sigma \subseteq \mathcal{M}_+(X)$. Then, if Σ is universally measurable we may take any plan $\boldsymbol{\eta}$ with barycenter in $L^q(X,\mathfrak{m})$ to obtain

$$\boldsymbol{\eta}(\Sigma) \le \int\int_X f \,\mathrm{d}\mu \,\mathrm{d}\boldsymbol{\eta}(\mu) \le c_q(\boldsymbol{\eta})\|f\|_p. \tag{2.2.5}$$

In particular we have

$$\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma) = 0 \quad \Longrightarrow \quad \boldsymbol{\eta}(\Sigma) = 0 \qquad \text{for all } \boldsymbol{\eta} \text{ with barycenter in } L^q(X,\mathfrak{m}). \tag{2.2.6}$$

In addition, taking in (2.2.5) the infimum over all the $f \in \mathcal{L}^p_+(X,\mathfrak{m})$ such that $\int f \,\mathrm{d}\mu \geq 1$ on Σ and, at the same time, the supremum with respect to all plans $\boldsymbol{\eta}$ with barycenter in $L^q(X,\mathfrak{m})$ and $c_q(\boldsymbol{\eta}) > 0$, we find

$$\sup_{c(\boldsymbol{\eta})>0} \frac{\boldsymbol{\eta}(\Sigma)}{c_q(\boldsymbol{\eta})} \leq \mathrm{Mod}_{p,\mathfrak{m}}(\Sigma)^{1/p}. \tag{2.2.7}$$

The inequality (2.2.7) motivates the next definition.

Definition 2.2.3 ($(p,\mathfrak{m})$-content) *If $\Sigma \subseteq \mathcal{M}_+(X)$ is a universally measurable set we define*

$$C_{p,\mathfrak{m}}(\Sigma) := \sup_{c_q(\boldsymbol{\eta})>0} \frac{\boldsymbol{\eta}(\Sigma)}{c_q(\boldsymbol{\eta})}. \tag{2.2.8}$$

By convention, we set $C_{p,\mathfrak{m}}(\Sigma) = \infty$ if $0 \in \Sigma$.

A first important implication of (2.2.7) is that for any family $\mathcal{F}$ of plans $\boldsymbol{\eta}$ with barycenter in $L^q(X,\mathfrak{m})$

$$C := \sup\{c_q(\boldsymbol{\eta}) \;:\; \boldsymbol{\eta} \in \mathcal{F}\} < \infty \quad \Longrightarrow \quad \mathcal{F} \text{ is tight.} \tag{2.2.9}$$

Indeed, $\boldsymbol{\eta}(E^c_{\varepsilon^p}) \leq \varepsilon c_q(\boldsymbol{\eta}) \leq C\varepsilon$, where the $E_\varepsilon \subseteq \mathcal{M}_+(X)$ are the compact sets provided by Lemma 2.1.4. This allows to prove existence of optimal $\boldsymbol{\eta}$'s in (2.2.8).

Lemma 2.2.4 *Let $\Sigma \subseteq \mathcal{M}_+(X)$ be a universally measurable set such that $C_{p,\mathfrak{m}}(\Sigma) > 0$ and $\sup_\Sigma \mu(X) < \infty$. Then there exists an optimal plan $\boldsymbol{\eta}$ with barycenter in $L^q(X,\mathfrak{m})$ in (2.2.8), and any optimal plan is concentrated on Σ. In particular*

$$C_{p,\mathfrak{m}}(\Sigma) = \frac{\boldsymbol{\eta}(\Sigma)}{c_q(\boldsymbol{\eta})} = \frac{1}{c_q(\boldsymbol{\eta})}.$$

Proof. First we claim that the supremum in (2.2.7) can be restricted to the plans with barycenter in $L^q(X,\mathfrak{m})$ concentrated on Σ. Indeed, given any admissible $\boldsymbol{\eta}$ with $\boldsymbol{\eta}(\Sigma) > 0$, defining $\boldsymbol{\eta}' = (\boldsymbol{\eta}(\Sigma))^{-1}\chi_\Sigma\boldsymbol{\eta}$ we obtain another plan with barycenter in $L^q(X,\mathfrak{m})$ satisfying $\boldsymbol{\eta}'(\Sigma) = 1$ and

$$\int\int_X f \,\mathrm{d}\mu \,\mathrm{d}\boldsymbol{\eta}'(\mu) = \frac{1}{\boldsymbol{\eta}(\Sigma)}\int_\Sigma\int_X f \,\mathrm{d}\mu \,\mathrm{d}\boldsymbol{\eta}(\mu) \leq \frac{1}{\boldsymbol{\eta}(\Sigma)}\int\int_X f \,\mathrm{d}\mu \,\mathrm{d}\boldsymbol{\eta}(\mu) \leq \frac{c_q(\boldsymbol{\eta})}{\boldsymbol{\eta}(\Sigma)}\|f\|_p$$

for all $f \in \mathcal{L}^p_+(X,\mathfrak{m})$. In particular the definition of $c_q(\boldsymbol{\eta}')$ gives

$$c_q(\boldsymbol{\eta}') \leq \frac{c_q(\boldsymbol{\eta})}{\boldsymbol{\eta}(\Sigma)},$$

and proves our claim. The same argument proves that $\boldsymbol{\eta}' = \boldsymbol{\eta}$ whenever $\boldsymbol{\eta}$ is a mazimizer. Now we know that

$$C_{p,\mathfrak{m}}(\Sigma) = \sup_{\boldsymbol{\eta}(\Sigma)=1} \frac{1}{c_q(\boldsymbol{\eta})},$$

where the supremum is made over plans with barycenter in $L^q(X,\mathfrak{m})$. We take a maximizing sequence $(\boldsymbol{\eta}_k)$; for this sequence we have that $c_q(\boldsymbol{\eta}_k) \leq C$, so that $(\boldsymbol{\eta}_k)$ is tight by (2.2.9). Assume with no loss of generality that $\boldsymbol{\eta}_k$ weakly converges to some $\boldsymbol{\eta}$, that is clearly a probability measure in $\mathcal{M}_+(X)$. To see that $\boldsymbol{\eta}$ is a plan with barycenter in $L^q(X,\mathfrak{m})$ and that

$c_q(\boldsymbol{\eta})$ is optimal, we notice that the continuity and boundedness of $\mu \mapsto \int_X f\,\mathrm{d}\mu$ in bounded sets of $\mathcal{M}_+(X)$ for $f \in \mathrm{C}_b(X)$ gives

$$\int \int_X f\,\mathrm{d}\mu\,\mathrm{d}\boldsymbol{\eta}(\mu) = \lim_{k\to\infty} \int \int_X f\,\mathrm{d}\mu\,\mathrm{d}\boldsymbol{\eta}_k(\mu) \le \lim_{k\to\infty} c_q(\boldsymbol{\eta}_k)\|f\|_p,$$

so that

$$\int \int_X f\,\mathrm{d}\mu\,\mathrm{d}\boldsymbol{\eta}(\mu) \le \frac{1}{C_{p,\mathfrak{m}}(\Sigma)}\|f\|_p \qquad \forall f \in \mathrm{C}_b(X).$$

The thesis follows from Proposition 2.2.2. □

2.3 Equivalence between $C_{p,\mathfrak{m}}$ and $\mathrm{Mod}_{p,\mathfrak{m}}$

In the previous two sections, under the standing assumptions (X, τ) Hausdorff topological space (Polish in the case of $C_{p,\mathfrak{m}}$), $\mu \in \mathcal{M}_+(X)$ and $p \in [1, \infty)$, we introduced a p-Modulus $\mathrm{Mod}_{p,\mathfrak{m}}$ and a p-content $C_{p,\mathfrak{m}}$, proving the direct inequalities (see (2.2.7))

$$C^p_{p,\mathfrak{m}} \le \mathrm{Mod}_{p,\mathfrak{m}} \le \mathrm{Mod}_{p,\mathfrak{m},c} \qquad \text{on Souslin subsets of } \mathcal{M}_+(X).$$

Under the same assumptions on (X, τ) and $\mathfrak{m} \in \mathcal{M}_+(X)$, our goal in this section is the following result:

Theorem 2.3.1 *Let (X, τ) be a Polish topological space and $p > 1$. Then $\mathrm{Mod}_{p,\mathfrak{m}}$ is a Choquet capacity in $\mathcal{M}_+(X)$, every Souslin set $\Sigma \subseteq \mathcal{M}_+(X)$ is capacitable and satisfies $\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma)^{1/p} = C_{p,\mathfrak{m}}(\Sigma)$. If moreover Σ is also compact we have $\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma) = \mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma)$.*

Proof. We split the proof in two steps:

- first, prove that $\mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma)^{1/p} \le C_{p,\mathfrak{m}}(\Sigma)$ if Σ is compact, so that in particulat $\mathrm{Mod}^{1/p}_{p,\mathfrak{m}} = C_{p,\mathfrak{m}}$ on compact sets;
- then, prove that $\mathrm{Mod}_{p,\mathfrak{m}}$ and $C_{p,\mathfrak{m}}$ are inner regular, and deduce that $\mathrm{Mod}^{1/p}_{p,\mathfrak{m}} = C_{p,\mathfrak{m}}$ on Souslin sets.

The two steps together yield $\mathrm{Mod}_{p,\mathfrak{m}} = \mathrm{Mod}_{p,\mathfrak{m},c}$ on compact sets, hence we can use Proposition 2.1.2(v,vi) to obtain that $\mathrm{Mod}_{p,\mathfrak{m}}$ is a Choquet capacity in $\mathcal{M}_+(X)$.

Step 1. Assume that $\Sigma \subseteq \mathcal{M}_+(X)$ is compact. In particular $\sup_\Sigma \mu(X)$ is finite and so we have that the linear map $\Phi : \mathrm{C}_b(X) \to \mathrm{C}(\Sigma) = \mathrm{C}_b(\Sigma)$ given by

$$f \mapsto \Phi_f(\mu) := \int_X f\,\mathrm{d}\mu$$

is a bounded linear operator.

If Σ contains the null measure there is nothing to prove, because $\mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma) = \infty$ by definition and $C_{p,\mathfrak{m}}(\Sigma) = \infty$ by convention. If not, by compactness, we obtain that $\inf_\Sigma \mu(X) > 0$, so that taking $f \equiv 1$ in (2.1.3) we obtain $\mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma) < \infty$. We can also assume that $\mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma) > 0$, otherwise there is nothing to prove.

Our first step is the construction of a plan $\boldsymbol{\eta}$ with barycenter in $L^q(X,\mathfrak{m})$ concentrated on Σ. By the equivalent definition analogous to (2.1.3) for $\mathrm{Mod}_{p,\mathfrak{m},c}$, the constant $\xi = \mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma)^{-1/p}$ satisfies

$$\inf_{\mu\in\Sigma} \Phi_f(\mu) \le \xi \|f\|_p \qquad \forall f \in \mathrm{C}(\Sigma). \tag{2.3.1}$$

Denoting by $v = v(\mu)$ the generic element of $\mathrm{C}(\Sigma)$, we will now consider two functions on $\mathrm{C}(\Sigma)$:

$$\begin{aligned} F_1(v) &= \inf \{\|f\|_p \ : \ f \in \mathrm{C}_b(X), \ \ \Phi_f \ge v \text{ on } \Sigma\} \\ F_2(v) &= \min \{v(\mu) \ : \ \mu \in \Sigma\}. \end{aligned}$$

The following properties are immediate to check, using the linearity of $f \mapsto \Phi_f$ for the first one and (2.3.1) for the third one:

- F_1 is convex;
- F_2 is continuous and concave;
- $F_2 \le \xi \cdot F_1$.

With these properties, standard Banach theory gives us a continuous linear functional $L \in (\mathrm{C}(\Sigma))^*$ such that

$$F_2(v) \le L(v) \le \xi \cdot F_1(v) \qquad \forall v \in \mathrm{C}(\Sigma). \tag{2.3.2}$$

For the reader's convenience we detail the argument: first we apply the geometric form of the Hahn-Banach theorem in the space $\mathrm{C}(\Sigma) \times \mathbb{R}$ to the convex sets $A = \{F_2(v) > t\}$ and $B = \{F_1(v) \le t/\xi\}$, where the former is also open, to obtain a continuous linear functional G in $\mathrm{C}(\Sigma) \times \mathbb{R}$ such that

$$G(v,t) < G(w,s) \qquad \text{whenever } F_2(v) > t,\ F_1(w) \le s/\xi.$$

Representing $G(v,t)$ as $H(v) + \beta t$ for some $H \in (\mathrm{C}(\Sigma))^*$ and $\beta \in \mathbb{R}$, the inequality reads

$$H(v) + \beta t < H(w) + \beta s \qquad \text{whenever } F_2(v) > t,\ F_1(w) \le s/\xi.$$

Since F_1 and F_2 are real-valued, $\beta > 0$; we immediately get $F_2 \le (\gamma - H)/\beta \le \xi F_1$, with $\gamma := \sup H(v) + \beta F_2(v)$. On the other hand, $F_1(0) = F_2(0) = 0$ implies $\gamma = 0$, so that we can take $L = -H/\beta$ in (2.3.2).

In particular from (2.3.2) we get that if $v \ge 0$ then $L(v) \ge F_2(v) \ge 0$ and so, since Σ is compact, we can apply Riesz theorem to obtain a nonnegative measure $\boldsymbol{\eta}$ in Σ representing L:

$$L(v) = \int_\Sigma v(\mu)\,\mathrm{d}\boldsymbol{\eta} \qquad \forall v \in \mathrm{C}(\Sigma).$$

Furthermore this measure can't be null since (here $\mathbb{1}$ is the function identically equal to 1).

$$\boldsymbol{\eta}(\Sigma) = L(\mathbb{1}) \ge F_2(\mathbb{1}) = 1,$$

and so $\boldsymbol{\eta}(\Sigma) \ge 1$. Now we claim that $\boldsymbol{\eta}$ is a plan with barycenter in $L^q(X,\mathfrak{m})$; first we prove that $\boldsymbol{\eta}(\Sigma) \le 1$, so that $\boldsymbol{\eta}$ will be a probability measure. In fact, we know $F_2(v)\boldsymbol{\eta}(\Sigma) \le L(v)$ because $v \ge F_2(v)$ on Σ, and then

$$F_2(v)\boldsymbol{\eta}(\Sigma) \le \xi F_1(v).$$

In particular, inserting in this inequality $v = \Phi_\varphi$ with $\varphi \in \mathrm{C}_b(X)$, we obtain

$$\inf_{\Sigma} \Phi_\varphi \le \frac{\xi}{\boldsymbol{\eta}(\Sigma)} \|\varphi\|_p$$

and so $\mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma) \ge (\boldsymbol{\eta}(\Sigma)/\xi)^p = \boldsymbol{\eta}(\Sigma)^p \mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma)$, which implies $\boldsymbol{\eta}(\Sigma) \le 1$. Now we have that

$$\int_\Sigma \left(\int_X f \,\mathrm{d}\mu \right) \mathrm{d}\boldsymbol{\eta} = L(\Phi_f) \le \xi \cdot F_1(\Phi_f) \le \xi \cdot \|f\|_p \qquad \forall f \in \mathrm{C}_b(X) \tag{2.3.3}$$

and so, by Proposition 2.2.2, this inequality is true for every $f \in \mathscr{L}^p_+(X, \mathfrak{m})$, showing that $\boldsymbol{\eta}$ is a plan with barycenter in $L^q(X, \mathfrak{m})$; as a byproduct we gain also that $c_q(\boldsymbol{\eta}) \le \xi$ that gives us, that $C_{p,\mathfrak{m}}(\Sigma) \ge \mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma)^{1/p}$, thus obtaining that

$$C_{p,\mathfrak{m}}(\Sigma) = \mathrm{Mod}_{p,\mathfrak{m}}(\Sigma)^{1/p} = \mathrm{Mod}_{p,\mathfrak{m},c}(\Sigma)^{1/p}.$$

Step 2. Now we will prove that $\mathrm{Mod}_{p,\mathfrak{m}}$ and $C_{p,\mathfrak{m}}$ are both inner regular, namely their value on Souslin sets is the supremum of their value on compact subsets. Inner regularity and equality on compact sets yield $C_{p,\mathfrak{m}}(B) = \mathrm{Mod}_{p,\mathfrak{m}}(B)^{1/p}$ on every Souslin subset B of $\mathcal{M}_+(X)$.

$\mathrm{Mod}_{p,\mathfrak{m}}$ **is inner regular.** Proposition 2.1.2(v,vi) and the fact that $\mathrm{Mod}_{p,\mathfrak{m},c} = \mathrm{Mod}_{p,\mathfrak{m}}$ if the set is compact, give us that $\mathrm{Mod}_{p,\mathfrak{m}}$ is a capacity. For any set $L \subseteq \mathcal{M}_+(X)$ we have $\mathrm{Mod}_{p,\mathfrak{m}}(L) = \sup_\varepsilon \mathrm{Mod}_{p,\mathfrak{m}}(L \cap E_\varepsilon)$, where E_ε are the compact sets given by Lemma 2.1.4. Therefore, suffices to show inner regularity for a Souslin set B contained in E_ε for some ε. Since E_ε is compact, B is a Souslin-compact set and from Choquet Theorem 1.1.5 it follows that for every $\delta > 0$ there is a compact set $K \subseteq B$ such that $\mathrm{Mod}_{p,\mathfrak{m}}(K) \ge \mathrm{Mod}_{p,\mathfrak{m}}(B) - \delta$.

$C_{p,\mathfrak{m}}$ **is inner regular.** Since Souslin sets are universally measurable and $\mathcal{M}_+(X)$ is Polish, we can apply (1.1.1) to any Souslin set B with $\sigma = \boldsymbol{\eta}$ to get

$$\sup_{K \subseteq B} C_{p,\mathfrak{m}}(K) = \sup_{K \subseteq B} \sup_{c_q(\boldsymbol{\eta})>0} \frac{\boldsymbol{\eta}(K)}{c_q(\boldsymbol{\eta})} = \sup_{c_q(\boldsymbol{\eta})>0} \sup_{K \subseteq B} \frac{\boldsymbol{\eta}(K)}{c_q(\boldsymbol{\eta})} = \sup_{c_q(\boldsymbol{\eta})>0} \frac{\boldsymbol{\eta}(B)}{c_q(\boldsymbol{\eta})} = C_{p,\mathfrak{m}}(B).$$

□

The duality formula and the existence of maximizers and minimizers provide the following result.

Corollary 2.3.2 (Necessary and sufficient optimality conditions) *Let* $p > 1$, *let* $\Sigma \subseteq \mathcal{M}_+(X)$ *be a Souslin set such that* $\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma) > 0$ *and* $\sup_\Sigma \mu(X)$ *is finite. Then:*

(a) there exists $f \in \mathscr{L}^p_+(X, \mathfrak{m})$, *unique up to* $\mathfrak{m}$*-negligible sets, such that* $\int_X f \,\mathrm{d}\mu \ge 1$ *for* $\mathrm{Mod}_{p,\mathfrak{m}}$*-a.e.* $\mu \in \Sigma$ *and such that* $\|f\|_p^p = \mathrm{Mod}_{p,\mathfrak{m}}(\Sigma)$;

(b) there exists a plan $\boldsymbol{\eta}$ *with barycenter in* $L^q(X, \mathfrak{m})$ *concentrated on* Σ *such that* $\mathrm{Mod}_{p,\mathfrak{m}}(\Sigma)^{1/p} = 1/c_q(\boldsymbol{\eta})$;

(c) for the function f *in (a) and any* $\boldsymbol{\eta}$ *in (b) there holds*

$$\int_X f \,\mathrm{d}\mu = 1 \ \textit{ for } \boldsymbol{\eta}\textit{-a.e. } \mu \qquad \textit{and} \qquad \int_X \mu \,\mathrm{d}\boldsymbol{\eta}(\mu) = \frac{f^{p-1}}{\|f\|_p^p}\mathfrak{m}. \tag{2.3.4}$$

Finally, if $f \in \mathcal{L}^p_+(X, \mathfrak{m})$ *is optimal in* (2.1.2), *then any plan* $\boldsymbol{\eta}$ *with barycenter in* $L^q(X, \mathfrak{m})$ *concentrated on* Σ *such that* $c_q(\boldsymbol{\eta}) = \|f\|_p^{-1}$ *is optimal in* (2.2.8). *Conversely, if* $\boldsymbol{\eta}$ *is optimal in* (2.2.8), $f \in \mathcal{L}^p_+(X, \mathfrak{m})$ *and* $\int_X f \, \mathrm{d}\mu = 1$ *for* μ*-a.e.* $\boldsymbol{\eta}$ *then* f *is optimal in* (2.1.2).

Proof. The existence of f follows by Proposition 2.1.2(iv). The existence of a maximizer $\boldsymbol{\eta}$ in the duality formula, concentrated on Σ and satisfying $C_{p,\mathfrak{m}}(\Sigma) = 1/c_q(\boldsymbol{\eta})$ follows by Lemma 2.2.4. Since (2.2.6) gives $\int_X f \, \mathrm{d}\mu \geq 1$ for $\boldsymbol{\eta}$-a.e. $\mu \in \Sigma$ we can still derive the inequality (2.2.5) and obtain from Theorem 2.3.1 that all inequalities are equalities. Hence, $\int_X f \, \mathrm{d}\mu = 1$ for $\boldsymbol{\eta}$-a.e. $\mu \in \mathcal{M}_+(X)$. Finally, setting $\underline{\mu} := \int \mu \, \mathrm{d}\boldsymbol{\eta}(\mu)$, from (2.2.3) we get $\underline{\mu} = g\mathfrak{m}$ with $\|g\|_q = c_q(\boldsymbol{\eta})$. This, in combination with

$$\int_X fg \, \mathrm{d}\mathfrak{m} = \int \int_X f \, d\mu \, \mathrm{d}\boldsymbol{\eta}(\mu) = c_q(\boldsymbol{\eta})\|f\|_p = \|g\|_q \|f\|_p,$$

gives $g = f^{p-1}/\|f\|_p^p$.

Finally, the last statements follow directly from (2.2.5) and Theorem 2.3.1. □

In particular, choosing $\boldsymbol{\eta}$ as in (b) and defining

$$\Sigma' := \left\{ \mu \in \mathcal{M}_+(X) : \int_X f \, d\mu = 1 \right\},$$

since $\boldsymbol{\eta}(\Sigma) = \boldsymbol{\eta}(\Sigma')$ we obtain a subfamily with the same p-modulus on which the constraint is saturated.

2.4 Modulus of families of non-parametric curves

In this section we assume that (X, d) is a complete and separable metric space and that $\mathfrak{m} \in \mathcal{M}_+(X)$.

In order to apply the results of the previous sections (with the topology τ induced by d) to families of non-parametric curves we consider the canonical map $\tilde{J} : \mathscr{C}(X) \to \mathcal{M}_+(X) \setminus \{0\}$ of Definition 1.2.5(d). In the sequel, for the sake of simplicity, we will not distinguish between J and $\tilde{J}$, writing $J\underline{\gamma}$ or $J[\gamma] = J\gamma$ (this is not a big abuse of notation, since $\tilde{J}$ is a quotient map).

Now we discuss the notion of $(p, \mathfrak{m})$-modulus, for $p \in [1, \infty)$. The $(p, \mathfrak{m})$-modulus for families $\Gamma \subseteq \mathscr{C}(X)$ of non-parametric curves is given by

$$\mathrm{Mod}_{p,\mathfrak{m}}(\Gamma) := \inf \left\{ \int_X g^p \, \mathrm{d}\mathfrak{m} \; : \; g \in \mathcal{L}^p_+(X, \mathfrak{m}), \; \int_{\underline{\gamma}} g \geq 1 \quad \text{for all } \underline{\gamma} \in \Gamma \right\}. \tag{2.4.1}$$

We adopted the same notation $\mathrm{Mod}_{p,\mathfrak{m}}$ because the identity $\int_{\underline{\gamma}} g = \int_X g \, \mathrm{d}J\underline{\gamma}$ immediately gives

$$\mathrm{Mod}_{p,\mathfrak{m}}(\Gamma) = \mathrm{Mod}_{p,\mathfrak{m}}(J(\Gamma)). \tag{2.4.2}$$

In a similar vein, setting $q = p'$, in the space $\mathscr{C}(X)$ we can define plans with barycenter in $L^q(X, \mathfrak{m})$ as Borel probability measures $\boldsymbol{\pi}$ in $\mathscr{C}(X)$ satisfying

$$\int_{\mathscr{C}(X)} J\underline{\gamma} \, \mathrm{d}\boldsymbol{\pi}(\underline{\gamma}) = g\mathfrak{m} \qquad \text{for some } g \in L^q(X, \mathfrak{m}).$$

Notice that the integral in the left hand side makes sense because the Borel regularity of J easily gives that $\underline{\gamma} \mapsto J\underline{\gamma}(A)$ is Borel in $\mathscr{C}(X)$ for all $A \in \mathscr{B}(X)$. We define, exactly as in (2.2.3), $c_q(\boldsymbol{\pi})$ to be the $L^q(X, \mathfrak{m})$ norm of the barycenter g. Then, the same argument leading to (2.2.5) gives

$$\frac{\boldsymbol{\pi}(\Gamma)}{c_q(\boldsymbol{\pi})} \leq \mathrm{Mod}_{p,\mathfrak{m}}(\Gamma)^{1/p} \qquad \text{for all } \boldsymbol{\pi} \in \mathcal{P}(\mathscr{C}(X)) \text{ with barycenter in } L^q(X, \mathfrak{m}) \tag{2.4.3}$$

for every universally measurable set Γ in $\mathscr{C}(X)$.

Remark 2.4.1 (Democratic plans) In more explicit terms, Borel probability measures $\boldsymbol{\pi}$ in $\mathscr{C}(X)$ with barycenter in $L^q(X, \mathfrak{m})$ satisfy

$$\int_0^1 (\mathrm{e}_t)_\sharp(|\dot{\gamma}_t|\boldsymbol{\pi})\, \mathrm{d}t = g\mathfrak{m} \quad \text{for some } g \in L^q(X, \mathfrak{m}) \tag{2.4.4}$$

when we view them as measures on nonconstant curves $\gamma \in \mathrm{AC}([0, 1]; X)$. For instance, in the particular case when $\boldsymbol{\pi}$ is concentrated on family of geodesics parameterized with constant speed and with length uniformly bounded from below, the case $q = \infty$ corresponds to the class of democratic plans considered in [63].

Defining $C_{p,\mathfrak{m}}(\Gamma)$ as the supremum in the right hand side of (2.4.3), we can now use Theorem 2.3.1 to show that even in this case there is no duality gap.

Theorem 2.4.2 *For every $p > 1$ and every Souslin set $\Gamma \subseteq \mathscr{C}(X)$ with $\mathrm{Mod}_{p,\mathfrak{m}}(\Gamma) > 0$ there exists a $\boldsymbol{\pi} \in \mathcal{P}\big(\mathscr{C}(X)\big)$ with barycenter in $L^q(X, \mathfrak{m})$, concentrated on Γ and satisfying $c_q(\boldsymbol{\pi}) = \mathrm{Mod}_{p,\mathfrak{m}}(\Gamma)^{-1/p}$.*

Proof. From Theorem 2.3.1 we deduce the existence of $\boldsymbol{\eta} \in \mathcal{P}\big(\mathcal{M}_+(X)\big)$ with barycenter in $L^q(X, \mathfrak{m})$ concentrated on the Souslin set $J(\Gamma)$ and satisfying

$$\frac{1}{c_q(\boldsymbol{\eta})} = \mathrm{Mod}_{p,\mathfrak{m}}(J(\Gamma))^{1/p} = \mathrm{Mod}_{p,\mathfrak{m}}(\Gamma)^{1/p}.$$

By a measurable selection theorem [20, Theorem 6.9.1] we can find a $\boldsymbol{\eta}$-measurable map $f : J(\Gamma) \to \mathscr{C}(X)$ such that $f(\mu) \in \Gamma \cap J^{-1}(\mu)$ for all $\mu \in J(\Gamma)$. The measure $\boldsymbol{\pi} := f_\sharp \boldsymbol{\eta}$ is concentrated on Γ and the equality between the barycenters

$$\int_{\mathscr{C}(X)} J\underline{\gamma}\, \mathrm{d}\boldsymbol{\pi}(\underline{\gamma}) = \int \mu\, \mathrm{d}\boldsymbol{\eta}(\mu)$$

gives $c_q(\boldsymbol{\pi}) = c_q(\boldsymbol{\eta})$. □

2.5 Modulus of families of parametric curves

In this section we still assume that (X, d) is a complete and separable metric space and that $\mathfrak{m} \in \mathcal{M}_+(X)$. We consider a notion of p-modulus for *parametric* curves, enforcing the condition (2.4.4) (at least when Lipschitz curves are considered), and we compare with the non-parametric counterpart. To this aim, we introduce the continuous map

$$M : \mathrm{C}([0, 1]; X) \to \mathcal{P}(X), \qquad M(\gamma) := \gamma_\sharp\big(\mathscr{L}^1 \llcorner [0, 1]\big). \tag{2.5.1}$$

Indeed, replacing $J\gamma = \gamma_\sharp(|\dot\gamma|\mathscr{L}^1 \llcorner [0,1])$ with M we can consider a "parametric" modulus of a family of curves $\Sigma \subseteq \mathrm{C}([0,1];X)$ just by evaluating $\mathrm{Mod}_{p,\mathfrak{m}}(M(\Sigma))$. By Proposition 2.1.2(vii), if $\Sigma \subseteq \mathrm{AC}^\infty_c([0,1];X)$ then

$$\mathrm{Mod}_{p,\mathfrak{m}}(M(\Sigma)) = 0 \quad \Longleftrightarrow \quad \mathrm{Mod}_{p,\mathfrak{m}}(J(\Sigma)) = 0. \tag{2.5.2}$$

On the other hand, things are more subtle when the speed is not constant.

Definition 2.5.1 (q-energy and parametric barycenter) *Let $\boldsymbol{\rho} \in \mathcal{P}\big(\mathrm{C}([0,1];X)\big)$ and $q \in [1,\infty)$. We say that $\boldsymbol{\rho}$ has finite q-energy if $\boldsymbol{\rho}$ is concentrated on $\mathrm{AC}^q([0,1];X)$ and*

$$\int\int_0^1 |\dot\gamma_t|^q \,\mathrm{d}t\,\mathrm{d}\boldsymbol{\rho}(\gamma) < \infty. \tag{2.5.3}$$

We say that $\boldsymbol{\rho}$ has parametric barycenter $h \in L^q(X,\mathfrak{m})$ if

$$\int\int_0^1 f(\gamma_t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\rho}(\gamma) = \int_X f\,h\,\mathrm{d}\mathfrak{m} \quad \forall f \in \mathrm{C}_b(X). \tag{2.5.4}$$

The finiteness condition (2.5.3) and the concentration on $\mathrm{AC}^q([0,1];X)$ can be also be written, recalling the definition (1.2.1) of $\mathcal{E}_q$, as follows:

$$\int \mathcal{E}_q(\gamma)\,\mathrm{d}\boldsymbol{\rho}(\gamma) < \infty.$$

Notice also that the definition (2.5.1) of M gives that (2.5.4) is equivalent to require the existence of a constant $C \geq 0$ such that

$$\int\int_X f\,\mathrm{d}M\gamma\,\mathrm{d}\boldsymbol{\rho}(\gamma) \leq C\Big(\int_X f^p\,\mathrm{d}\mathfrak{m}\Big)^{1/p} \quad \forall f \in \mathrm{C}_b(X),\ f \geq 0. \tag{2.5.5}$$

In this case the best constant C in (2.5.5) corresponds to $\|h\|_{L^q(X,\mathfrak{m})}$ for h as in (2.5.4).

Remark 2.5.2 It is not difficult to check that a Borel probability measure $\boldsymbol{\rho}$ concentrated on a set $\Gamma \subseteq \mathrm{AC}^\infty([0,1];X)$ with $\boldsymbol{\rho}$-essentially bounded Lipschitz constants and parametric barycenter in $L^q(X,\mathfrak{m})$ has also (nonparametric) barycenter in $L^q(X,\mathfrak{m})$. Conversely, if $\boldsymbol{\pi} \in \mathcal{P}\big(\mathscr{C}(X)\big)$ with barycenter in $L^q(X,\mathfrak{m})$ and $\boldsymbol{\pi}$-essentially bounded length $\ell(\gamma)$, then $\mathsf{j}_\sharp\boldsymbol{\pi}$ has parametric barycenter in $L^q(X,\mathfrak{m})$.

Now, arguing as in the proof of Theorem 2.4.2 (which provided existence of plans $\boldsymbol{\pi}$ in $\mathscr{C}(X)$) we can use a measurable selection theorem to deduce from our basic duality Theorem 2.3.1 the following result.

Theorem 2.5.3 *For every $p > 1$ and every Souslin set $\Sigma \subseteq \mathrm{C}([0,1];X)$, $\mathrm{Mod}_{p,\mathfrak{m}}(M(\Sigma)) > 0$ is equivalent to the existence of $\boldsymbol{\rho} \in \mathcal{P}\big(\mathrm{C}([0,1];X)\big)$ concentrated on Σ with parametric barycenter in $L^q(X,\mathfrak{m})$.*

Our next goal is to use reparameterizations to improve the parametric barycenter from $L^q(X,\mathfrak{m})$ to $L^\infty(X,\mathfrak{m})$. To this aim, we begin by proving the Borel regularity of some parametrization maps. Let $h : X \to (0,\infty)$ be a Borel map with $\sup_X h < \infty$ and for every $\sigma \in \mathrm{C}([0,1];X)$ let us set

$$G(\sigma) := \int_0^1 h(\sigma_r)\,\mathrm{d}r, \qquad \mathsf{t}_\sigma(s) := \frac{1}{G(\sigma)}\int_0^s h(\sigma_r)\,\mathrm{d}r : [0,1] \to [0,1]. \tag{2.5.6}$$

Since t_σ is Lipschitz and $\mathsf{t}'_\sigma > 0$ $\mathscr{L}^1$-a.e. in $(0,1)$, its inverse $\mathsf{s}_\sigma : [0,1] \to [0,1]$ is absolutely continuous and we can define

$$H : \mathrm{AC}([0,1];X) \to \mathrm{AC}([0,1];X), \qquad H\sigma(t) := \sigma(\mathsf{s}_\sigma(t)). \tag{2.5.7}$$

Notice that $H\big(\mathrm{AC}^\infty_c([0,1];X)\big) \subseteq \mathrm{AC}_0([0,1];X)$.

Lemma 2.5.4 *If $h : X \to \mathbb{R}$ is a bounded Borel function, the map G in* (2.5.6) *is Borel. If we assume, in addition, that $h > 0$ in X, then also t_σ in* (2.5.6) *is Borel and the map H in* (2.5.7) *is Borel and injective.*

Proof. Let us prove first that the map

$$\sigma \mapsto \tilde{\mathsf{t}}_\sigma(t) = \int_0^t h(\sigma_r)\,\mathrm{d}r$$

is Borel from $\mathrm{C}([0,1];X)$ to $\mathrm{C}([0,1])$ for any bounded Borel function $h : X \to \mathbb{R}$. This follows by a monotone class argument (see for instance [20, Theorem 2.12.9(iii)]), since class of functions h for which the statement is true is a vector space containing bounded continuous functions and stable under equibounded pointwise limits. By the continuity of the integral operator, the map G is Borel as well.

Now we turn to H, assuming that $h > 0$. By Proposition 1.1.3(iii) it will be sufficient to show that the inverse of H, namely the map $\sigma \mapsto \sigma \circ \mathsf{t}_\sigma$, is Borel. Since the map $(\sigma, \mathsf{t}) \mapsto \sigma \circ \mathsf{t}$ is continuous from $\mathrm{C}([0,1];X) \times \mathrm{C}([0,1])$ to $\mathrm{C}([0,1];X)$, the Borel regularity of the inverse of H follows by the Borel regularity of $\sigma \mapsto \mathsf{t}_\sigma$. □

Theorem 2.5.5 *Let $q \in (1,\infty)$ and $p = q'$. If $\boldsymbol{\rho} \in \mathcal{P}\big(\mathrm{C}([0,1];X)\big)$ has finite q-energy and parametric barycenter $h \in L^\infty(X,\mathfrak{m})$, then $\boldsymbol{\pi} = \mathsf{i}_\sharp\boldsymbol{\rho}$ has barycenter in $L^q(X,\mathfrak{m})$ and*

$$c_q(\boldsymbol{\pi}) \le \Big(\int \mathcal{E}_q(\gamma)\,\mathrm{d}\boldsymbol{\rho}(\gamma)\Big)^{1/q} \|h\|^{1/p}_{L^\infty(X,\mathfrak{m})}. \tag{2.5.8}$$

Conversely, if $\boldsymbol{\pi} \in \mathcal{P}\big(\mathscr{C}(X)\big)$ has barycenter in $L^q(X,\mathfrak{m})$ and $\boldsymbol{\pi}$-essentially bounded length $\ell(\gamma)$, concentrated on a Souslin set $\Gamma \subseteq \mathscr{C}(X)$, there exists $\boldsymbol{\rho} \in \mathcal{P}\big(\mathrm{C}([0,1];X)\big)$ with finite q-energy and parametric barycenter in $L^\infty(X,\mathfrak{m})$ concentrated in a Souslin set contained in $[\mathsf{j}(\Gamma)]$.

More generally, let $\boldsymbol{\sigma} \in \mathcal{P}\big(\mathrm{C}([0,1];X)\big)$ be concentrated on a Souslin set $\Gamma \subseteq \mathrm{AC}^\infty([0,1];X)$, with parametric barycenter in $L^q(X,\mathfrak{m})$ and with $\boldsymbol{\sigma}$-essentially bounded Lipschitz constants. Then there exists $\boldsymbol{\rho} \in \mathcal{P}\big(\mathrm{C}([0,1];X)\big)$ with finite q-energy and parametric barycenter in $L^\infty(X,\mathfrak{m})$ concentrated on a Souslin set contained in $[\Gamma]$.

Proof. Notice that for every nonnegative Borel f there holds

$$\iint_{\underline{\gamma}} f\,\mathrm{d}\boldsymbol{\pi}(\underline{\gamma}) = \iint_0^1 f(\gamma_t)\,|\dot\gamma_t|\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\rho}(\gamma) \le \Big(\int \mathcal{E}_q\,\mathrm{d}\boldsymbol{\rho}\Big)^{1/q}\Big(\iint_0^1 f^p(\gamma_t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\rho}(\gamma)\Big)^{1/p}$$
$$\le \Big(\int \mathcal{E}_q\,\mathrm{d}\boldsymbol{\rho}\Big)^{1/q}\Big(\int_X f^p\,h\,\mathrm{d}\mathfrak{m}\Big)^{1/p} \le \Big(\int \mathcal{E}_q\,\mathrm{d}\boldsymbol{\rho}\Big)^{1/q}\|h\|^{1/p}_{L^\infty(X,\mathfrak{m})}\|f\|_{L^p(X,\mathfrak{m})},$$

so that (2.5.8) holds.

Let us now prove the last statement from $\boldsymbol{\sigma}$ to $\boldsymbol{\rho}$, since the "converse" statement from $\boldsymbol{\pi}$ to $\boldsymbol{\rho}$ simply follows by applying the last statement to $\boldsymbol{\sigma} := \mathsf{j}_\sharp\boldsymbol{\pi}$ and recalling Remark 2.5.2. Let $g \in L^q(X,\mathfrak{m})$ be the parametric barycenter of $\boldsymbol{\sigma}$ and let us set $h := 1/(\varepsilon \vee g)$, with $\varepsilon > 0$ fixed. Up to a modification of g in a $\mathfrak{m}$-negligible set, it is not restrictive to assume that h is Borel and with values in $(0, 1/\varepsilon]$, so that the corresponding maps G and H defined as in (2.5.6) and (2.5.7) are Borel.

We set $\hat{\boldsymbol{\rho}} := z^{-1}G(\cdot)\boldsymbol{\sigma}$, where $z \in (0, 1/\varepsilon]$ is the normalization constant $\int G(\gamma)\,\mathrm{d}\boldsymbol{\sigma}(\gamma)$. Let us consider the inverse $\mathsf{s}_\sigma : [0,1] \to [0,1]$ of the map t_σ in (2.5.6), which is absolutely continuous for every σ and the corresponding transformation $H\sigma$ in (2.5.7). We denote by L the $\boldsymbol{\sigma}$-essential supremum of the Lipschitz constants of the curves in Γ. Notice that for $\boldsymbol{\sigma}$-a.e. σ

$$|(H\sigma)'|(t) \le L\mathsf{s}_\sigma'(t) = \frac{L\,G(\sigma)}{h(H\sigma(t))} \quad \mathscr{L}^1\text{-a.e. in } (0,1), \tag{2.5.9}$$

and for every nonnegative Borel function f there holds

$$\int_0^1 f(H\sigma(t))\,\mathrm{d}t = \int_0^1 f(\sigma(\mathsf{s}_\sigma(t)))\,\mathrm{d}t = \int_0^1 f(\sigma(s))\mathsf{t}_\sigma'(s)\,\mathrm{d}s = \frac{1}{G(\sigma)}\int_0^1 f(\sigma(s))h(\sigma(s))\,\mathrm{d}s,$$

so that choosing $f = h^{-q}$ yields

$$\mathcal{E}_q(H\sigma) \le L^q G^q(\sigma)\int_0^1 h^{-q}(H\sigma(t))\,\mathrm{d}t \le \frac{L^q}{\varepsilon^{q-1}}\int_0^1 h^{1-q}(\sigma(s))\,\mathrm{d}s. \tag{2.5.10}$$

Now we set $\boldsymbol{\rho} := H_\sharp\hat{\boldsymbol{\rho}}$ and notice that, by construction, $\boldsymbol{\rho}$ is concentrated on the Souslin set $H(\Gamma) \subseteq [\Gamma]$. Integrating the q-energy with respect to $\boldsymbol{\rho}$ we obtain

$$\begin{aligned}\int \mathcal{E}_q(\theta)\,\mathrm{d}\boldsymbol{\rho}(\theta) &= \int \mathcal{E}_q(H\sigma)\,\mathrm{d}\hat{\boldsymbol{\rho}}(\sigma) \le \frac{L^q}{z\varepsilon^{q-1}}\int G(\sigma)\int_0^1 h^{1-q}(\sigma(s))\,\mathrm{d}s\,\mathrm{d}\boldsymbol{\pi}(\sigma)\\ &\le \frac{L^q}{z\varepsilon^q}\int_X h^{1-q}g\,\mathrm{d}\mathfrak{m} \le \frac{L^q}{z\varepsilon^q}\left(\varepsilon^{q-1}\int_X g\,\mathrm{d}\mathfrak{m} + \int_X g^q\,\mathrm{d}\mathfrak{m}\right),\end{aligned}$$

thus obtaining that $\boldsymbol{\rho}$ has finite q-energy. Similarly

$$\begin{aligned}\int\int_0^1 f(\theta(t))\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\rho}(\theta) &= \int\int_0^1 f(H\sigma(t))\,\mathrm{d}t\,d\hat{\boldsymbol{\rho}}(\sigma)\\ &= \frac{1}{z}\int G(\sigma)\int_0^1 f(\sigma(s))h(\sigma(s))\,\mathrm{d}s\,\mathrm{d}\boldsymbol{\sigma}(\sigma) \le \frac{1}{\varepsilon z}\int_X fgh\,\mathrm{d}\mathfrak{m}.\end{aligned}$$

Since $gh \le 1$, this shows that $\boldsymbol{\rho}$ has parametric barycenter in $L^\infty(X,\mathfrak{m})$. □

Corollary 2.5.6 *A Souslin set $\Gamma \subseteq \mathscr{C}(X)$ is* $\mathrm{Mod}_{p,\mathfrak{m}}$*-negligible if and only if $\boldsymbol{\rho}_*([\mathsf{j}\Gamma]) = 0$ for every $\boldsymbol{\rho} \in \mathcal{P}\big(\mathrm{C}([0,1];X)\big)$ concentrated on $\mathrm{AC}^q([0,1];X)$ and with parametric barycenter in $L^\infty(X,\mathfrak{m})$.*

Proof. Let us first suppose that Γ is $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible and let us denote by $h \in L^\infty(X,\mathfrak{m})$ the parametric barycenter of $\boldsymbol{\rho}$ and let us prove that $\boldsymbol{\rho}_*([\mathsf{j}\Gamma]) = 0$. Since $\boldsymbol{\rho}$ is concentrated on $\mathrm{AC}^q([0,1];X)$ we can assume with no loss of generality (possibly restricting $\boldsymbol{\rho}$ to the class of curves σ with $\mathcal{E}_q(\sigma) \le n$ and normalizing) that $\boldsymbol{\rho}$ has finite q-energy. We observe that if $\sigma \in \mathrm{AC}([0,1];X)$ and $f : X \to [0,\infty]$ is Borel, there holds

$$\int_0^1 f(\sigma(t))|\dot\sigma(t)|\,\mathrm{d}t \le \Big(\int_0^1 f^p(\sigma(t))\,\mathrm{d}t\Big)^{1/p}\Big(\mathcal{E}_q(\sigma)\Big)^{1/q}. \tag{2.5.11}$$

If f satisfies

$$\int_{\underline{\gamma}} f \ge 1 \qquad \forall \underline{\gamma} \in \Gamma$$

we obtain that $\int_\sigma f \ge 1$ for all $\sigma \in [\mathrm{j}\Gamma]$. We can now integrate w.r.t. $\boldsymbol{\rho}$ and use (2.5.11) to get

$$\begin{aligned}\boldsymbol{\rho}_*([\mathrm{j}\Gamma]) &\le \Big(\int\int_0^1 f^p(\sigma(t))\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\rho}(\sigma)\Big)^{1/p}\Big(\int \mathcal{E}_q(\sigma)\,\mathrm{d}\boldsymbol{\rho}(\sigma)\Big)^{1/q}\\ &= \Big(\int_X f^p\,h\,\mathrm{d}\mathfrak{m}\Big)^{1/p}\Big(\int \mathcal{E}_q(\sigma)\,\mathrm{d}\boldsymbol{\rho}(\sigma)\Big)^{1/q} \le \|f\|_p\|h\|_\infty^{1/p}\Big(\int \mathcal{E}_q(\sigma)\,\mathrm{d}\boldsymbol{\rho}(\sigma)\Big)^{1/q}. \end{aligned} \tag{2.5.12}$$

By minimizing with respect to f we obtain that $\boldsymbol{\rho}_*([\mathrm{j}\Gamma]) = 0$.

Conversely, suppose that $\mathrm{Mod}_{p,\mathfrak{m}}(\Gamma) > 0$; possibly passing to a smaller set, by the countable subadditivity of $\mathrm{Mod}_{p,\mathfrak{m}}$ we can assume that ℓ is bounded on Γ: then by Theorem 2.4.2 there exists $\boldsymbol{\pi} \in \mathcal{P}\big(\mathscr{C}(X)\big)$ with barycenter in $L^q(X,\mathfrak{m})$ concentrated on Γ and therefore the boundedness of ℓ allows to apply the final statement of Theorem 2.5.5 to obtain $\boldsymbol{\rho} \in \mathcal{P}\big(\mathrm{C}([0,1];X)\big)$ with finite q-energy, parametric barycenter in $L^\infty(X,\mathfrak{m})$ and concentrated on a Souslin subset of $[\mathrm{j}\Gamma]$. $\square$

In the next corollary, in order to avoid further measurability issues, we state our result with the inner measure

$$\mu_*(E) := \sup\{\mu(B) \,:\, B \text{ Borel}, B \subseteq E\}.$$

This formulation is sufficient for our purposes.

Corollary 2.5.7 *Let* $\Gamma \subseteq \mathrm{AC}^\infty([0,1];X)$ *be a Souslin set such that* $\boldsymbol{\rho}_*\big([\Gamma]\big) = 0$ *for every plan* $\boldsymbol{\rho} \in \mathcal{P}(\mathrm{C}([0,1];X))$ *concentrated on* $\mathrm{AC}^q([0,1];X)$ *and with parametric barycenter in* $L^\infty(X,\mathfrak{m})$*. Then* $M(\Gamma)$ *is* $\mathrm{Mod}_{p,\mathfrak{m}}$*-negligible.*

Proof. Suppose by contradiction that $\mathrm{Mod}_{p,\mathfrak{m}}(M(\Gamma)) > 0$; possibly passing to a smaller set, by the countable subadditivity of $\mathrm{Mod}_{p,\mathfrak{m}}$ we can assume that Lip is bounded on Γ. By Theorem 2.5.3 there exists $\boldsymbol{\pi} \in \mathcal{P}\big(\mathrm{C}([0,1];X)\big)$ with parametric barycenter in $L^q(X,\mathfrak{m})$ concentrated on Γ. The boundedness of Lip on Γ allows to appy the second part of Theorem 2.5.5 to obtain $\boldsymbol{\rho} \in \mathcal{P}\big(\mathrm{C}([0,1];X)\big)$ with parametric barycenter in $L^\infty(X,\mathfrak{m})$, finite q-energy and concentrated on a Souslin subset of $[\Gamma]$. $\square$

2.6 Test plans and their null sets

In this section we will assume that (X,d) is a complete and separable metric space and $\mathfrak{m} \in \mathcal{M}_+(X)$. The following notions have already been used in [11] ($q = 2$) and [9] (in connection with the Sobolev spaces with gradient in $L^p(X,\mathfrak{m})$, with $q = p'$; see also [5] in connection with the BV theory), with a slight difference: in [9], [11] the authors use only q-test plans that satisfy the additional condition $\int \mathcal{E}_q \, \mathrm{d}\boldsymbol{\rho} < \infty$. Here we drop this assumption, requiring only that $\boldsymbol{\rho}$ is concentrated on $\mathrm{AC}^q([0,1];X) = \{\mathcal{E}_q < \infty\}$. However it is obvious that the negligible sets described by the two approaches are the same, since every q-plan $\boldsymbol{\rho}$ without the integrability condition can be approximated by q-plans σ satisfying even (2.6.1) below.

Definition 2.6.1 (q-test plans and negligible sets) *Let $\boldsymbol{\rho} \in \mathcal{P}(\mathrm{C}([0,1];X))$ and $q \in [1,\infty]$. We say that $\boldsymbol{\rho}$ is a q-test plan if*

(i) $\boldsymbol{\rho}$ is concentrated on $\mathrm{AC}^q([0,1];X)$;

(ii) there exists a constant $C = C(\boldsymbol{\rho}) > 0$ satisfying $(\mathrm{e}_t)_\sharp\boldsymbol{\rho} \le C\mathfrak{m}$ for all $t \in [0,1]$.

We say that a universally measurable set $\Gamma \subseteq \mathrm{C}([0,1];X)$ is q-negligible if $\boldsymbol{\rho}(\Gamma) = 0$ for all q-test plans $\boldsymbol{\rho}$.

Notice that, by definition, $\mathrm{C}([0,1];X)\setminus\mathrm{AC}^q([0,1];X)$ is q-negligible. The lack of invariance of these concepts, even under bi-Lipschitz reparameterizations (dependent on the curve) is due to condition (ii), which is imposed at any given time and with no averaging (and no dependence on speed as well). Since condition (ii) is more restrictive compared for instance to the notion of democratic test plan of [63] (see Remark 2.4.1), this means that sets of curves have higher chances of being negligible w.r.t. this notion, as the next elementary example shows.

We now want to relate null sets according to Definition 2.6.1 to null sets in the sense of p-modulus. Notice first that in the definition of q-negligible set we might consider only plans $\boldsymbol{\rho}$ satisfying the stronger condition

$$\operatorname{esssup}\{\mathcal{E}_q(\sigma)\} < \infty \tag{2.6.1}$$

because any q-test plan can be monotonically be approximated by q-test plans satisfying this condition. Arguing as in the proof of (2.5.12) we easily see that

$$\Gamma \subseteq \mathscr{C}(X)\ \mathrm{Mod}_{p,\mathfrak{m}}\text{-negligible} \qquad \Longrightarrow \qquad \mathsf{i}^{-1}(\Gamma)\ q\text{-negligible}. \tag{2.6.2}$$

The following simple example shows that the implication can't be reversed, namely sets whose images under i^{-1} are q-negligible need not be $\mathrm{Mod}_{p,\mathfrak{m}}$-null.

Example 2.6.2 *Let $X = \mathbb{R}^2$, d the Euclidean distance, $\mathfrak{m} = \mathscr{L}^2$. The family of parametric segments*

$$\Sigma = \{\gamma^x :\ x \in [0,1]\} \subseteq \mathrm{AC}([0,1];\mathbb{R}^2)$$

with $\gamma^x_t = (x,t)$ is q-negligible for any q, but $\mathsf{i}(\Sigma)$ has p-modulus equal to 1.

In the previous example the implication fails because the trajectories γ^x fall, at any given time t, into a $\mathfrak{m}$-negligible set, and actually the same would be true if this concentration property holds at some fixed time. It is tempting to imagine that the implication is restored if we add to the initial family of curves all their reparameterizations (an operation that leaves the p-modulus invariant). However, since any reparameterization fixes the endpoints, even this fails. However, in the following, we will see that the implication

$$\Gamma\ q\text{-negligible} \qquad \Longrightarrow \qquad \mathrm{Mod}_{p,\mathfrak{m}}(\mathsf{i}(\Gamma)) = 0$$

could be restored if we add some structural assumptions on Γ (in particular a "stability" condition); the collections of curves we are mainly interested in are those connected with the theory of Sobolev spaces in [11], [9], and we will find a new proof of the fact that if we define weak upper gradients according to the two notions, the Sobolev spaces are eventually the same.

We now fix some additional notation: for $I=[a,b]\subseteq[0,1]$ we define the "stretching" map $\mathsf{s}_I:\mathrm{AC}([0,1];X)\to\mathrm{AC}([0,1];X)$, mapping γ to $\gamma\circ s_I$, where $s_I:[0,1]\to[a,b]$ is the affine map with $s_I(0)=a$ and $s_I(1)=b$. Notice that this map acts also in all the other spaces AC^q, AC_0, AC^∞_c of parametric curves we are considering. Recall also the definition of k given in Proposition 1.2.1

Definition 2.6.3 (Stable and invariant sets of curves)

(i) *We say that $\Gamma\subseteq\{\gamma\in\mathrm{AC}([0,1];X):\ \ell(\gamma)>0\}$ is invariant under constant speed reparameterization if $\mathsf{k}\gamma\in\Gamma$ for all $\gamma\in\Gamma$;*

(ii) *We say that $\Gamma\subseteq\mathrm{AC}([0,1];X)$ is $\sim$-invariant if $[\gamma]\subseteq\Gamma$ for all $\gamma\in\Gamma$;*

(iii) *We say that $\Gamma\subseteq\mathrm{AC}([0,1];X)$ is stable if for every $\gamma\in\Gamma$ there exists $\varepsilon\in(0,1/2)$ such that $s_I\gamma\in\Gamma$ whenever $I=[a,b]\subseteq[0,1]$ and $|a|+|1-b|\le\varepsilon$.*

The following theorem provides key connections between q-negligibility and $\mathrm{Mod}_{p,\mathfrak{m}}$-negligibility, both in the nonparametric sense (statement (i)) and in the parametric case (statement (ii)), for stable sets of curves.

Theorem 2.6.4 *Let $\Gamma\subseteq\mathrm{AC}([0,1];X)$ be a Souslin and stable set of curves.*

(i) *If, in addition, $\ell(\gamma)>0$ for all $\gamma\in\Gamma$ and Γ is both $\sim$-invariant and invariant under constant speed reparameterization, then Γ is q-negligible if and only if $J(\Gamma)$ is $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible in $\mathcal{M}_+(X)$ (equivalently, $\mathsf{i}(\Gamma)$ is $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible in $\mathscr{C}(X)$).*

(ii) *If Γ is q-negligible and $[\Gamma\cap\mathrm{AC}^\infty([0,1];X)]\subseteq\Gamma$, then $M\big(\Gamma\cap\mathrm{AC}^\infty([0,1];X)\big)$ is $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible in $\mathcal{M}_+(X)$. If Γ is also $\sim$-invariant then the converse holds, too.*

Proof. (i) The proof of the nontrivial implication, from positivity of $\mathrm{Mod}_{p,\mathfrak{m}}(J(\Gamma))$ to Γ being not q-negligible is completely analogous to the proof of (ii), given below, by applying Corollary 2.5.6 to $\mathsf{i}(\Gamma)$ in place of Corollary 2.5.7 to $\Gamma\cap\mathrm{AC}^\infty([0,1];X)$ and the same rescaling technique. Since we will only need (ii) in the sequel, we only give a detailed proof of (ii).

(ii) Let us prove that the positivity of $\mathrm{Mod}_{p,\mathfrak{m}}\big(M(\Gamma\cap\mathrm{AC}^\infty([0,1];X))\big)$ implies that Γ is not q-negligible. Since $\Gamma\cap\mathrm{AC}^\infty([0,1];X)$ is stable, we can assume the existence of $\varepsilon\in(0,1/2)$ such that $s_I\gamma\in\Gamma$ whenever $I=[a,b]\subseteq[0,1]$ and $|a|+|1-b|\le\varepsilon$.

By applying Corollary 2.5.7 to $\Gamma\cap\mathrm{AC}^\infty([0,1];X)$ we obtain the existence of $\boldsymbol{\rho}\in\mathcal{P}\big(\mathrm{AC}^q([0,1];X)\big)$ concentrated on a Souslin subset of $[\Gamma\cap\mathrm{AC}^\infty([0,1];X)]$, and then on Γ, with L^∞ parametric barycenter, i.e. such that

$$\int_0^1(\mathrm{e}_t)_\sharp\boldsymbol{\rho}\,\mathrm{d}t\le C\mathfrak{m}\qquad\text{for some } C>0. \tag{2.6.3}$$

Let's define a family of reparametrization maps $F^\tau_\varepsilon:\mathrm{AC}^q([0,1];X)\to\mathrm{AC}^q([0,1];X)$:

$$F^\tau_\varepsilon\gamma(t)=\gamma\Big(\frac{t+\tau}{1+\varepsilon}\Big)\qquad t\in[0,1],\qquad\forall\gamma\in\mathrm{AC}^q([0,1];X),\ \ \forall\,\tau\in[0,\varepsilon]. \tag{2.6.4}$$

Let us consider now the measure

$$\boldsymbol{\rho}_\varepsilon = \frac{1}{\varepsilon}\int_0^\varepsilon (F_\varepsilon^\tau)_\sharp \boldsymbol{\rho}\, d\tau.$$

We claim that $\boldsymbol{\rho}_\varepsilon$ is a q-plan: it is clear that $\boldsymbol{\rho}_\varepsilon$ is a probability measure on $\mathrm{AC}^q([0,1];X)$, and so we have to check only the marginals at every time:

$$\begin{aligned}(\mathrm{e}_t)_\sharp \boldsymbol{\rho}_\varepsilon &= \frac{1}{\varepsilon}\int_0^\varepsilon (\mathrm{e}_t)_\sharp\big((F_\varepsilon^\tau)_\sharp \boldsymbol{\rho}\big)\, d\tau = \frac{1}{\varepsilon}\int_0^\varepsilon (\mathrm{e}_{\frac{t+\tau}{1+\varepsilon}})_\sharp \boldsymbol{\rho}\, d\tau \\ &= \frac{1+\varepsilon}{\varepsilon}\int_{\frac{t}{1+\varepsilon}}^{\frac{t+\varepsilon}{1+\varepsilon}} (\mathrm{e}_s)_\sharp \boldsymbol{\rho}\, ds \le \frac{1+\varepsilon}{\varepsilon}\int_0^1 (\mathrm{e}_s)_\sharp \boldsymbol{\rho}\, ds \le C\frac{1+\varepsilon}{\varepsilon}\mathfrak{m} \qquad \text{for all } t\in[0,1].\end{aligned}$$

Now we reach the absurd if we show that $\boldsymbol{\rho}_\varepsilon$ is concentrated on Γ; in order to do so it is sufficient to notice that $F_\varepsilon^\tau = s_I$ with $I = I_\varepsilon^\tau = [\frac{\tau}{1+\varepsilon}, \frac{1+\tau}{1+\varepsilon}]$ and $\tau\in[0,\varepsilon]$.

Now if we assume also that $[\Gamma]\subseteq\Gamma$ then we know that given a curve $\gamma\in\Gamma$ then $\gamma\circ s_1^{-1} =: \eta\in\Gamma\cap\mathrm{AC}^\infty([0,1];X)$, where s_1 is the parametrization defined in Proposition 1.2.1. We recall that by definition we have $(1+\ell(\gamma))s'(t) = 1+|\dot\gamma_t|$; in particular, by the change of variable formula

$$\int_0^1 (1+|\dot\gamma_t|)g(\gamma_t)\,\mathrm{d}t = (1+\ell(\gamma))\int_0^1 g(\eta_s)\,\mathrm{d}s \qquad \forall g \text{ Borel function.} \tag{2.6.5}$$

We suppose that $M\big(\Gamma\cap\mathrm{AC}^\infty([0,1];X)\big)$ is $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible; this gives us a p-integrable Borel function f such that

$$\int_0^1 f(\gamma_t)\,\mathrm{d}t = \infty \qquad \forall\gamma\in\Gamma\cap\mathrm{AC}^\infty([0,1];X). \tag{2.6.6}$$

Now given any q-plan $\boldsymbol{\pi}$ we have that

$$\begin{aligned}\iint_0^1 (|\dot\gamma_t|+1)f(\gamma_t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi} &\le \left(\iint_0^1 (|\dot\gamma_t|+1)^q \mathrm{d}t\,\mathrm{d}\boldsymbol{\pi}\right)^{1/q}\left(\iint_0^1 f(\gamma_t)^p\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi}\right)^{1/p} \\ &\le \left(\left(\int \mathcal{E}_q(\gamma)\,\mathrm{d}\boldsymbol{\pi}\right)^{1/q}+1\right)\left(C(\boldsymbol{\pi})\cdot\int_X f^p\,\mathrm{d}\mathfrak{m}\right)^{1/p} < \infty.\end{aligned} \tag{2.6.7}$$

Now, using (2.6.6), (2.6.7) and (2.6.5) we get precisely that $\boldsymbol{\pi}(\Gamma)=0$.

□

Remark 2.6.5 We note that the proof shows that if Γ is $\sim$-invariant and $M\big(\Gamma\cap\mathrm{AC}^\infty([0,1];X)\big)$ is $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible in $\mathcal{M}_+(X)$ then Γ is q-negligible, also if the stability assumption is dropped.

2.7 Weak upper gradients

As in the previous sections, (X,d) will be a complete and separable metric space and $\mathfrak{m}\in\mathcal{M}_+(X)$.

Recall that a Borel function $g : X \to [0, \infty]$ is an upper gradient of $f : X \to \mathbb{R}$ if

$$|f(\underline{\gamma}_{fin}) - f(\underline{\gamma}_{ini})| \leq \int_{\underline{\gamma}} g \tag{2.7.1}$$

holds for all $\underline{\gamma} \in \mathscr{C}(X)$. Here, the curvilinear integral $\int_{\underline{\gamma}} g$ is given by $\int_J g(\gamma_t)|\dot{\gamma}_t| \, \mathrm{d}t$, where $\gamma : J \to X$ is any parameterization of the curve $\underline{\gamma}$ (i.e., $\underline{\gamma} = \mathsf{i}\gamma$, and one can canonically take $\gamma = \mathsf{j}\underline{\gamma}$). It follows from Proposition 1.2.4 that the upper gradient property can be equivalently written in the form

$$|f(\underline{\gamma}_{fin}) - f(\underline{\gamma}_{ini})| \leq \int_X g \, \mathrm{d}J\underline{\gamma}.$$

Now we introduce two different notions of Sobolev function and a corresponding notion of p-weak gradient; the first one was first given in [75] while the second one [11] in for $p = 2$ and in [9] for general exponent. When discussing the corresponding notions of (minimal) weak gradient we will follow the terminology of [9].

Definition 2.7.1 ($N^{1,p}$ and p-upper gradient) *Let f be a $\mathfrak{m}$-measurable and p-integrable function on X. We say that f belongs to the space $N^{1,p}(X, \mathsf{d}, \mathfrak{m})$ if there exists $g \in \mathscr{L}^p_+(X, \mathfrak{m})$ such that* (2.7.1) *is satisfied for* $\mathrm{Mod}_{p,\mathfrak{m}}$*-a.e. curve $\underline{\gamma}$.*

Functions in $N^{1,p}$ have the important Beppo-Levi property of being absolutely continuous along $\mathrm{Mod}_{p,\mathfrak{m}}$-a.e. curve $\underline{\gamma}$ (more precisely, this means $f \circ \mathsf{j}\underline{\gamma} \in \mathrm{AC}([0,1]; X)$), see [75, Proposition 3.1]. Because of the implication (2.6.2), functions in $N^{1,p}(X, \mathsf{d}, \mathfrak{m})$ belongs to the Sobolev space defined below (see [11], [9]) where (2.7.1) is required for q-a.e. curve γ.

Definition 2.7.2 ($W^{1,p}$ and p-weak upper gradient) *Let f be a $\mathfrak{m}$-measurable and p-integrable function on X. We say that f belongs to the space $W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ if there exists $g \in \mathscr{L}^p_+(X, \mathfrak{m})$ such that*

$$|f(\gamma_1) - f(\gamma_0)| \leq \int_0^1 g(\gamma_t)|\dot{\gamma}_t| \, \mathrm{d}t$$

is satisfied for q-a.e. curve $\gamma \in \mathrm{AC}^q([0,1]; X)$.

We remark that there is an important difference between the two definitions, namely the first one is a priori not invariant if we change the function f on a $\mathfrak{m}$-negligible set, while the second one has this kind of invariace, because for any q-test plan $\boldsymbol{\rho}$, any $\mathfrak{m}$-negligible Borel set N and any $t \in [0,1]$ the set $\{\gamma \ : \ \gamma_t \in N\}$ is $\boldsymbol{\rho}$-negligible. Associated to these two notions are the minimal p-upper gradient and the minimal p-weak upper gradient, both uniquely determined up to $\mathfrak{m}$-negligible sets (for a more detailed discussion, see [9], [75]).

As an application of Theorem 2.6.4, we show that these two notions are essentially equivalent modulo the choice of a representative in the equivalence class: more precisely, for any $f \in W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ there exists a $\mathfrak{m}$-measurable representative $\tilde{f}$ of f which belongs to $N^{1,p}(X, \mathsf{d}, \mathfrak{m})$. This result is not new, because in [11] and [9] the equivalence has already been shown. On the other hand, the proof of the equivalence in [11] and [9] is by no means elementary, it passes through the use of tools from the theory of gradient flows and optimal transport theory and it provides the equivalence with another relevant notion of "relaxed" gradient based on the approximation through Lipschitz functions. We provide a totally different proof, using the results proved in this paper about negligibility of sets of curves.

In the following theorem we provide, first, existence of a "good representative" of f. Notice that the standard theory of Sobolev spaces provides existence of this representative via approximation with Lipschitz functions.

Theorem 2.7.3 (Good representative) *Let $f : X \to \mathbb{R}$ be a Borel function and let us set*

$$\Gamma = \big\{\gamma \in \mathrm{AC}^\infty([0,1];X) \ : \ f \circ \gamma \textit{ has a continuous representative } f_\gamma : [0,1] \to \mathbb{R}\big\}.$$

If $\mathrm{Mod}_{p,\mathfrak{m}}\big(M(\mathrm{AC}^\infty([0,1];X) \setminus \Gamma)\big) = 0$ *there exists a $\mathfrak{m}$-measurable representative $\tilde f : X \to \mathbb{R}$ of f satisfying*

$$\mathrm{Mod}_{p,\mathfrak{m}}\big(M(\{\gamma \in \Gamma \ : \ \tilde f \circ \gamma \not\equiv f_\gamma\})\big) = 0. \tag{2.7.2}$$

In particular

(i) for q-a.e. curve γ there holds $\tilde f \circ \gamma \equiv f_\gamma$;

(ii) for $\mathrm{Mod}_{p,\mathfrak{m}}$-a.e. curve $\underline{\gamma}$ there holds $\tilde f \circ \mathrm{j}\underline{\gamma} \equiv f_{\mathrm{j}\underline{\gamma}}$.

Proof. Let us set $\tilde\Gamma := \mathrm{AC}^\infty([0,1];X) \setminus \Gamma$, so that our assumption reads $\mathrm{Mod}_{p,\mathfrak{m}}(M(\tilde\Gamma)) = 0$. Notice first that the (ii) makes sense because $f_{\mathrm{j}\underline{\gamma}}$ exists for $\mathrm{Mod}_{p,\mathfrak{m}}$-a.e. curve $\underline{\gamma}$ thanks to (2.5.2) and $\mathrm{Mod}_{p,\mathfrak{m}}(M(\tilde\Gamma \cap \mathrm{AC}^\infty_c([0,1];X))) = 0$ (also, constant curves are all contained in Γ). Also (i) makes sense thanks to 2.6.5 and the fact that the property of having a continuous representative is $\sim$-invariant.

Step 1. (Construction of a good set Γ_g of curves). Since we have $\mathrm{Mod}_{p,\mathfrak{m}}(M(\tilde\Gamma)) = 0$, there exists $h \in \mathcal{L}^p_+(X,\mathfrak{m})$ such that $\int_0^1 h \circ \sigma = \infty$ for every $\sigma \in \tilde\Gamma$. Starting from Γ and h, we can define the set $\Gamma_g = \big\{\eta \in \Gamma \ : \ \int_0^1 h \circ \eta < \infty\big\}$ of "good" curves, satisfying the following three conditions:

(a) $f \circ \eta$ has a continuous representative for all $\eta \in \Gamma_g$;

(b) $\int_0^1 h \circ \eta < \infty$ for all $\eta \in \Gamma_g$;

(c) $M\big(\mathrm{AC}^\infty([0,1];X) \setminus \Gamma_g\big)$ is $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible.

Indeed, properties (a) and (b) follow easily by definition, while (c) follows by the inclusion

$$M\big(\mathrm{AC}^\infty([0,1];X) \setminus \Gamma_g\big) \subseteq M\big(\mathrm{AC}^\infty([0,1];X) \setminus \Gamma\big) \cup \Big\{\mu : \int_X h \,\mathrm{d}\mu = \infty\Big\}.$$

Step 2. (Construction of $\tilde f$). For every point $x \in X$ we consider the set of pairs good curves-times that pass through x at time t:

$$\Theta_x = \{(\eta, t) \in \Gamma_g \times [0,1] \ : \ \eta(t) = x\},$$

and, thanks to property (a) of Γ_g, we can partition this set according to the value of the continuous representative f_η at t:

$$\Theta_x = \bigcup_{r \in \mathbb{R}} \Theta_x^r \qquad \text{with} \qquad \Theta_x^r = \{(\eta, t) \in \Theta_x \ : \ f_\eta(t) = r\}.$$

Now, the key point is that for every $x \in X$ there exists at most one r such Θ_x^r is not empty. Indeed, suppose that $r_1 \neq r_2$ and that there exist $(\eta_1, t_1) \in \Theta_x^{r_1}$, $(\eta_2, t_2) \in \Theta_x^{r_2}$, so that $r_1 = f_{\eta_1}(t_1) \neq f_{\eta_2}(t_2) = r_2$; since η_1, $\eta_2 \in \Gamma_g$, property (b) of Γ_g gives

$$\int_0^1 h \circ \eta_1 \, \mathrm{d}t + \int_0^1 h \circ \eta_2 \, \mathrm{d}t < \infty. \tag{2.7.3}$$

Suppose to fix the ideas that $t_1 > 0$ and $t_2 < 1$ (otherwise we reverse time for one curve, or both, in the following argument). Now we create a new curve $\eta_3 \in \mathrm{AC}^\infty([0,1]; X)$ by concatenation:

$$\eta_3(s) := \begin{cases} \eta_1(2st_1) & \text{if } s \in [0, 1/2], \\ \eta_2(1 - 2(1-s)(1-t_2)) & \text{if } s \in [1/2, 1]. \end{cases}$$

This curve is clearly absolutely continuous and it follows first η_1 for half of the time and then it follows η_2; it is clear that, since $f \circ \eta_3$ coincides $\mathscr{L}^1$-a.e. in $(0,1)$ with the function

$$a(s) := \begin{cases} f_{\eta_1}(2st_1) & \text{if } s \in [0, 1/2], \\ f_{\eta_2}(1 - 2(1-s)(1-t_2)) & \text{if } s \in [1/2, 1] \end{cases}$$

which has a jump discontinuity at $s = 1/2$, $f \circ \eta_3$ has no continuous representative. It follows that η_3 belongs to $\tilde{\Gamma}$ and therefore $\int_0^1 h \circ \eta_3 = \infty$. But, since

$$\frac{1}{2t_1} \int_0^1 h \circ \eta_1 \, \mathrm{d}t + \frac{1}{2(1-t_2)} \int_0^1 h \circ \eta_2 \, \mathrm{d}t \geq \int_0^1 h \circ \eta_3 \, \mathrm{d}t$$

we get a contradiction with (2.7.3).

Now we define

$$\tilde{f}(x) := \begin{cases} f_\eta(t) & \text{if } (\eta, t) \in \Theta_x \text{ for some } \eta \in \Gamma_g,\ t \in [0,1] \\ f(x) & \text{otherwise.} \end{cases}$$

By construction, $\tilde{f}(\eta(t)) = f_\eta(t)$ for all $t \in [0,1]$ and $\eta \in \Gamma_g$, so that property (c) of Γ_g shows (2.7.2) which implies also that that

$$\mathrm{Mod}_{p,\mathfrak{m}}\big(M(\{\gamma \in \Gamma \cap \mathrm{AC}_c^\infty([0,1]; X) \ : \ \tilde{f} \circ \gamma \not\equiv f_\gamma\})\big) = 0.$$

Recalling (2.5.2) and the fact that j is a Borel isomorphism, we can rewrite this last equation as

$$\mathrm{Mod}_{p,\mathfrak{m}}\big(J(\{\underline{\gamma} \in \mathscr{C}(X) \ : \ \tilde{f} \circ \mathsf{j}\underline{\gamma} \not\equiv f_{\mathsf{j}\underline{\gamma}}\})\big) = 0,$$

and so we proved (ii).Using 2.6.5 and the fact that $\{\gamma \ : \ \tilde{f} \equiv f_\gamma\}$ is clearly a $\sim$-invariant set gives (i).

Step 3. (The set $F := \{f \neq \tilde{f}\}$ is $\mathfrak{m}$-negligible.) Let γ^x be the curve identically equal x, that is $\gamma_t^x = x$ for all $t \in [0,1]$. It is clear that γ^x belongs to Γ for every $x \in X$: in particular $f_{\gamma^x}(t) = f(x)$ for every $t \in [0,1]$. The basic observation is that if we consider the set $\tilde{\Gamma}_c$ of constant curves γ^x satisfying $\tilde{f} \circ \gamma^x \not\equiv f_{\gamma^x}$, then $f(x) \neq \tilde{f}(x)$ for every such curve, hence $\tilde{\Gamma}_c = \{\gamma^x \ : \ x \in F\}$. In particular we have that $M(\tilde{\Gamma}_c) = \{\delta_x \ : \ x \in F\}$. Now, from (2.7.2), we know that $\mathrm{Mod}_{p,\mathfrak{m}}(M(\tilde{\Gamma}_c)) = 0$; this provides the existence of $g \in \mathcal{L}_+^p(X, \mathfrak{m})$ such that $g(x) = \infty$ for every $x \in F$, and so we get that F is contained in a $\mathfrak{m}$-negligible set. $\square$

The following simple example shows that, in Theorem 2.7.3, the "nonparametric" assumption that $J(\mathrm{AC}([0,1];X)\setminus\Gamma)$ is $\mathrm{Mod}_{p,\mathfrak{m}}$-negligible is not sufficient to conclude that $\tilde f = f$ $\mathfrak{m}$-a.e. in X.

Example 2.7.4 Let $X=[0,1]$, d the Euclidean distance, $\mathfrak{m}=\mathscr{L}^1+\delta_{1/2}$, $p\in[1,\infty)$. The function f identically equal to 0 on $X\setminus\{1/2\}$ and equal to 1 at $x=1/2$ has a continuous (actually, identically equal to 0) representative $f_{\mathrm{j}\underline{\gamma}}$ for $\mathrm{Mod}_{p,\mathfrak{m}}$-a.e. curve $\underline{\gamma}$, but any function $\tilde f$ such that $\tilde f\circ \mathrm{j}\underline{\gamma}\equiv f_{\mathrm{j}\underline{\gamma}}$ for $\mathrm{Mod}_{p,\mathfrak{m}}$-a.e. $\underline{\gamma}$ should be equal to 0 also at $x=1/2$, so that $\mathfrak{m}(\{f\neq\tilde f\})=1$.

Now, we are going to apply Theorem 2.7.3 to the problem of equivalence of Sobolev spaces. We begin with a few preliminary results and definitions.

Let $f:X\to\mathbb{R}$, $g:X\to[0,\infty]$ be Borel functions. We consider the sets

$$\mathcal{I}(g):=\Big\{\gamma\in\mathrm{AC}([0,1];X):\int_\gamma g<\infty\Big\},\tag{2.7.4}$$

and

$$\mathcal{B}(f,g):=\Big\{\gamma\in\mathcal{I}(g):f\circ\gamma\in W^{1,1}(0,1),\quad |\frac{\mathrm{d}}{\mathrm{d}t}(f\circ\gamma)|\le|\dot\gamma|g\circ\gamma\ \mathscr{L}^1\text{-a.e. in }(0,1)\Big\}.\tag{2.7.5}$$

We will need the following simple measure theoretic lemma, which says that integration in one variable maps Borel functions to Borel functions. Its proof is an elementary consequence of a monotone class argument (see for instance [20, Theorem 2.12.9(iii)]) and of the fact that the statement is true for F bounded and continuous.

Lemma 2.7.5 *Let (Y,d_Y) be a metric space and let $F:[0,1]\times Y\to[0,\infty]$ be Borel. Then the function $\mathcal{I}_F:Y\to[0,\infty]$ defined by $y\mapsto\int_0^1 F(t,y)\,\mathrm{d}t$ is a Borel function.*

Lemma 2.7.6 *Let $f:X\to\mathbb{R}$, $g:X\to[0,\infty]$ be Borel functions. Then $\mathcal{I}(g)\setminus\mathcal{B}(f,g)$ is a Borel set, stable and $\sim$-invariant.*

Proof. Stability is simple to check: if, by contradiction, it were $\gamma\in\mathcal{I}(g)\setminus\mathcal{B}(f,g)$ and $\mathsf{s}_{[a_n,b_n]}\gamma\in\mathcal{B}(f,g)$ with $a_n\downarrow 0$ and $b_n\uparrow 1$, we would get $f\circ\gamma\in W^{1,1}(a_n,b_n)$ and $|\frac{\mathrm{d}}{\mathrm{d}t}f\circ\gamma|\le|\dot\gamma|g\circ\gamma\in L^1(0,1)$ $\mathscr{L}^1$-a.e. in (a_n,b_n). Taking limits, we would obtain $\gamma\in\mathcal{B}(f,g)$, a contradiction.

For the proof of $\sim$-invariance we note that, first of all, that Lemma 2.7.5 with $F(t,\gamma):=g(\gamma_t)|\dot\gamma_t|$ guarantees that $\mathcal{I}(g)$ is a $\sim$-invariant Borel set, provided we define F using a Borel representative of $|\dot\gamma|$; this can be achieved, for instance, using the lim inf of the metric difference quotients. Analogously, the set

$$\mathsf{L}:=\big\{\gamma\in\mathrm{AC}([0,1];X)\ :\ \int_0^1|f(\gamma_t)|\,\mathrm{d}t<\infty\big\}$$

is Borel. Now, $\gamma\in\mathcal{B}(f,g)$ if and only if $\gamma\in\mathcal{I}(g)\cap\mathsf{L}$ and

$$\Big|\int_0^1\varphi'(t)f(\gamma_t)\,\mathrm{d}t\Big|\le\int_0^1|\varphi(t)|g(\gamma_t)|\dot\gamma_t|\,\mathrm{d}t\qquad\text{for all }\varphi\in W\tag{2.7.6}$$

with $W = \{\varphi \in \mathrm{AC}([0,1];[0,1]) \ : \ \varphi(0) = \varphi(1) = 0\}$. Now, if both s and s^{-1} are absolutely continuous from $[0,1]$ to $[0,1]$, setting $\eta := \gamma \circ \mathsf{s}$, we can use the change of variables formula to obtain that $(\varphi \circ \mathsf{s})' f \circ \eta \in L^1(0,1)$ for all $\varphi \in W$ and that

$$\left| \int_0^1 (\varphi \circ \mathsf{s})'(r) f(\eta_r)\, \mathrm{d}r \right| \le \int_0^1 |\varphi \circ \mathsf{s}(r)| g(\eta_r) |\dot\eta_r|\, \mathrm{d}r \qquad \text{for all } \varphi \in W.$$

Since $W \circ \mathsf{s} = W$ we eventually obtain $\varphi' f \circ \eta \in L^1(0,1)$ for all $\varphi \in W$ (so that $f \circ \eta$ is locally integrable in $(0,1)$) and

$$\left| \int_0^1 \varphi'(r) f(\eta_r)\, \mathrm{d}r \right| \le \int_0^1 |\varphi(r)| g(\eta_r) |\dot\eta_r|\, \mathrm{d}r \qquad \text{for all } \varphi \in W.$$

It is easy to check that these two conditions, in combination with $\int_\eta g < \infty$, imply that $\eta \in \mathsf{L}$, therefore $f \circ \eta$ belongs to $\mathcal{B}(f,g)$ and $\sim$-invariance is proved.

In order to prove that $\mathcal{B}(f,g)$ is Borel we follow a similar path: we already know that both $\mathcal{I}(g)$ and L are Borel, and then in the class $\mathcal{I}(g) \cap \mathsf{L}$ the condition (2.7.6), now with W replaced by a countable dense subset of $\mathrm{C}^1_c(0,1)$ for the C^1 norm, provides a characterization of $\mathcal{B}(f,g)$. Since for $\varphi \in \mathrm{C}^1_c(0,1)$ fixed the maps

$$\eta \in \mathsf{L} \mapsto \int_0^1 \varphi'(r) f(\eta_r)\, \mathrm{d}r, \qquad \eta \mapsto \int_0^1 |\varphi(r)| g(\eta_r) |\dot\eta_r|\, \mathrm{d}r$$

are easily seen to be Borel in $\mathrm{AC}([0,1];X)$ (as a consequence of Lemma 2.7.5, splitting in positive and negative part the first integral and using once more a Borel representative of $|\dot\eta|$ in the second integral) we obtain that $\mathcal{B}(f,g)$ is Borel. □

Theorem 2.7.7 (Equivalence theorem) *Any $f \in N^{1,p}(X, \mathsf{d}, \mathfrak{m})$ belongs to $W^{1,p}(X, \mathsf{d}, \mathfrak{m})$. Conversely, for any $f \in W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ there exists a $\mathfrak{m}$-measurable representative $\tilde f$ that belongs to $N^{1,p}(X, \mathsf{d}, \mathfrak{m})$. More precisely, $\tilde f$ satisfies:*

(i) $\tilde f \circ \gamma \in \mathrm{AC}([0,1];X)$ *for q-a.e. curve* $\gamma \in \mathrm{AC}^\infty([0,1];X)$;

(ii) $\tilde f \circ \mathsf{j}\underline{\gamma} \in \mathrm{AC}([0,1];X)$ *for* $\mathrm{Mod}_{p,\mathfrak{m}}$*-a.e. curve* $\underline{\gamma}$.

Proof. We already discussed the easy implication from $N^{1,p}$ to $W^{1,p}$, so let us focus on the converse one.In the sequel we fix $f \in W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ and a p-weak upper gradient g. By Fubini's theorem, it is easily seen that the space $W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ is invariant under modifications in $\mathfrak{m}$-negligible sets; as a consequence, since the Borel σ-algebra is countably generated, we can assume with no loss of generality that f is Borel. Another simple application of Fubini's theorem (see [9, Remark 4.10]) shows that for q-a.e. curve γ there exists an absolutely continuous function $f_\gamma : [0,1] \to \mathbb{R}$ such that $f_\gamma = f$ $\mathscr{L}^1$-a.e. in $(0,1)$ and $|\frac{\mathrm{d}}{\mathrm{d}t} f_\gamma| \le |\dot\gamma| g \circ \gamma$ $\mathscr{L}^1$-a.e. in $(0,1)$. Since the L^q integrability of g yields that the complement of $\mathcal{I}(g)$ is q-negligible, we can use Lemma 2.7.6 and Theorem 2.6.4(ii) to infer that $\Sigma = \mathcal{I}(g) \setminus \mathcal{B}(f,g)$ satisfies $\mathrm{Mod}_{p,\mathfrak{m}}\big(M(\Sigma \cap AC^\infty([0,1];X))\big) = 0$.

By Theorem 2.7.3 we obtain a $\mathfrak{m}$-measurable representative $\tilde f$ of f such that $\tilde f \circ \gamma \equiv f_\gamma$ for q-a.e. curve γ and $\tilde f \circ \mathsf{j}\underline{\gamma} \equiv f_{\mathsf{j}\underline{\gamma}}$ for $\mathrm{Mod}_{p,\mathfrak{m}}$-a.e. $\underline{\gamma}$. Hence, the fundamental theorem of calculus for absolutely continuous functions gives

$$|\tilde f(\underline{\gamma}_{fin}) - \tilde f(\underline{\gamma}_{ini})| = |f_{\mathsf{j}\underline{\gamma}}(1) - f_{\mathsf{j}\underline{\gamma}}(0)| \le \int_0^1 g((\mathsf{j}\underline{\gamma})_t) |(\dot{\mathsf{j}\underline{\gamma}})_t|\, \mathrm{d}t = \int_{\underline{\gamma}} g$$

for $\mathrm{Mod}_{p,\mathfrak{m}}$-a.e. $\underline{\gamma}$. □

CHAPTER 3

Orlicz-Sobolev Spaces

Let (X, d) be a complete and separable metric space and let $\mathfrak{m}$ be a nonnegative Borel measure in X that is finite on bounded sets. In this chapter we introduce and compare two notions of Orlicz-Sobolev space on X, the first obtained by relaxation of the asymptotic Lipschitz constant, the second obtained by a suitable weak upper gradient property. Eventually we will show that the two notions of Sobolev functions coincide; the equivalence is valid for any N-function Φ. In a subsequent section we illustrate how this result generalize [9], showing that in the case Ψ (the convex conjugate of Φ) is doubling, we can define also a notion of modulus of gradient that coincides with other notions of gradient, generalization of the ones introduced in [25], [57], [75], described in the appendix. The proof follows closely [9], but, choosing properly the energy used for the constriction of the gradient flow, we are able to achieve the proof looking at dissipation of a functional independent of the function Φ, namely the squared norm of the function.

We briefly summarize the proof: in Section 1.6 we studied the properties of the Hopf-Lax semigroup

$$Q_t f(x) := \inf_{y\in X} f(y) + t\Psi\Big(\frac{\mathsf{d}(x,y)}{t}\Big),$$

for which we proved the differential inequality

$$\frac{\mathrm{d}}{\mathrm{d}t} Q_t f(x) + \Phi\big(\mathrm{lip}_a(Q_t f, x)\big) \leq 0,$$

that will play an important role in our analysis. In Section 3.1 and Section 3.2 we present and compare the two definitions of Φ-Sobolev spaces we already mentioned in the Introduction, while in Section 3.1.1 we gather a few facts on the gradient flow of $\mathcal{F}_v^\Phi$ that are used in Section 3.3 to prove our main result.

Basically, the proof is achieved controlling the dissipation of the function $f \mapsto \|f\|_2^2$ along the gradient flow of $\mathcal{F}_v^\Phi$ in two different ways: on one side we use properties of the gradient flow that involves $\mathcal{F}_v^\Phi$. On the other side we consider $f_t\mathfrak{m}$ as a absolutely continuous curve in the probability space with respect to W_Ψ. Then, thanks to superposition principle 1.7.2, we lift this curve of measures to a measure on the space of curves, and then we can use the BL definition to estimate the dissipation with $\mathcal{F}_{BL}^\Phi$.

The proof can be extended also to the degenerate case $\Phi(t) = t$, provided that a suitable version of the Hamilton-Jacobi inequality for the Hopf-Lax formula is found: this is done in Chapter 4, that is a smaller and improved version of [5].

3.1 Variational definition and relaxed energy $\mathcal{F}_v^\Phi$

In this section a notion of Orlicz-Sobolev space by a relaxation procedure is presented. First we define

$$\mathrm{F}_\Phi(f) = \begin{cases} \|\mathrm{lip}_a(f,\cdot)\|_{(\Phi),\mathfrak{m}} & \text{if } f \in \mathrm{Lip}_0(X,\mathsf{d}) \\ +\infty & \text{otherwise,} \end{cases}$$

where the dual Orlicz norm $\|\cdot\|_{(\Phi),\mathfrak{m}}$ is recalled in Definition 1.5.1. Then we consider $\mathcal{F}_v^\Phi$, the lower semicontinuous relaxation of $\mathcal{F}_v^\Phi$ with respect to the L^1 topology:

$$\mathcal{F}_v^\Phi(f) = \inf\Big\{\liminf_{n\to\infty} \mathrm{F}_\Phi(f_n) \;:\; f_n \to f \text{ in } L^1(X,\mathfrak{m})\Big\}.$$

We will call this function the Φ-relaxed energy functional. We recall that $D(\mathcal{F})$ is defined as the domain of finiteness of a functional $\mathcal{F}$.

Definition 3.1.1 (Variational H-definition) *The space $H_v^{1,\Phi}(X,\mathsf{d},\mathfrak{m})$ is defined as $D(\mathcal{F}_v^\Phi)$. In particular a function $f \in L^1(X,\mathfrak{m})$ belongs to the space $H_v^{1,\Phi}(X,\mathsf{d},\mathfrak{m})$ if and only if there are Lipschitz functions with bounded support f_n such that $f_n \to f$ in $L^1(X,\mathfrak{m})$ and*

$$\sup_{n\in\mathbb{N}} \|\mathrm{lip}_a(f_n,\cdot)\|_{(\Phi),\mathfrak{m}} < \infty.$$

Here the subscript v stands for *variational*. It is easy to see that $H^{1,\Phi}$ is a vector space: it follows from the fact that $\mathcal{F}_v^\Phi$ is convex and positively 1-homogeneous. Convexity of $\mathcal{F}_v^\Phi$ follows by the more precise inequality for the asymptotic Lipschitz constant

$$\mathrm{lip}_a(\lambda f + \mu g) \le |\lambda|\mathrm{lip}_a(f) + |\mu|\mathrm{lip}_a(g) \tag{3.1.1}$$

which simply follows by homogeneity and convexity of $f \mapsto \mathrm{lip}_a(f,x)$. Moreover these properties yield that the map $N : f \mapsto \|f\|_1 + \mathcal{F}_v^\Phi(f)$ is actually a norm. Using the semicontinuity of $\mathcal{F}_v^\Phi$ with respect to the L^1 convergence we find also that $H^{1,\Phi}$ is complete with respect to the norm N. We will call this norm the $H^{1,\Phi}$ norm, and denoted by $\|f\|_{H^{1,\Phi}} := N(f)$. Unlike the case $\Phi(t) = t^p/p$, at this level of generality we can't expect to find a *modulus of the gradient* (see Section 3.4.1); however we will see that if the convex conjugate of Φ is doubling we have its existence (Theorem 3.4.8).

Remark 3.1.2 *It is obvious that $\mathcal{F}_v^\Phi(f) \le F_\Phi(f)$, in particular if $f \in \mathrm{Lip}_0(X,\mathsf{d})$ then we have that $f \in H_v^{1,\Phi}$ and $\mathcal{F}_v^\Phi(f) \le \|\mathrm{lip}_a(f)\|_{(\Phi),\mathfrak{m}}$.*

However the same thing is not obvious when f is a Lipschitz function and $f \in L^1$, so we prove it in the next proposition:

Proposition 3.1.3 *Let $f \in L^1(X,\mathfrak{m})$ be a bounded Lipschitz function. Then we have that $\mathcal{F}_v^\Phi(f) \le \|\mathrm{lip}_a(f)\|_{(\Phi),\mathfrak{m}}$. In particular, if $\mathfrak{m}$ is finite then the constant functions have null energy, and this implies that $\mathcal{F}_v^\Phi(C + f) = \mathcal{F}_v^\Phi(f)$ for every $f \in L^1$, $C > 0$.*

Proof. Fix a point $x \in X$, and let χ_r be a sequence of 1-Lipschitz function such that $\chi_{B_r(x)} \leq \chi_r \leq \chi_{B_{r+2}(x)}$. Then let us consider the sequence $(\chi_n f) \subset \mathrm{Lip}_0(X, \mathsf{d})$; it is obvious that $\chi_n f \to f$ in L^1 so we know that

$$\mathcal{F}_v^\Phi(f) \leq \liminf_{n\to\infty} \|\chi_n f\|_{(\Phi)}.$$

We can assume $\|\mathrm{lip}_a(f)\|_{(\Phi)} < \infty$; now we have $\mathrm{lip}_a(\chi_n f) \leq \chi_n \mathrm{lip}_a(f) + f\chi_{B_{n+2}\setminus B_n}$ and so

$$\|\chi_n f\|_{(\Phi)} \leq \|\mathrm{lip}_a(f)\|_{(\Phi)} + \|f\chi_{B_{n+2}\setminus B_n}\|_{(\Phi)};$$

now, using that $f\chi_{B_{n+2}\setminus B_n} \to 0$ pointwise and $f \in L^\infty \cap L^1 \subset M^\Phi$, we get by dominated convergence (Lemma 1.5.3) that the last term in the right hand side is going to zero and so we get the thesis.

Whenever $\mathfrak{m}$ is finite we have that the constant functions are bounded integrable Lipschitz functions and so their energy can be estimated with the (Φ)-norm of their asymptotic Lipschitz constant, that is 0. In particular, by convexity and homogeneity, we have

$$\mathcal{F}_v^\Phi(f) - \mathcal{F}_v^\Phi(C) \leq \mathcal{F}_v^\Phi(C+f) \leq \mathcal{F}_v^\Phi(f) + \mathcal{F}_v^\Phi(C),$$

and so, since $\mathcal{F}_v^\Phi(C) = 0$ we have proved also the last assertion. □

3.1.1 Gradient flow of $\mathcal{F}_v^\Phi$

In this subsection we assume that $\mathfrak{m}(X) < \infty$. In the proof of equivalence a relevant role is retained by the gradient flow of the convex and lower semicontinuous functional $\mathcal{F}_v^\Phi : L^2(X, \mathfrak{m}) \to [0, \infty]$; we can consider this functional, thanks to the fact that $\mathcal{F}_v^\Phi$ is defined on L^1, but since $\mathfrak{m}$ is finite, we have $L^2 \subset L^1$. With a slight abuse of notation we will keep the notation $\mathcal{F}_v^\Phi$ for this restricted functional. The convexity has been already proved, while the lower semicontinuity in L^2 simply follows by the lower semicontinuity in L^1 and the fact that $\mathfrak{m}$ is finite. In addition, the domain of $\mathcal{F}_v^\Phi$,

$$D(\mathcal{F}_v^\Phi) = H_v^{1,\Phi}(X, \mathsf{d}, \mathfrak{m}) \cap L^2(X, \mathfrak{m})$$

is dense in $L^2(X, \mathfrak{m})$, because it contains $\mathrm{Lip}_0(X, \mathsf{d})$. Thanks to these facts we can apply the standard theory of gradient flows [23] of convex lower semicontinuous functionals in Hilbert spaces, recalled in Section 1.4 to obtain, starting from any $f_0 \in L^2(X, \mathfrak{m})$, a curve f_t such that:

(a) $t \mapsto f_t$ is locally Lipschitz from $(0, \infty)$ to $L^2(X, \mathfrak{m})$ and $f_t \to f_0$ strongly in L^2 as $t \downarrow 0$;

(b) $t \mapsto \mathcal{F}_v^\Phi(f_t)$ is locally absolutely continuous in $(0, \infty)$;

(c) $\frac{\mathrm{d}}{\mathrm{d}t} f_t = \Delta_\Phi f_t$ for a.e. $t \in (0, \infty)$.

Here $\Delta_\Phi f$ denotes the opposite of the element of minimal norm of the subdifferential $\partial^- \mathcal{F}_v^\Phi(f)$, when this set is not empty. Namely, $\xi = -\Delta_\Phi f$ satisfies

$$\mathcal{F}_v^\Phi(h) \geq \mathcal{F}_v^\Phi(f) + \int_X \xi(h - f)\,\mathrm{d}\mathfrak{m} \qquad \forall h \in L^2(X, \mathfrak{m}) \tag{3.1.2}$$

and is the vector with smallest $L^2(X,\mathfrak{m})$ norm among those with this property. We will denote by $D(\Delta_\Phi)$ the set of functions for which the subdifferential is not empty.

We can think of the gradient flow also as a semigroup S_t that maps f_0 in f_t. When $\mathfrak{m}(X)$ is finite, a property that will be used is that $S_t(af_0+C) = a\cdot S_{t/a}(f_0)+C$ for all $C\in\mathbb{R}$, $a\in(0,\infty)$; this is true because $\mathcal{F}_v^\Phi$ is positively 1-homogeneus and invariant by addition of a constant (Proposition 3.1.3) and so we get that $\partial^-\mathcal{F}_v^\Phi$ is positively 0-homogeneus and also invariant by addition.

Proposition 3.1.4 (Integration by parts) *For all $f\in D(\Delta_\Phi)$ and $g\in D(\mathrm{Ch}_1)$ it holds*

$$-\int_X g\Delta_\Phi f\,\mathrm{d}\mathfrak{m} \le \mathcal{F}_v^\Phi(g), \tag{3.1.3}$$

with equality if $g=f$.

Proof. Since $-\Delta_\Phi f\in\partial^-\mathcal{F}_1(f)$ it holds

$$\mathcal{F}_v^\Phi(f) - \int_X g\Delta_\Phi f\,\mathrm{d}\mathfrak{m} \le \mathcal{F}_v^\Phi(f+g), \qquad \forall g\in L^2(X,\mathfrak{m}).$$

Now we can use (3.1.1) to estimate $\mathcal{F}_v^\Phi(f+g)$ with $\mathcal{F}_v^\Phi(f)+\mathcal{F}_v^\Phi(g)$, and so we get the first statement. For the second statement we need the converse inequality when $f=g$; but this is easy, because it is sufficient to put $h=0$ in (3.1.2).

□

Proposition 3.1.5 (Some properties of the gradient flow of $\mathcal{F}_v^\Phi$) *Let $f_0\in L^2(X,\mathfrak{m})$ and let (f_t) be the gradient flow of $\mathcal{F}_v^\Phi$ starting from f_0. Then:*
(Mass preservation) $\int f_t\,\mathrm{d}\mathfrak{m} = \int f_0\,\mathrm{d}\mathfrak{m}$ for any $t\ge 0$.
(Maximum principle) If $f_0\le C$ (resp. $f_0\ge c$) $\mathfrak{m}$-a.e. in X, then $f_t\le C$ (resp $f_t\ge c$) $\mathfrak{m}$-a.e. in X for any $t\ge 0$.
(Energy dissipation) Suppose $0<c\le f_0\le C<\infty$ $\mathfrak{m}$-a.e. in X and let $\Theta\in C^2([c,C])$. Then $t\mapsto\int\Theta(f_t)\,\mathrm{d}\mathfrak{m}$ is locally absolutely continuous in $(0,\infty)$ and it holds

$$-\frac{\mathrm{d}}{\mathrm{d}t}\int\Theta(f_t)\,\mathrm{d}\mathfrak{m} \le \mathcal{F}_v^\Phi(\Theta'(f_t)) \qquad \text{for a.e. } t\in(0,\infty),$$

with equality if $\Theta(t)=t^2$.

Proof. (Mass preservation) Just notice that from (3.1.3) we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \pm\mathbf{1} f_t\,\mathrm{d}\mathfrak{m} = \int \pm\mathbf{1}\cdot\Delta_\Phi f_t\,\mathrm{d}\mathfrak{m} \le \mathcal{F}_v^\Phi(\pm\mathbf{1}) = 0 \quad \text{for a.e. } t>0,$$

where $\mathbf{1}$ is the function identically equal to 1, which has Φ-relaxed energy equal to 0 by Proposition 3.1.3.
(Maximum principle) Fix $f\in L^2(X,\mathfrak{m})$, $\tau>0$ and, according to the so-called implicit Euler scheme, let f^τ be the unique minimizer of

$$g \quad\mapsto\quad \mathcal{F}_v^\Phi(g)+\frac{1}{2\tau}\int_X |g-f|^2\,\mathrm{d}\mathfrak{m}.$$

Assume that $f \leq C$. We claim that in this case $f^\tau \leq C$ as well. Indeed, if this is not the case we can consider the competitor $g := \min\{f^\tau, C\}$ in the above minimization problem. By Lemma 3.3.2 we get $\mathcal{F}(g) \leq \mathcal{F}(f^\tau)$ and the L^2 distance of f and g is strictly smaller than the one of f and f^τ as soon as $\mathfrak{m}(\{f^\tau > C\}) > 0$, which is a contradiction. Starting from f_0, iterating this procedure, and using the fact that the implicit Euler scheme converges as $\tau \downarrow 0$ (see [23], [8] for details) to the gradient flow we get the conclusion.
(Energy dissipation) Since $t \mapsto f_t \in L^2(X, \mathfrak{m})$ is locally absolutely continuous and, by the maximum principle, f_t take their values in $[c, C]$ $\mathfrak{m}$-a.e., from the fact that Θ is Lipschitz in $[c, C]$ we get the claimed absolute continuity statement. Now, we know from the Lagrange mean value theorem that exists a function $\xi_t^h : X \to [c, C]$ such that:

$$\Theta(f_{t+h}) - \Phi(f_t) = \Theta'(f_t)(f_{t+h} - f_t) + \frac{1}{2}\Theta''(\xi_t^h)(f_{t+h} - f_t)^2.$$

Dividing by h and integrating in space, we get that, for times where the L^2 derivative of f_t exists (i.e., for almost every t):

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_X \Theta(f_t)\,\mathrm{d}\mathfrak{m} = \int_X \Theta'(f_t)\Delta_\Phi f_t\,\mathrm{d}\mathfrak{m}.$$

We can now use Lemma 3.1.4 with $g = \Theta'(f_t)$ in the right hand side to get the last statement.
□

3.2 Beppo Levi definition and Φ-weak energy

Recall that the evaluation maps $\mathrm{e}_t : C([0,1], X) \to X$ are defined by $\mathrm{e}_t(\gamma) := \gamma_t$. We also introduce the restriction maps $\mathrm{restr}_t^s : C([0,1], X) \to C([0,1], X)$, $0 \leq t \leq s \leq 1$, given by

$$\mathrm{restr}_t^s(\gamma)_r := \gamma_{(1-r)t+rs}, \tag{3.2.1}$$

so that restr_t^s "stretches" the restriction of the curve to $[s, t]$ to the whole of $[0, 1]$.

Our definition of Φ-weak upper gradient is inspired by [5], [9], [11], allowing for exceptional curves in (1.3.5), but with a different notion of exceptional set, compared to [57], [75]. What follows is a generalization, for a general N-function, of the theory of test plans we already encountered in Definition 2.6.1; notice that however here we require $C(\boldsymbol{\pi}) \geq 1$, since this condition will be needed in the proof and we don't have a good homogeneity of the Φ norm with respect to the change of measure. We recall that Ψ is the convex conjugate of Φ.

Definition 3.2.1 (Test plans and negligible sets of curves) *We say that a measure $\boldsymbol{\pi} \in \mathcal{M}_+(C([0,1], X))$ is a* Ψ-test plan *if $\boldsymbol{\pi}$ is concentrated on $AC([0,1], X)$, $\int_0^1 \||\dot\gamma_t|\|_{\Psi,\boldsymbol{\pi}}\mathrm{d}t < \infty$ and there exists a constant $C(\boldsymbol{\pi}) \geq 1$ such that*

$$(\mathrm{e}_t)_\sharp\boldsymbol{\pi} \leq C(\boldsymbol{\pi})\mathfrak{m} \qquad \forall t \in [0,1]. \tag{3.2.2}$$

A set $A \subset C([0,1], X)$ is said to be Φ-negligible *if it is contained in a $\boldsymbol{\pi}$-negligible set for any Ψ-test plan $\boldsymbol{\pi}$. A property which holds for every $\gamma \in C([0,1], X)$, except possibly a Φ-negligible set, is said to hold for Φ-almost every curve.*

Observe that, by definition, $C([0,1],X) \setminus AC([0,1],X)$ is Φ-negligible, so the notion starts to be meaningful when we look at subsets of $AC([0,1],X)$. Notice also that in our definition we let $\boldsymbol{\pi}$ be a finite measure, not only a probability measure, as in the homogeneous case ($\Phi(t) = t^q$).

Remark 3.2.2 An easy consequence of condition (3.2.2) is that if two $\mathfrak{m}$-measurable functions f, $g : X \to \mathbb{R}$ coincide up to a $\mathfrak{m}$-negligible set and $\mathcal{T}$ is an at most countable subset of $[0,1]$, then the functions $f \circ \gamma$ and $g \circ \gamma$ coincide in $\mathcal{T}$ for Φ-almost every curve γ.

Moreover, choosing an arbitrary Ψ-test plan $\boldsymbol{\pi}$ and applying Fubini's Theorem to the product measure $\mathscr{L}^1 \times \boldsymbol{\pi}$ in $(0,1) \times C([0,1];X)$ we also obtain that $f \circ \gamma = g \circ \gamma$ $\mathscr{L}^1$-a.e. in $(0,1)$ for $\boldsymbol{\pi}$-a.e. curve γ; since $\boldsymbol{\pi}$ is arbitrary, the same property holds for Φ-a.e. γ.

Remark 3.2.3 Differently from the same notion for the case $\Phi(t) = t^p$, it is not obvious how the constant $C(\boldsymbol{\pi})$ behave through the localization $\boldsymbol{\pi}^A = \frac{1}{\boldsymbol{\pi}(A)}\boldsymbol{\pi}|_A$. But in our definition we are not forced to have $\boldsymbol{\pi}$ a probability measure, so when we want to localize to a Borel set $A \subset C([0,1];X)$ we simply take $\boldsymbol{\pi}_A = \boldsymbol{\pi}|_A$.

Coupled with the definition of Φ-negligible set of curves, there is the definition of $BL^{1,\Phi}(X, \mathsf{d}, \mathfrak{m})$. In order to find a proper definition, let's try to do some calculation when g is an upper gradient for f, that is, $|f(\gamma_1) - f(\gamma_0)| \leq \int_\gamma g\, ds$ for every $\gamma \in AC([0,1];X)$. In particular we can integrate this inequality with respect to a Ψ-plan $\boldsymbol{\pi}$:

$$\int |f(\gamma_1) - f(\gamma_0)| \,\mathrm{d}\boldsymbol{\pi} \leq \iint_0^1 g(\gamma_t)|\dot{\gamma}_t| \,\mathrm{d}t \,\mathrm{d}\boldsymbol{\pi}. \tag{3.2.3}$$

If we want to let $\|g\|_{(\Phi),\mathfrak{m}}$ appear, we can use Hölder's inequality time by time, with respect to the measure $\boldsymbol{\pi}$, and then use Lemma 1.5.5 and $C(\boldsymbol{\pi}) \geq 1$:

$$\begin{aligned} \int_0^1 \int g(\gamma_t)|\dot{\gamma}_t| \,\mathrm{d}\boldsymbol{\pi} \,\mathrm{d}t &\leq \int_0^1 \|g(\gamma_t)\|_{(\Phi),\boldsymbol{\pi}} \cdot \||\dot{\gamma}_t|\|_{\Psi,\boldsymbol{\pi}} \,\mathrm{d}t \\ &= \int_0^1 \|g\|_{(\Phi),(\mathrm{e}_t)_\sharp\boldsymbol{\pi}} \cdot \||\dot{\gamma}_t|\|_{\Psi,\boldsymbol{\pi}} \,\mathrm{d}t \\ &\leq C(\boldsymbol{\pi})\|g\|_{(\Phi),\mathfrak{m}} \int_0^1 \||\dot{\gamma}_t|\|_{\Psi,\boldsymbol{\pi}} \,\mathrm{d}t. \end{aligned} \tag{3.2.4}$$

Now we are ready to state the definition of Beppo Levi space:

Definition 3.2.4 (Beppo-Levi space) *The Beppo Levi space $BL^{1,\Phi}(X, \mathsf{d}, \mathfrak{m})$ is defined as the set of functions $f \in L^1(X, \mathfrak{m})$ for which there exists a constant $E \geq 0$ such that for every Ψ-plan $\boldsymbol{\pi}$ we have*

$$\int |f(\gamma_1) - f(\gamma_0)| \,\mathrm{d}\boldsymbol{\pi} \leq E \cdot C(\boldsymbol{\pi}) \int_0^1 \||\dot{\gamma}_t|\|_{\Psi,\boldsymbol{\pi}} \,\mathrm{d}t. \tag{3.2.5}$$

The least constant E such that the above inequality holds is called $\mathcal{F}^\Phi_{BL}(f)$, the Φ-weak energy of f.

Remark 3.2.5 It is very easy to see that also $\mathcal{F}^\Phi_{BL}$ is l.s.c. with respect to L^1-convergence; this is true thanks to the fact that the left hand side of (3.2.5) is continuous with respect to convergence in L^1 for every $\boldsymbol{\pi}$, since we have $(\mathrm{e}_t)_\sharp\boldsymbol{\pi} \le C(\boldsymbol{\pi})\mathfrak{m}$ for $t = 0, 1$. Moreover, the left hand side is continuous with respect to $\mathfrak{m}$-a.e. convergence: $f_n \to f$ $\mathfrak{m}$-a.e. implies that for a fixed $t > 0$, $f_n(\gamma_t) \to f(\gamma_t)$ for Φ-a.e. curve; in particular this is true for $t = 0$ and $t = 1$ and we conclude that $\mathcal{F}^\Phi_{BL}$ is l.s.c. with respect to $\mathfrak{m}$-a.e. convergence.

Note that this definition is very mild, compared to the other ones present in the literature (even the one in [5]), but we will see that, thanks to the equivalence theorem, $f \in BL^{1,\Phi}$ implies that $f \circ \gamma$ is also BV along Φ-almost every curve, and moreover if Ψ is doubling then $f \circ \gamma \in W^{1,1}((0,1), \mathscr{L}^1)$ for Φ-almost every γ.

We conclude this definition with a remark about rescaling of plans:

Remark 3.2.6 For every $0 \le s_1 < s_2 \le 1$ let $\mathrm{rest}_{s_1}^{s_2}$ be the restriction map in the interval $[s_1, s_2]$, namely

$$\mathrm{rest}_{s_1}^{s_2}\gamma(t) = \gamma(s_1(1-t) + ts_2).$$

It is straightforward that for every Ψ-plan and every $0 \le s_1 < s_2 \le 1$ we have that $(\mathrm{rest}_{s_1}^{s_2})_\sharp\boldsymbol{\pi}$ is still a Ψ-plan. In particular whenever $\boldsymbol{\pi}$ is a Ψ-plan and $f \in BL^{1,\Phi}(X, \mathsf{d}, \mathfrak{m})$, applying (3.2.5) with $(\mathrm{rest}_{s_1}^{s_2})_\sharp\boldsymbol{\pi}$ we obtain

$$\int |f(\gamma_{s_1}) - f(\gamma_{s_2})| \,\mathrm{d}\boldsymbol{\pi} \le \mathcal{F}^\Phi_{BL}(f) \cdot C(\boldsymbol{\pi}) \int_{s_1}^{s_2} |||\dot\gamma_t|||_{\Psi,\boldsymbol{\pi}} \,\mathrm{d}t \qquad \forall 0 \le s_1 < s_2 \le 1. \tag{3.2.6}$$

3.3 Proof of equivalence

Here we want to show that for every function $f \in L^1$ we have that $\mathcal{F}^\Phi_v(f) = \mathcal{F}^\Phi_{BL}(f)$.

Theorem 3.3.1 *Let $(X, \mathsf{d}, \mathfrak{m})$ be a complete and separable metric measure space, with $\mathfrak{m}$ non-negative Borel measure finite on bounded sets. Then the spaces*

$$H^{1,\Phi}_v(X, \mathsf{d}, \mathfrak{m}), \qquad BL^{1,\Phi}(X, \mathsf{d}, \mathfrak{m})$$

and the corresponding Φ-energies $\mathcal{F}^\Phi_v$ and $\mathcal{F}^\Phi_{BL}$ coincide.

First we state two lemmas that enable us to look only at $f \in L^\infty$ with bounded support.

Lemma 3.3.2 (Continuity on truncations) *Let $h : \mathbb{R} \to \mathbb{R}$ be a 1-Lipschitz function such that $h(0) = 0$. Then for every $f \in L^1$ we have that*

- $\mathcal{F}^\Phi_v(h(f)) \le \mathcal{F}^\Phi_v(f)$;
- $\mathcal{F}^\Phi_{BL}(h(f)) \le \mathcal{F}^\Phi_{BL}(f)$;

in particular, letting $f^N = (f \wedge N) \vee (-N)$ we have that $\mathcal{F}^\Phi_v(f^N) \to \mathcal{F}^\Phi_v(f)$ and the same is true for $\mathcal{F}^\Phi_{BL}$.

Proof. The first assertion follows by the inequality at the level of asymptotic Lipschitz constant $\mathrm{lip}_a(h \circ f) \le \mathrm{lip}_a(f)$ and the fact that $f_n \xrightarrow{L^1} f$ implies $h(f_n) \xrightarrow{L^1} h(f)$. As for the Φ-weak energy, suffices to notice that $|h(f(\gamma_1)) - h(f(\gamma_0))| \le |f(\gamma_1) - f(\gamma_0)|$.

We prove continuity only for $\mathcal{F}_v^\Phi$, the proof for $\mathcal{F}_{BL}^\Phi$ being exactly the same. It is clear that

$$h_N(t) = \begin{cases} N & \text{if } t \geq N \\ t & \text{if } |t| < N \\ -N & \text{if } t \leq -N \end{cases}$$

satisfies the assumption of this lemma and so we have $\mathcal{F}_v^\Phi(f^N) = \mathcal{F}_v^\Phi(h_N(f)) \leq \mathcal{F}_v^\Phi(f)$. Since $f \in L^1$ we have also $h_N(f) \stackrel{L^1}{\to} f$ and so, using the lower semicontinuity of $\mathcal{F}_v^\Phi$ we obtain

$$\mathcal{F}_v^\Phi(f) \geq \limsup_{N\to\infty} \mathcal{F}_v^\Phi(f^N) \geq \liminf_{N\to\infty} \mathcal{F}_v^\Phi(f^N) \geq \mathcal{F}_v^\Phi(f).$$

It follows that $\mathcal{F}_v^\Phi(f^N) \uparrow \mathcal{F}_v^\Phi(f)$. □

Lemma 3.3.3 (Reduction to bounded support) *For every $x \in X$, let χ_r be a family of 1-Lipschitz function such that $\chi_{B_r(x)} \leq \chi_r \leq \chi_{B_{r+2}(x)}$; then we have $\mathcal{F}_v^\Phi(\chi_r f) \to \mathcal{F}_v^\Phi(f)$ for every $f \in L^1 \cap L^\infty$. The same is true for $\mathcal{F}_{BL}^\Phi$.*

Moreover if f has support in B_r, then in the definition of $\mathcal{F}_v^\Phi$ we can take f_n to be Lipschitz function with support contained in B_{r+2}: in formulae

$$\mathcal{F}_v^\Phi(f) = \inf\Big\{\liminf_{n\to\infty} F_\Phi(f_n) \;:\; f_n \to f \text{ in } L^1(X, \mathfrak{m}), \quad \operatorname{supp}(f_n) \subset B_{r+2}\Big\}. \tag{3.3.1}$$

Here $\operatorname{supp}(f)$ *is the smallest closed set S such that $f = 0$ $\mathfrak{m}$-almost everywhere in S^c.*

Proof. First let us note that $\chi_r f \to f$ in L^1 so that by the lower semicontinuity of $\mathcal{F}_v^\Phi$ (and the same for $\mathcal{F}_{BL}^\Phi$)

$$\liminf_{r\to\infty} \mathcal{F}_v^\Phi(\chi_r f) \geq \mathcal{F}_v^\Phi(f).$$

It remains to show the other inequality; in particular we can assume $\mathcal{F}_v^\Phi(f) < \infty$. Let $(f_n) \subset \operatorname{Lip}_0(X, \mathsf{d})$ be an optimal sequence given in the definition of $\mathcal{F}_v^\Phi(f)$; note that, letting $C = \|f\|_\infty$, we can assume $|f_n| \leq C$ otherwise we can take f_n^C as approximating functions and we have $\operatorname{lip}_a(f_n^C) \leq \operatorname{lip}_a(f_n)$. Now consider $(\chi_r f_n) \subset \operatorname{Lip}_0(X, \mathsf{d})$ as an approximating sequence for $\chi_r f$: we have $\operatorname{lip}_a(\chi_r f_n) \leq \chi_r \operatorname{lip}_a(f_n) + f_n \chi_{B_{r+2}\setminus B_r}$ and so

$$\liminf_{n\to\infty} \|\operatorname{lip}_a(\chi_r f_n)\|_{(\Phi),\mathfrak{m}} \leq \liminf_{n\to\infty} \|\operatorname{lip}_a(f_n)\|_{(\Phi),\mathfrak{m}} + \limsup \|f_n \chi_{B_{r+2}\setminus B_r}\|_{(\Phi),\mathfrak{m}}$$

Up to subsequences we have $f_n \to f$ pointwise and since we have $f_n \chi_{B_{r+2}} \leq \|f\|_\infty \chi_{B_{r+2}} \in M^\Phi$, using Lemma 1.5.3, we have also $f_n \chi_{B_{r+2}} \to f \chi_{B_{r+2}}$ strongly in L^Φ. Taking limits:

$$\mathcal{F}_v^\Phi(\chi_r f) \leq \liminf_{n\to\infty} \|\operatorname{lip}_a(\chi_r f_n)\|_{(\Phi)} \leq \mathcal{F}_v^\Phi(f) + \|f \chi_{B_{r+2}\setminus B_r}\|_{(\Phi)}; \tag{3.3.2}$$

again, using $f \chi_{B_{r+2}\setminus B_r} \to 0$ pointwise and $f \in L^1 \cap L^\infty \subset M^\Phi$ we get that the last term is going to 0 and so

$$\limsup_{n\to\infty} \mathcal{F}_v^\Phi(\chi_r f) \leq \mathcal{F}_v^\Phi(f).$$

As for $\mathcal{F}_{BL}^\Phi$ we begin with the obvious inequality

$$|\chi_r(x) f(x) - \chi_r(y) f(y)| \leq |f(x) - f(y)| + |\chi_r(x) - \chi_r(y)| \min\{f(x), f(y)\} \qquad \forall x, y \in X;$$

Then we know that if $x, y \in B_r$ then $\chi_r(x) - \chi_r(y) = 0$ so we can exploit this fact and get

$$|\chi_r(x)f(x) - \chi_r(y)f(y)| \le |f(x) - f(y)| + |\chi_r(x) - \chi_r(y)|\big(|f(x)|\chi_{B_r^c}(x) + |f(y)|\chi_{B_r^c}(y)\big).$$

Using the fact that χ_r is 1-Lipschitz, putting $\gamma_0 = y$ and $\gamma_1 = x$ and $f^r(x) = f(x)\chi_{B_r^c}(x)$ we get

$$\left|\int_{\partial\gamma} \chi_r f\right| \le \left|\int_{\partial\gamma} f\right| + \ell(\gamma)|f^r(\gamma_1)| + \ell(\gamma)|f^r(\gamma_0)| \qquad \forall \gamma \in AC([0,1]; X);$$

Integrating over a Ψ-plan $\boldsymbol{\pi}$, and using that $f \in BL^{1,\Phi}$, we get:

$$\int |(\chi_r f)(\gamma_1) - (\chi_r f)(\gamma_0)| \,\mathrm{d}\boldsymbol{\pi} \le C(\boldsymbol{\pi}) \cdot \mathcal{F}^{\Phi}_{BL}(f) \int_0^1 \|\dot{\gamma_t}\|_{\Psi,\boldsymbol{\pi}} + C_1 + C_0, \tag{3.3.3}$$

where $C_t = \int |f^r(\gamma_t)|\ell(\gamma)\,\mathrm{d}\boldsymbol{\pi}$. Now using Hölder inequality and subadditivity of the norm:

$$\begin{aligned} C_t = \int |f^r(\gamma_t)|\ell(\gamma)\,\mathrm{d}\boldsymbol{\pi} &\le \|f^r(\gamma_t)\|_{(\Phi),\boldsymbol{\pi}} \cdot \|\ell(\gamma)\|_{\Psi,\boldsymbol{\pi}} \\ &= \|f^r\|_{(\Phi),(\mathrm{e}_t)_\sharp\boldsymbol{\pi}} \cdot \left\|\int_0^1 |\dot{\gamma_s}|\,\mathrm{d}s\right\|_{\Psi,\boldsymbol{\pi}} \le C(\boldsymbol{\pi}) \cdot \|f^r\|_{(\Phi),\mathfrak{m}} \int_0^1 \||\dot{\gamma_s}|\|_{\Psi,\boldsymbol{\pi}}\,\mathrm{d}s. \end{aligned}$$

Using this estimate in (3.3.3) we get that

$$\mathcal{F}^{\Phi}_{BL}(\chi_r f) \le \mathcal{F}^{\Phi}_{BL}(f) + 2\|f^r\|_{(\Phi),\mathfrak{m}};$$

letting $r \to \infty$ and noticing, as before, that $\|f^r\|_{(\Phi),\mathfrak{m}} \to 0$, we get the desired inequality.

For the last assertion suffices to notice that in (3.3.2) we have that the last term in the right hand side is equal to 0 and also $\chi_r f = f$ and so $\mathcal{F}^{\Phi}_v(\chi_r f) = \mathcal{F}^{\Phi}_v(f) = \liminf \|\chi_r f_n\|_{(\Phi)}$ and $\chi_r f_n$ are Lipschitz functions with support contained in B_{r+2}. □

Now we prove the easy inequality $\mathcal{F}^{\Phi}_v(f) \ge \mathcal{F}^{\Phi}_{BL}(f)$; notice that in Section 3.2 we proved that $\mathcal{F}^{\Phi}_{BL}$ is lower semicontinuous with respect to the L^1 convergence, and that $\mathcal{F}^{\Phi}_{BL}(f) \le \|g\|_{(\Phi),\mathfrak{m}}$ for every g upper gradient of f (see (3.2.4)). Since for every $f \in \mathrm{Lip}_0(X, \mathsf{d})$ we have that $\mathrm{lip}_a(f)$ is an upper gradient for f, we have that $\mathcal{F}^{\Phi}_{BL} \le F_\Phi$; passing to the L^1-lower semicontinuous relaxations:

$$\mathcal{F}^{\Phi}_{BL}(f) \le \mathcal{F}^{\Phi}_v(f) \qquad \forall f \in L^1(X, \mathfrak{m}).$$

Now we are ready to prove the converse inequality, namely from a function $f \in BL^{1,\Phi}(X, \mathsf{d}, \mathfrak{m})$ we want to build a sequence of approximating Lipschitz functions in such a way that

$$\limsup_{n\to\infty} \|\mathrm{lip}_a(f_n)\|_{(\Phi),\mathfrak{m}} \le \mathcal{F}^{\Phi}_{BL}(f). \tag{3.3.4}$$

As in [9] for the case $\Phi(t) = t^q$ with $1 < q < \infty$ and [5] for the BV case $q = 1$, our main tool in the construction is the gradient flow in $L^2(X, \mathfrak{m})$ of the functional $\mathcal{F}^{\Phi}_v$, starting from f. We initially assume that (X, d) is a complete and separable space and that $\mathfrak{m}$ is a finite Borel measure, so that the L^2-gradient flow of $\mathcal{F}^{\Phi}_v$ can be used. Furthermore, in order to apply the results of Section 1.6, we will assume also that Ψ is a strictly convex function with continuous derivative. The finiteness assumption on $\mathfrak{m}$ and the hypothesis on Ψ will be eventually removed in the proof of the equivalence result.

We start with the following proposition, which relates energy dissipation to a sharp combination of Φ-weak energy and metric dissipation in W_Ψ.

Proposition 3.3.4 *Let $\mu_t = f_t\mathfrak{m}$ be a curve in $AC([0,1],(\mathcal{M}_+(X),W_\Psi))$. Assume that for some $0 < c < C < \infty$ it holds $c \le f_t \le C$ $\mathfrak{m}$-a.e. in X for any $t \in [0,1]$, and that $f_0 \in BL^{1,\Phi}(X,\mathsf{d},\mathfrak{m})$. Then for all $\Theta \in C^2([c,C])$ convex it holds*

$$\int \Theta(f_0)\,\mathrm{d}\mathfrak{m} - \int \Theta(f_s)\,\mathrm{d}\mathfrak{m} \le \mathrm{Lip}(\Theta')\cdot\max\{C,1\}\cdot\mathcal{F}^{\Phi}_{BL}(f)\int_0^s |\dot\mu_t|\,\mathrm{d}t \qquad \forall s>0.$$

Proof. Let $\boldsymbol{\pi} \in \mathcal{M}_+(C([0,1],X))$ be a plan associated to the curve (μ_t) as in Proposition 1.7.2. The assumption $f_t \le C$ $\mathfrak{m}$-a.e. and the fact that $\|\dot\gamma_t\|_{\Psi,\boldsymbol{\pi}} = |\dot\mu_t| \in L^1(0,1)$ guarantee that $\boldsymbol{\pi}$ is an Ψ-test plan, such that $1 \le C(\boldsymbol{\pi}) \le \max\{C,1\}$.

Now, using $f_0 \in BL^{1,\Phi}$ and (3.2.6) with $s_0 = 0$ and $s_1 = s$, we get that:

$$\begin{aligned}\int \Theta(f_0) - \int \Theta(f_s)\,\mathrm{d}\mathfrak{m} &\le \int \Theta'(f_0)(f_0 - f_s)\,\mathrm{d}\mathfrak{m} = \int \Theta'(f_0)\circ \mathrm{e}_0 - \Theta'(f_0)\circ \mathrm{e}_s\,\mathrm{d}\boldsymbol{\pi}\\ &\le \int |\Theta'(f_0(\gamma_s)) - \Theta'(f_0(\gamma_0))|\,\mathrm{d}\boldsymbol{\pi}(\gamma)\\ &\le \mathrm{Lip}(\Theta')\int |f_0(\gamma_s) - f_0(\gamma_0)|\,\mathrm{d}\boldsymbol{\pi}(\gamma)\\ &\le \mathrm{Lip}(\Theta')\cdot C(\boldsymbol{\pi})\cdot \mathcal{F}^{\Phi}_{BL}(f_0)\int_0^s \||\dot\gamma_t|\|_{\Psi,\boldsymbol{\pi}}\,\mathrm{d}t\\ &= \mathrm{Lip}(\Theta')\cdot\max\{C,1\}\cdot\mathcal{F}^{\Phi}_{BL}(f_0)\int_0^s |\dot\mu_t|\,\mathrm{d}t.\end{aligned}$$

□

The key argument to achieve the identification is the following lemma which gives a sharp bound on the W_Ψ-speed of the L^2-gradient flow of $\mathcal{F}^\Phi_v$. A similar lemma, in the W_p case, has been introduced in [58] and then used in [11], [41] to study the heat flow on metric measure spaces; also, the W_∞ case, most similar to this general one, has been studied in [5].

Lemma 3.3.5 (Kuwada's lemma for $\mathcal{F}^\Phi_v$) *Let $f_0 \in L^2(X,\mathfrak{m})$ and let (f_t) be the gradient flow of $\mathcal{F}^\Phi_v$ starting from f_0. Assume that for some $0 < c < C < \infty$ it holds $c \le f_0 \le C$ $\mathfrak{m}$-a.e. in X. Then the curve $t \mapsto \mu_t := f_t\mathfrak{m} \in \mathcal{M}_+(X)$ is absolutely continuous w.r.t. W_Ψ and it holds*

$$|\dot\mu_t| \le \frac{1}{A_\Phi(c)} \qquad \text{for a.e. } t \in (0,\infty),$$

where A_Φ is the character of Φ, defined in (1.5.6).

Proof. We start from the duality formula (1.7.3)

$$W_\Psi^{(s)}(\mu,\nu) = \sup_{\varphi\in \mathrm{Lip}_0(X,\mathsf{d})} \int_X Q_s\varphi\,d\nu - \int_X \varphi\,d\mu. \tag{3.3.5}$$

where $Q_t\varphi$ is defined in (1.6.1) and (1.6.2). Fix $\varphi \in \mathrm{Lip}_0(X,\mathsf{d})$ and recall (Theorem 1.6.5) that the map $t \mapsto Q_t\varphi$ is Lipschitz with values in $C(X)$, in particular also as a $L^2(X,\mathfrak{m})$-valued map.

Fix also $0 \le t < r$, set $\ell = (r-t)$ and recall that since (f_t) is a gradient flow of $\mathcal{F}^\Phi_v$ in $L^2(X,\mathfrak{m})$, the map $[0,\ell] \ni \tau \mapsto f_{t+\tau}$ is absolutely continuous with values in $L^2(X,\mathfrak{m})$.

Therefore, since both factors are uniformly bounded, the map $[0,\ell] \ni \tau \mapsto Q_{\frac{s\tau}{\ell}}\varphi f_{t+\tau}$ is absolutely continuous with values in $L^2(X,\mathfrak{m})$. In addition, the equality

$$\frac{Q_{\frac{s(\tau+h)}{\ell}}\varphi f_{t+\tau+h} - Q_{\frac{s\tau}{\ell}}\varphi f_{t+\tau}}{h} = f_{t+\tau}\frac{Q_{\frac{s(\tau+h)}{\ell}} - Q_{\frac{s\tau}{\ell}}\varphi}{h} + Q_{\frac{s(\tau+h)}{\ell}}\varphi\frac{f_{t+\tau+h} - f_{t+\tau}}{h},$$

together with the uniform continuity of $(x,\tau) \mapsto Q_{\frac{s\tau}{\ell}}\varphi(x)$ shows that the derivative of $\tau \mapsto Q_{\frac{s\tau}{\ell}}\varphi f_{t+\tau}$ can be computed via the Leibniz rule.

We have:

$$\begin{aligned}\int_X Q_s\varphi\,\mathrm{d}\mu_r - \int_X \varphi\,\mathrm{d}\mu_t &= \int Q_s\varphi f_{t+\ell}\,\mathrm{d}\mathfrak{m} - \int_X \varphi f_t\,\mathrm{d}\mathfrak{m} = \int_X\int_0^1 \frac{\mathrm{d}}{\mathrm{d}\tau}\big(Q_{s\tau}\varphi f_{t+\ell\tau}\big)d\tau\,\mathrm{d}\mathfrak{m}\\ &\leq \int_X\int_0^\ell \Big(-s\Phi\big(\mathrm{lip}_a(Q_{s\tau}\varphi)\big)f_{t+\ell\tau} + \ell Q_{s\tau}\varphi\Delta_\Phi f_{t+\ell\tau}\,\mathrm{d}\tau\Big)\,\mathrm{d}\mathfrak{m},\end{aligned} \tag{3.3.6}$$

having used Theorem 1.6.5.

Observe that by inequality (3.1.3) and Proposition 3.1.3 we have

$$\int_X Q_{s\tau}\varphi\Delta_\Phi f_{t+\ell\tau}\,\mathrm{d}\mathfrak{m} \leq \mathcal{F}^\Phi_v(Q_{s\tau}\varphi) \leq \|\mathrm{lip}_a(Q_{s\tau}\varphi)\|_{(\Phi),\mathfrak{m}}. \tag{3.3.7}$$

Plugging this inequality in (3.3.6), and taking $s = \frac{\ell}{A_\Phi(c)}$ we obtain

$$\begin{aligned}\int_X Q_s\varphi\,\mathrm{d}\mu_r - \int_X \varphi\,\mathrm{d}\mu_t \leq -\int_0^1\int_X s\Phi\big(\mathrm{lip}_a(Q_{s\tau}\varphi)\big)f_{t+\ell\tau}\,\mathrm{d}\mathfrak{m}\,\mathrm{d}\tau&\\ + \int_0^1 sA_\Phi(c)\|\mathrm{lip}_a(Q_{s\tau}\varphi)\|_{(\Phi),\mathfrak{m}}\,\mathrm{d}\tau&\end{aligned}$$

Now using the definition of $A_\Phi(c)$ we know that

$$A_\Phi(c)\|g\|_{(\Phi),\mathfrak{m}} \leq 1 + c\int_X \Phi(g)\,\mathrm{d}\mathfrak{m}. \tag{3.3.8}$$

Using this inequality with $g = \mathrm{lip}_a(Q_{s\tau}\varphi)$ in the end we get that

$$\int_X Q_s\varphi\,\mathrm{d}\mu_r - \int_X \varphi\,\mathrm{d}\mu_t \leq \int_0^1 s - \int_X s\Phi\big(\mathrm{lip}_a(Q_{s\tau}\varphi)\big)(f_{t+\ell\tau} - c)\,\mathrm{d}\mathfrak{m}\,\mathrm{d}\tau \leq s$$

This latter bound obviously doesn't depend on φ, so from (3.3.5) and (1.7.1) we deduce

$$W_\Psi(\mu_t,\mu_r) \leq \frac{(r-t)}{A_\Phi(c)}.$$

In particular, we showed that the curve μ_t is $\frac{1}{A_\Phi(c)}$-Lipschitz. □

We can now prove our main theorem:

Proof. [of Theorem 3.3.1] Recalling the results at the beginning of this Section, in order to conclude the proof we are only left to show that a bounded function of belonging to $BL^{1,\Phi}$, has finite Φ-relaxed energy, and that the two energies coincide; then Lemma 3.3.2 will give the equivalence for all functions.

We first assume that $\mathfrak{m}(X) < \infty$. First note that for every $a, b \in \mathbb{R}$ we have $\mathcal{F}_v^\Phi(a + bf) = |b|\mathcal{F}_v^\Phi(f)$ thanks to 1-homogeneity and Proposition 3.1.3, and the same is true for $\mathcal{F}_{BL}^\Phi$. Now take a function $f \in BL^{1,\Phi}$ such that $|f| \leq M$ and consider the scaled functions $f^\varepsilon = 1 + \frac{\varepsilon f}{M}$; we have that $1 - \varepsilon < f^\varepsilon < 1 + \varepsilon$. So, for any $0 < \varepsilon < 1$ we can put $f_0 = f^\varepsilon$ and consider the gradient flow f_t in $L^2(X, \mathfrak{m})$ with respect to $\mathcal{F}_v^\Phi$, starting from f_0. Let $\Theta(x) = x^2$ and use the energy dissipation estimate in Proposition 3.1.5; finally use Lemma 3.3.5 combined with Proposition 3.3.4 with $f_0 = f^\varepsilon$ to obtain:

$$\begin{aligned} 2\int_0^s \mathcal{F}_v^\Phi(f_t)\,\mathrm{d}t &= \int_X (f_0)^2\,\mathrm{d}\mathfrak{m} - \int_X (f_s)^2\,\mathrm{d}\mathfrak{m} \\ &\leq 2s \cdot \mathcal{F}_{BL}^\Phi(f_0) \cdot \frac{1+\varepsilon}{A_\Phi(1-\varepsilon)}. \end{aligned}$$

Now, knowing that $\mathcal{F}_v^\Phi(f_t)$ is nonincreasing in t we can say

$$s\mathcal{F}_v^\Phi(f_s) \leq \int_0^s \mathcal{F}_v^\Phi(f_t)\,\mathrm{d}t \leq s \cdot \mathcal{F}_{BL}^\Phi(f_0)\frac{1+\varepsilon}{A_\Phi(1-\varepsilon)}$$

and thus, first dividing by s, then letting $s \to 0$, taking the lower semicontinuity of $\mathcal{F}_v^\Phi$ into account we get

$$\mathcal{F}_v^\Phi(f^\varepsilon) \leq \mathcal{F}_{BL}^\Phi(f^\varepsilon)\frac{1+\varepsilon}{A_\Phi(1-\varepsilon)}.$$

Eventually we use that $\mathcal{F}_v^\Phi(f^\varepsilon) = \frac{\varepsilon}{M}\mathcal{F}_v^\Phi(f)$ and the same is true for $\mathcal{F}_{BL}^\Phi$, and then we divide by ε/M and let $\varepsilon \to 0$; now by Definition 1.5.6 we have that $A_\Phi(1-\varepsilon) \to 1$ and so

$$\mathcal{F}_v^\Phi(f) \leq \mathcal{F}_{BL}^\Phi(f),$$

which let us conclude.

Now let us consider a measure $\mathfrak{m}$ that is finite on bounded sets: using again Lemma 3.3.2 we need only to consider $f \in L^1 \cap L^\infty$. Let us fix a point $x \in X$ and consider χ_r as in Lemma 3.3.3. Then we consider the space $X_r = (\bar{B}_{r+4}, \mathsf{d}, \mathfrak{m}_r)$, where $\mathfrak{m}_r = \chi_{B_{r+3}}\mathfrak{m}$, and the function $f\chi_r$. It is obvious that we always have $\mathcal{F}_{BL}^{\Phi,X_r} \leq \mathcal{F}_{BL}^{\Phi,X}$ since a Ψ-plan in X_r is also a Ψ-plan in X.

The crucial point is that $\mathcal{F}_v^{\Phi,X_r}(f\chi_r) = \mathcal{F}_v^{\Phi,X}(f\chi_r)$; again it is obvious that $\mathcal{F}_v^{\Phi,X_r} \leq \mathcal{F}_v^{\Phi,X}$, since for every sequence of function in $\mathrm{Lip}_0(X,\mathsf{d})(X)$ we can recover a sequence of functions in $\mathrm{Lip}_0(X,\mathsf{d})(X_r)$ by restriction, and this latter sequence has less energy. But then, thanks to the last assertion in Lemma 3.3.3 we know that we can restrict the admissible sequence in the definition of $\mathcal{F}_v^{\Phi,X_r}(f\chi_r)$ to be supported in B_{r+2} and so, extending them to 0 outside B_{r+2}, they are admissible also in the definition of $\mathcal{F}_v^{\Phi,X}(f\chi_r)$, with the same energy.

Now we can prove that

$$\begin{aligned} \mathcal{F}_v^\Phi(f) &= \lim_{r\to\infty} \mathcal{F}_v^\Phi(f\chi_r) = \lim_{r\to\infty} \mathcal{F}_v^{\Phi,X_r}(\chi_r f) \\ &= \lim_{r\to\infty} \mathcal{F}_{BL}^{\Phi,X_r}(\chi_r f) \leq \lim_{r\to\infty} \mathcal{F}_{BL}^\Phi(\chi_r f) = \mathcal{F}_{BL}^\Phi(f). \end{aligned}$$

In order to remove the smoothness assumption on Ψ, we use Lemma 3.3.6 below, and so we can consider complementary couples $(\Phi_\varepsilon, \Psi_\varepsilon)$ sufficiently near to (Φ, Ψ), given by that lemma.

Using (3.3.10), we have $(1-\varepsilon)F_\Phi \leq F_{\Phi_\varepsilon} \leq F_\Phi$. In particular, taking the lower semicontinuous relaxation, we obtain

$$(1-\varepsilon)\mathcal{F}_v^\Phi \leq \mathcal{F}_v^{\Phi_\varepsilon} \leq \mathcal{F}_v^\Phi;$$

letting $\varepsilon \to 0$ we find that $\mathcal{F}_v^{\Phi_\varepsilon}(f) \to \mathcal{F}_v^\Phi(f)$, for all $f \in L^1(X, \mathfrak{m})$. Using again (3.3.10) it is clear that Ψ_ε-plans are also Ψ-plans and vice versa, and more precisely

$$(1-\varepsilon)\mathcal{F}_{BL}^\Phi \leq \mathcal{F}_{BL}^{\Phi_\varepsilon} \leq \mathcal{F}_{BL}^\Phi;$$

as before we get $\mathcal{F}_{BL}^{\Phi_\varepsilon} \to \mathcal{F}_{BL}^\Phi$ pointwise. It is now obvious that the equivalence for Φ_ε-Sobolev spaces extends to an equivalence for the Φ-Sobolev spaces. □

Lemma 3.3.6 *Let us consider an N-function Ψ. Then, for every $\varepsilon > 0$ there exists a complementary couple of N-function $(\Phi_\varepsilon, \Psi_\varepsilon)$, where Ψ_ε is of class C^1 and strictly convex, such that*

$$\Psi(x) \leq \Psi_\varepsilon(x) \leq \Psi\Big(\frac{x}{1-\varepsilon}\Big) \qquad \forall x \geq 0; \tag{3.3.9}$$

in particular, for every measure space (E,μ) we have $L^\Psi(\mu) = L^{\Psi_\varepsilon}(\mu)$ and $L^\Phi(\mu) = L^{\Phi_\varepsilon}(\mu)$; more precisely for every $f \in L^\Psi(E,\mu)$, $g \in L^\Phi(E,\mu)$ we have

$$\|f\|_{\Psi,\mu} \leq \|f\|_{\Psi_\varepsilon,\mu} \leq \frac{1}{1-\varepsilon}\|f\|_{\Psi,\mu} \qquad (1-\varepsilon)\|g\|_{(\Phi),\mu} \leq \|g\|_{(\Phi_\varepsilon),\mu} \leq \|g\|_{(\Phi),\mu} \tag{3.3.10}$$

Proof. We consider $\psi : [0,\infty) \to [0,\infty)$, the right derivative of Ψ. This is an increasing function, left continuous; the fact that Ψ is an N-function gives us that $\psi(0) = 0$ and that ψ is unbounded. In particular it is easy to see that it can be represented as

$$\psi(x) = c_0(x) + \sum_{i=0}^{\infty} c_i \rho_{x_i}(x), \tag{3.3.11}$$

where c_0 is a continuous function, c_i are positive real numbers whose sum is locally finite[1], and $\rho_y(x) = H(x-y)$, with H the Heaviside function

$$H(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Now we can represent also the right derivative of the function $\Psi(\frac{x}{1-\varepsilon})$:

$$\begin{aligned} \psi_1(x) := \frac{d}{dx}^+ \Psi\Big(\frac{x}{1-\varepsilon}\Big) &= \frac{1}{1-\varepsilon}\psi\Big(\frac{x}{1-\varepsilon}\Big) \\ &= \frac{1}{1-\varepsilon}c_0\Big(\frac{x}{1-\varepsilon}\Big) + \sum_{i=0}^{\infty} \frac{c_i}{1-\varepsilon}\rho_{x_i}\Big(\frac{x}{1-\varepsilon}\Big). \end{aligned}$$

Now let us consider, for every $i \geq 1$, $i \in \mathbb{N}$, the Lipschitz function ρ_i:

$$\rho_i(x) = \begin{cases} 0 & \text{if } x < x_i(1-\varepsilon) \\ \frac{x - x_i(1-\varepsilon)}{\varepsilon x_i} & \text{if } x_i(1-\varepsilon) \leq x \leq x_i \\ 1 & \text{if } x > x_i; \end{cases}$$

[1] for every $M > 0$ we have that $\sum_{x_i < M} c_i < \infty$

it is obvious that ρ_i are continuous functions such that

$$\rho_{x_i}(x) \leq \rho_i(x) \leq \rho_{x_i}\Big(\frac{x}{1-\varepsilon}\Big) \qquad \forall x \in [0,\infty). \tag{3.3.12}$$

In order to achieve the strict convexity, we have to add also a strictly increasing function. It is sufficient to consider the function

$$A(x) = \sum_{i=0}^{\infty} \varepsilon 2^{-i}\psi(2^{-i})\rho(x-2^{-i}),$$

where ρ is any continuous Heviside function, for example $\rho(x) = (x/(1+x))_+$. Then it is obvious that $\rho(x) \leq H(x)$ and so

$$0 \leq A(x) \leq \sum_{2^{-i}\leq x} \varepsilon 2^{-i}\psi(2^{-i}) \leq \varepsilon\psi(x). \tag{3.3.13}$$

Summing all, we consider $\psi_\varepsilon(x) = A(x) + c_0(x) + \sum_{i=0}^{\infty} c_i\rho_i(x)$; this is a strictly increasing and continuous function, and using (3.3.12), (3.3.13) and $1+\varepsilon \leq (1-\varepsilon)^{-1}$, we get

$$\psi(x) \leq \psi_\varepsilon(x) \leq \psi_1(x) \qquad \forall x \geq 0.$$

In particular, integrating this inequality, we deduce we can take $\Psi_\varepsilon(x) = \int_0^x \psi_\varepsilon(t)\,\mathrm{d}t$.

The inequalities for $\|\cdot\|_{\Psi_\varepsilon,\mu}$ are clear thanks to (3.3.9), while the ones for $\|\cdot\|_{(\Phi_\varepsilon),\mu}$ follow by duality, and the coincidence of the Orlicz spaces is then clear.

□

3.4 Consequences of the equivalence theorem

The first consequence we state about the equivalence theorem is that from the definition of $BL^{1,\Phi}$ we deduce stronger informations (see similar properties in the definition of BV functions in [5]), in particular we have that every $f \in BL^{1,\Phi}$ is BV along almost every curve.

If we don't add any other assumption on Φ, we can't expect to find anything better, in particular we can't expect $W^{1,1}$ regularity along Φ-almost every curve, and we can't expect any kind of modulus of gradient. This is shown in the Subsection 3.4.1 below.

Theorem 3.4.1 (Strong Beppo Levi property) *A function $f \in L^1(X,\mathfrak{m})$ belongs to the Φ-Beppo Levi space $BL^{1,\Phi}(X,\mathsf{d},\mathfrak{m})$ if and only if:*

- *for Φ-a.e. curve γ we have $f\circ\gamma \in BV(0,1)$ and 0 and 1 are approximate continuity points for $f\circ\gamma$; in particular*

$$|f(\gamma_1) - f(\gamma_0)| \leq |D(f\circ\gamma)|(0,1) \qquad \textit{for } \Phi\textit{-a.e. curve } \gamma;$$

- *there exists a constant $E>0$ for any Ψ-plan $\boldsymbol{\pi}$ the following inequality holds*

$$\int |D(f\circ\gamma)|(0,1)\,\mathrm{d}\boldsymbol{\pi} \leq E\cdot C(\boldsymbol{\pi})\int_0^1 \||\dot\gamma_t|\|_{\Psi,\boldsymbol{\pi}}\,\mathrm{d}t. \tag{3.4.1}$$

The least constant E for which (3.4.1) *holds is exactly* $\mathcal{F}^{\Phi}_{BL}(f)$.

Proof. Let us suppose we have a function f that satisfies the two assumption of the theorem, with a constant $E = E(f)$; combining the two inequalities one gets easily that $f \in BL^{1,\Phi}$ and that $E(f)$ is a good constant for (3.2.5) so that $E(f) \geq \mathcal{F}^{\Phi}_{BL}(f)$.

Now let us suppose that $f \in BL^{1,\Phi}$; thanks to the equivalence theorem we have a sequence of bounded Lipschitz functions (f_n) such that $f_n \to f$ in $L^1(X, \mathfrak{m})$ and such that $\lim \|\mathrm{lip}_a(f)\|_{(\Phi),\mathfrak{m}} = \mathcal{F}^{\Phi}_{BL}(f)$. We can assume this sequence is "fast converging", that is, we have $\sum_n \|f - f_n\|_1 < \infty$. Now, calling $E_n(\gamma) = \|f \circ \gamma - f_n \circ \gamma\|_1$ and $E = \sum_n E_n$:

$$\int E(\gamma)\,\mathrm{d}\boldsymbol{\pi} = \iint_0^1 \sum_{n=1}^{\infty} |f(\gamma_t) - f_n(\gamma_t)|\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi} \leq C(\boldsymbol{\pi}) \sum_{n=1}^{\infty} \int_X |f - f_n|\,\mathrm{d}\mathfrak{m} < \infty;$$

and so we have that $E < \infty$ for $\boldsymbol{\pi}$-almost every curve γ. For those curves we have $E_n(\gamma) \to 0$ and so $f \circ \gamma \to f_n \circ \gamma$ in $L^1(0,1)$; by the lower semicontinuity of the total variation we have $\liminf_n |D(f_n \circ \gamma)|(0,1) \geq |D(f \circ \gamma)|(0,1)$. Now, using Beppo Levi, and exploiting the upper gradient property of the asymptotic Lipschitz constant we get

$$\begin{aligned}
\int |D(f \circ \gamma)|(0,1)\,\mathrm{d}\boldsymbol{\pi} &\leq \liminf_{n\to\infty} \int |D(f_n \circ \gamma)|(0,1)\,\mathrm{d}\boldsymbol{\pi} \\
&\leq \liminf_{n\to\infty} \iint_0^1 \mathrm{lip}_a(f_n, \gamma_t)|\dot{\gamma}_t|\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi} \\
&\leq \liminf_{n\to\infty} \|\mathrm{lip}_a(f_n)\|_{(\Phi),\mathfrak{m}} \cdot C(\boldsymbol{\pi}) \int_0^1 \|\dot{\gamma}_t\|_{\Psi,\boldsymbol{\pi}}\,\mathrm{d}t \\
&= \mathcal{F}^{\Phi}_{BL}(f) \cdot C(\boldsymbol{\pi}) \int_0^1 \|\dot{\gamma}_t\|_{\Psi,\boldsymbol{\pi}}\,\mathrm{d}t,
\end{aligned}$$

in particular $E(f) \leq \mathcal{F}^{\Phi}_{BL}(f)$. We are left to show the approximate continuity property: with a very similar calculation as before we can estimate $|f_n(\gamma_s) - f_n(\gamma_0)|$ with the upper gradient property and use that $f_n(\gamma_t) \to f(\gamma_t)$ for $r = s, 0$ for Φ-almost every curve γ, to conclude that

$$\begin{aligned}
\int |f(\gamma_s) - f(\gamma_0)|\,\mathrm{d}\boldsymbol{\pi} &= \lim_{n\to\infty} \int |f_n(\gamma_s) - f_n(\gamma_0)|\,\mathrm{d}\boldsymbol{\pi} \\
&\leq \mathcal{F}^{\Phi}_v(f) \cdot C(\boldsymbol{\pi}) \int_0^s \|\dot{\gamma}_t\|_{\Psi,\boldsymbol{\pi}}\,\mathrm{d}t;
\end{aligned}$$

integrating this inequality in the s variable from 0 to s_0, and then dividing by s_0 we get

$$\int \frac{1}{s_0} \int_0^{s_0} |f(\gamma_s) - f(\gamma_0)|\,\mathrm{d}s\,\mathrm{d}\boldsymbol{\pi} \leq \mathcal{F}^{\Phi}_v(f) \cdot C(\boldsymbol{\pi}) \int_0^{s_0} \|\dot{\gamma}_s\|_{\Psi,\boldsymbol{\pi}}\,\mathrm{d}s \to 0 \qquad \text{as } s_0 \to 0;$$

in particular, letting $H_t(g) = \fint_0^t |g(s) - g(0)|\,\mathrm{d}s$, we get that $\liminf_{t\to 0} H_t(f \circ \gamma) = 0$ for $\boldsymbol{\pi}$-almost every curve γ, but suffices to imply that 0 is a point of approximate continuity since we know also that $f \circ \gamma \in BV$; in this case in fact we know that H_t has always limit as $t \to 0$, and this limit is equal to $|g(0) - \tilde{g}(0)|$.

The same reasoning applies also to conclude that for any $t \in [0,1]$, we have that $f(\gamma_t)$ coincides with the precise representative for Φ-almost every curve γ, that is

$$f(\gamma_t) = \lim_{\varepsilon\to 0} \fint_{t-\varepsilon}^{t+\varepsilon} f(\gamma_s)\,\mathrm{d}s \qquad \text{for } \Phi\text{-almost every curve } \gamma.$$

In particular also $t = 1$ is a point of approximate continuity for Φ-a.e. curve γ. $\square$

3.4.1 Example of Φ-Sobolev function that is not absolutely continuous along almost every curve

Let us consider $\mathbb{R}$ with the euclidean norm $|\cdot|_e$ and define recursively

$$\begin{cases} a_1 = 0 \\ a_n = a_{n-1} + \frac{ln(n)}{n^2} & \text{if } n > 1. \end{cases}$$

Denote also $a = \lim_{n\to\infty} a_n < \infty$; define $A_n = [a_{n-1}, a_n)$ and

$$\mathfrak{m} = \mathscr{L}^1|_{[a,a+1]} + \sum_{n=2}^{\infty} \frac{1}{ln(n)} \mathscr{L}^1|_{A_n};$$

the singularity of this measure we are interested in is at the point $x = a$, where the density is decreasing from the left; notice also that $\mathfrak{m}(\mathbb{R}) = \pi^2/6$. Let us consider

$$\Phi(t) = (t+1)\ln(t+1) - t \qquad \text{and} \qquad \Psi(t) = e^t - t - 1,$$

which are easily seen to be a pair of complementary N-functions. We will see that for functions in $BL^{1,\Phi}(\mathbb{R}, |\cdot|_e, \mathfrak{m})$ we can't go beyond BV regularity along curves, proved in Theorem 3.4.1.

Proposition 3.4.2 *Let a, $\mathfrak{m}$ defined as before. Let us consider $f(x) = \chi_{[a,\infty)}$. then*

(i) $f \in H^{1,\Phi}(\mathbb{R}, |\cdot|_e, \mathfrak{m})$;

(ii) there is a set of curves Γ and a Ψ-plan $\boldsymbol{\pi}$ such that $\boldsymbol{\pi}(\Gamma) > 0$ and $f \circ \gamma \notin W^{1,1}(0,1)$ for all $\gamma \in \Gamma$.

Proof. In order to prove (i) we will explicitly find a sequence (f_n) approximating f: let

$$f_n(x) = \begin{cases} 0 & \text{if } x < a_n \\ \frac{x - a_n}{a_{n+1} - a_n} & \text{if } x \in A_{n+1} \\ 1 & \text{if } x \geq a_{n+1}; \end{cases}$$

then (f_n) is a sequence of Lipschitz function[2] such that $\mathrm{lip}_a(f_n) = |f_n'| = \frac{n^2}{\ln(n)}\chi_{A_n}$; furthermore $f_n \to f$ in $L^1(\mathbb{R}, \mathfrak{m})$ and we can compute $|||f_n'|||_{(\Phi)}$ since it is the norm of a characteristic function (see [72], Example 9, Section 1.2):

$$|||f_n'|||_{(\Phi)} = \frac{n^2}{ln(n)} \|\chi_{A_n}\|_{(\Phi)} = \frac{n^2}{\ln(n)} \mathfrak{m}(A_n) \Psi^{-1}\Big(\frac{1}{\mathfrak{m}(A_n)}\Big).$$

Now for $t > 2$ we have $\Psi(t) \geq \frac{1}{2}e^t$ and so, for n big enough, we can estimate $\Psi^{-1}(1/\mathfrak{m}(A_n)) \leq \ln(2) - \ln(\mathfrak{m}(A_n)) = \ln(2) + 2\ln(n)$, getting

$$\mathcal{F}_v^\Phi(f) \leq \liminf_{n\to\infty} |||f_n'|||_{(\Phi)} \leq \liminf_{n\to\infty} \frac{n^2}{\ln(n)} \cdot \frac{\ln(2) + 2\ln(n)}{n^2} = 2,$$

[2] The bounded support hypothesis can be easily dropped when $\mathfrak{m}$ has bounded support; in fact it is sufficient to consider a 1-Lipschitz function with bounded support g such that $g = 1$ on a neighborhood of the support of $\mathfrak{m}$ and then for every Lipschitz function ρ we have that $g\rho$ has bounded support, $\|\mathrm{lip}_a(g\rho)\|_{(\Phi),\mathfrak{m}} = \|\mathrm{lip}_a(\rho)\|_{(\Phi),\mathfrak{m}}$ and $\|f - \rho\|_1 = \|f - g\rho\|_1$ for every $f \in L^1(X, \mathfrak{m})$.

so that $f \in H^{1,\Phi}(\mathbb{R}, |\cdot|_e, \mathfrak{m})$.

In order to achieve (ii) we have to consider curves that meets a in their interior: let $T(t)$ be the map that maps monotonically $\mathscr{L}^1|_{[0,\pi^2/6]} = \mathfrak{m}_1$ to $\mathfrak{m}$, in particular defining $b_1 = 0$ and $b_n = b_{n-1} + \frac{1}{n^2}$ for $n > 1$, and B_n defined similarly to how we defined A_n before, we have that T maps B_n to A_n linearly, and here $T' = \ln(n)$. Let $F : [0,1] \to C([0,1];\mathbb{R})$ be the map defined by $F(t)(s) = T(s+t)$; now let us consider $\boldsymbol{\pi} = F_\sharp(2\mathscr{L}^1|_{[0,1/2]})$. We have to verify this is a Ψ-plan.

- $(\mathrm{e}_t)_\sharp \boldsymbol{\pi} = T_\sharp(2\mathscr{L}^1|_{[t,t+1/2]}) \le T_\sharp(2\mathfrak{m}_1) = 2\mathfrak{m}$, and so we can take $C(\boldsymbol{\pi}) = 2$;
- $\| |\gamma'(t)| \|_{\Psi,\boldsymbol{\pi}} = \|T'(t+\cdot)\|_{\Psi,2\mathscr{L}^1|_{[0,1/2]}} \le \|T'\|_{\Psi,2\mathfrak{m}_1}$, and we have that $T' \in L^\Psi(2\mathfrak{m}_1)$ since

$$2\int \Psi\Big(\frac{T'}{2}\Big)\,\mathrm{d}\mathfrak{m}_1 \le 2\int e^{\frac{T'}{2}}\,\mathrm{d}\mathfrak{m}_1 = 2\sum_{n=2}^{\infty} \frac{n^{1/2}}{n^2} < \infty.$$

In particular $\boldsymbol{\pi}$ is a Ψ-plan but for $t < \pi^2/6 - 1$ we have that $F(t)(s) = a$ for $s \in (0,1)$ and so $f \circ F(t)$ is a step function, and $f \circ F(t) \notin W^{1,1}$. So we can take $\boldsymbol{\pi}$ as the Ψ-plan for (ii) and $\Gamma = \{F(t) \;:\; 0 \le t < \pi^2/6 - 1\}$ as the bad set; it is obvious that $\boldsymbol{\pi}(\Gamma) = \pi^2/6 - 1 > 0$ and for the reasoning above we have $f \circ \gamma \notin W^{1,1}(0,1)$ for every $\gamma \in \Gamma$.

$\square$

3.4.2 Ψ doubling: existence of the gradient

In this section we assume Ψ to be doubling. By the results recalled in Section 1.5 in this case we have that $L^\Psi(X,\mathfrak{m}) = M^\Phi(X,\mathfrak{m})$ and so it is separable and we have also $(L^\Psi)^* = L^\Phi$; in particular we can consider the weak-$*$ topology $\sigma(L^\Phi, L^\Psi)$, and we know that in every ball this topology is metrizable and moreover closed balls are compact sets. In particular we will use several times this result:

Lemma 3.4.3 *Let $(f_n) \subset L^\Phi(X,\mathfrak{m})$ be a sequence such that*

$$\liminf_{n\to\infty} \|f_n\|_{(\Phi),\mathfrak{m}} < \infty;$$

then there exists a subsequence (not relabeled for convenience) and a function $f \in L^\Phi(X,\mathfrak{m})$ such that

$$\int_X f_n g \,\mathrm{d}\mathfrak{m} \to \int_X f g\,\mathrm{d}\mathfrak{m} \qquad \text{for all } g \in L^\Psi(X,\mathfrak{m});$$

we have also $\|f\|_{(\Phi),\mathfrak{m}} \le \liminf_{n\to\infty} \|f_n\|_{(\Phi),\mathfrak{m}}$.

Moreover if we have $(f_n) \subset L^\infty(X,\mathfrak{m}) \cap L^1(X,\mathfrak{m})$ and $f \ge 0$, there exist convex combinations $\hat{f}_n$ of (f_n) and a sequence $(h_n) \subset L^\Phi(X,\mathfrak{m})$ such that

$$\hat{f}_n \le h_n \quad \forall n \in \mathbb{N} \qquad \text{and} \qquad h_n \to f \text{ strongly in } L^\Phi(X,\mathfrak{m}).$$

This can be seen as a sort of weak-$*$ Mazur lemma.

Proof. We know that L^Ψ is separable and that $(L^\Psi)^* = L^\Phi$. In particular the weak-$*$ topology on balls of L^Φ is metrizable and compact. Given the hypothesis, this gives us a subsequence of f_n converging weakly-$*$ to some $f \in L^\Phi(X, \mathfrak{m})$, and we have lower semicontinuity for the norm.

For the second part, let us consider a bounded set B. Since $\mathfrak{m}(B) < \infty$ we have that

$$L^\Phi(B, \mathfrak{m}) \subset L^1(B, \mathfrak{m}) \qquad \text{and} \qquad L^\infty(B, \mathfrak{m}) \subset L^\Psi(X, \mathfrak{m});$$

in particular the weak-$*$ convergence in $L^\Phi(B, \mathfrak{m})$ implies that $f_n \rightharpoonup f$ weakly in $L^1(B)$, and thanks to Mazur lemma there exist convex combinations $\tilde{f}_n$ of f_n that converge strongly to f in $L^1(B)$. By a diagonal argument we can assume that this is true for every bounded set B. Moreover we have $\sup_{n\in\mathbb{N}} \|\tilde{f}_n\|_{(\Phi),\mathfrak{m}} \le \sup_{n\in\mathbb{N}} \|f_n\|_{(\Phi),\mathfrak{m}}$.
Since $\tilde{f}_n \in L^\infty(X, \mathfrak{m}) \cap L^1(X, \mathfrak{m})$ we have also $\tilde{f}_n \in M^\Phi(X, \mathfrak{m})$. Let us call $k_n = \sup\{\tilde{f}_n, f\} - f$. It is clear that $0 \le k_n \le \tilde{f}_n$ and so $k_n \in M^\Phi(X, \mathfrak{m})$, but we have also that $\|k_n\|_{1,B} \le \|\tilde{f}_n - f\|_{1,B} \to 0$ for every bounded set and in particular we have

$$\int_X k_n g \, \mathrm{d}\mathfrak{m} \to 0 \qquad \text{for all } g \in L^\infty(X, \mathfrak{m}) \text{ with bounded support}$$

using the fact that k_n are bounded in (Φ)-norm and that $\overline{\mathrm{Lip}_0(X, \mathsf{d})}^{\|\cdot\|_\Psi} = M^\Psi = L^\Psi$ (since Ψ is doubling), we can conclude that $k_n \rightharpoonup 0$ in $M^\Phi(X, \mathfrak{m})$ (we recall that $(M^\Phi)^* = L^\Psi$).
Applying again Mazur lemma (this time in M^Φ) we can find convex combination $\hat{k}_n$ of k_n such that $\hat{k}_n \to 0$ strongly in M^Φ and so also strongly in L^Φ. Now taking convex combination of the inequalities $k_n \ge \tilde{f}_n - f$ we get convex combination $\hat{f}_n$ of f_n such that $\hat{k}_n \ge \hat{f}_n - f$. Consider $h_n = \hat{k}_n + f$ and we get the thesis. □

Now we are ready to define the weak gradient: in order to simplify the arguments we present only the gradients in the Beppo-Levi context.

Definition 3.4.4 (Φ-weak upper gradients) *A Borel function $g : X \to [0, \infty]$ is a Φ-weak upper gradient of $f : X \to \mathbb{R}$ if*

$$\left| \int_{\partial\gamma} f \right| \le \int_\gamma g < \infty \qquad \text{for } \Phi\text{-a.e. } \gamma. \tag{3.4.2}$$

Definition 3.4.5 (Sobolev functions along Φ-a.e. curve) *A function $f : X \to \mathbb{R}$ is Sobolev along Φ-a.e. curve if for Φ-a.e. curve γ the function $f \circ \gamma$ coincides a.e. in $[0, 1]$ and in $\{0, 1\}$ with an absolutely continuous map $f_\gamma : [0, 1] \to \mathbb{R}$.*

By Remark 3.2.2 applied to $\mathcal{T} := \{0, 1\}$, (3.4.2) does not depend on the particular representative of f in the class of $\mathfrak{m}$-measurable functions coinciding with f up to a $\mathfrak{m}$-negligible set. The same Remark also shows that the property of being Sobolev along Φ-q.e. curve γ is independent of the representative in the class of $\mathfrak{m}$-measurable functions coinciding with f $\mathfrak{m}$-a.e. in X.

In the next proposition, based on Lemma 1.3.3, we prove that the existence of a Φ-weak upper gradient g implies Sobolev regularity along Φ-a.e. curve.

Proposition 3.4.6 *Let $f : X \to \mathbb{R}$ be $\mathfrak{m}$-measurable, and let g be a Φ-weak upper gradient of f. Then f is Sobolev along Φ-a.e. curve.*

Proof. Notice that if $\boldsymbol{\pi}$ is a Ψ-test plan, so is $(\mathrm{restr}_t^s)_\sharp\boldsymbol{\pi}$. Hence if g is a Φ-weak upper gradient of f such that $\int_\gamma g < \infty$ for Φ-a.e. γ, then for every $t < s$ in $[0,1]$ it holds

$$|f(\gamma_s) - f(\gamma_t)| \le \int_t^s g(\gamma_r)|\dot\gamma_r|\,\mathrm{d}r \qquad \text{for } \Phi\text{-a.e. } \gamma.$$

Let $\boldsymbol{\pi}$ be a Ψ-test plan: by Fubini's theorem applied to the product measure $\mathscr{L}^2 \times \boldsymbol{\pi}$ in $(0,1)^2 \times C([0,1];X)$, it follows that for $\boldsymbol{\pi}$-a.e. γ the function f satisfies

$$|f(\gamma_s) - f(\gamma_t)| \le \left|\int_t^s g(\gamma_r)|\dot\gamma_r|\,\mathrm{d}r\right| \qquad \text{for } \mathscr{L}^2\text{-a.e. } (t,s) \in (0,1)^2.$$

An analogous argument shows that for $\boldsymbol{\pi}$-a.e. γ

$$\begin{cases} |f(\gamma_s) - f(\gamma_0)| \le \int_0^s g(\gamma_r)|\dot\gamma_r|\,\mathrm{d}r \\ |f(\gamma_1) - f(\gamma_s)| \le \int_s^1 g(\gamma_r)|\dot\gamma_r|\,\mathrm{d}r \end{cases} \qquad \text{for } \mathscr{L}^1\text{-a.e. } s \in (0,1). \tag{3.4.3}$$

Since $g \circ \gamma|\dot\gamma| \in L^1(0,1)$ for $\boldsymbol{\pi}$-a.e. γ, by Lemma 1.3.3 it follows that $f \circ \gamma \in W^{1,1}(0,1)$ for $\boldsymbol{\pi}$-a.e. γ, and

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}(f\circ\gamma)\right| \le g\circ\gamma|\dot\gamma| \quad \text{a.e. in } (0,1), \text{ for } \boldsymbol{\pi}\text{-a.e. } \gamma. \tag{3.4.4}$$

Since $\boldsymbol{\pi}$ is arbitrary, we conclude that $f \circ \gamma \in W^{1,1}(0,1)$ for Φ-a.e. γ, and therefore it admits an absolutely continuous representative f_γ; moreover, by (3.4.3), it is immediate to check that $f(\gamma_t) = f_\gamma(t)$ for $t \in \{0,1\}$ and Φ-a.e. γ. □

The last statement of the proof above and (3.4.4) yield the following

$$g_i,\ i=1,2,\ \ \Phi\text{-weak upper gradients of } f \quad \Longrightarrow \quad \min\{g_1,g_2\}\ \Phi\text{-weak upper gradient of } f. \tag{3.4.5}$$

Using this stability property we can recover a distinguished minimal object.

Definition 3.4.7 (Minimal Φ-weak upper gradient) *Let $f : X \to \mathbb{R}$ be a $\mathfrak{m}$-measurable function having at least a Φ-weak upper gradient $g_0 : X \to [0,\infty]$. The minimal Φ-weak upper gradient $|\nabla f|_{w,\Phi}$ of f is the Φ-weak upper gradient characterized, up to $\mathfrak{m}$-negligible sets, by the property*

$$|\nabla f|_{w,\Phi} \le g \qquad \mathfrak{m}\text{-a.e. in } X, \text{ for every } \Phi\text{-weak upper gradient } g \text{ of } f. \tag{3.4.6}$$

We will refer to it also as the Φ-weak gradient of f.

Uniqueness of the minimal weak upper gradient is obvious. For existence, let $\theta : X \to (0,\infty)$ be a $\mathfrak{m}$-integrable function (the existence of such θ is granted since $\mathfrak{m}$ is σ-finite), then we can prove $|\nabla f|_{w,\Phi} := \inf_n g_n$, where g_n are Φ-weak upper gradients which provide a minimizing sequence in

$$\inf\left\{\int_X \theta\tan^{-1} g\,\mathrm{d}\mathfrak{m} :\ g \le g_0 \text{ is a } \Phi\text{-weak upper gradient of } f\right\}.$$

We immediately see, thanks to (3.4.5), that we can assume with no loss of generality that $g_{n+1} \le g_n$. Hence, applying (3.4.2) to g_n and by monotone convergence, the function $|\nabla f|_{w,\Phi}$ is a Φ-weak upper gradient of f and $\int_X \theta\tan^{-1} g\,\mathrm{d}\mathfrak{m}$ is minimal at $g = |\nabla f|_{w,\Phi}$. This minimality, in conjunction with (3.4.5), gives (3.4.6). Now we are ready to state the main result of this section.

Theorem 3.4.8 *Let $f \in L^1(X, \mathfrak{m})$; then the following are equivalent*

(i) $f \in H_v^{1,\Phi}(X, \mathsf{d}, \mathfrak{m})$*;*

(ii) there exists a function $g \in L^\Phi(X, \mathfrak{m})$ that is a Φ-weak upper gradient for f. We have also $\||\nabla f|_{w,\Phi}\|_{(\Phi),\mathfrak{m}} = \mathcal{F}_v^\Phi(f)$.

(iii) There exist $g \in L^\Phi(X, \mathfrak{m})$ and a sequence $(f_n) \subset \mathrm{Lip}_0(X, \mathsf{d})$ such that

$$f_n \to f \text{ strongly in } L^1(X, \mathfrak{m}), \qquad \mathrm{lip}_a(f_n) \stackrel{*}{\rightharpoonup} g \text{ weakly-}* \text{ in } L^\Phi(X, \mathfrak{m}). \tag{3.4.7}$$

(iv) There exist $g, g_n \in L^\Phi(X, \mathfrak{m})$ and a sequence $(f_n) \subset \mathrm{Lip}_0(X, \mathsf{d})$ such that $g_n \geq \mathrm{lip}_a(f_n)$ and

$$f_n \to f \text{ strongly in } L^1(X, \mathfrak{m}), \qquad g_n \to g \text{ strongly in } L^\Phi(X, \mathfrak{m}). \tag{3.4.8}$$

Moreover, every g in (iii) (or (iv)) is also a Φ-weak upper gradient for f and, conversely, $|\nabla f|_{w,\Phi}$ satisfies (iii) and (iv).

We observe that (i) doesn't always imply (ii), (iii) or (iv) if we don't require that Ψ is doubling. In fact (iii) or (iv) imply (ii) and in turn, (ii) implies that f is Sobolev along Φ-almost every curve (Proposition 3.4.6), but this is not true always, as we proved in Subsection 3.4.1.

Proof. (ii) $\Rightarrow$ (i): Suppose f has a Φ-weak upper gradient g; then integrating (3.4.2) with respect to a Ψ-plan $\boldsymbol{\pi}$ we obtain (3.2.3), that as usual reduces to (3.2.5) with $E = \|g\|_{(\Phi),\mathfrak{m}}$. This shows that $f \in BL^{1,\Phi}$ and minimizing in g we get

$$\mathcal{F}_v^\Phi(f) = \mathcal{F}_{BL}^\Phi(f) \leq \||\nabla f|_{w,\Phi}\|_{(\Phi)}. \tag{3.4.9}$$

(i) $\Rightarrow$ (iii) Let us suppose now that $f \in H_v^{1,\Phi}$, so that there exists a sequence of Lipschitz functions (f_n) such that $\lim_n \|\mathrm{lip}_a(f_n)\|_{(\Phi)} = \mathcal{F}_v^\Phi(f)$. Using Lemma 3.4.3, up to subsequences we get the existence of a function $g \in L^\Phi$ such that $\mathrm{lip}_a(f_n) \stackrel{*}{\rightharpoonup} g$. By lower semicontinuity of the norm with respect to weak convergence we have that

$$\|g\|_{(\Phi)} \leq \lim_{n\to\infty} \|\mathrm{lip}_a(f_n)\|_{(\Phi)} = \mathcal{F}_v^\Phi(f). \tag{3.4.10}$$

(iii) $\Rightarrow$ (ii) We can use (3.2.3) with f_n and $\mathrm{lip}_a(f_n)$ and pass to the limit as $n \to \infty$ to get[3]:

$$\int |f(\gamma_1) - f(\gamma_0)| \,\mathrm{d}\boldsymbol{\pi} \leq \iint_0^1 g(\gamma_t)|\dot{\gamma}_t| \,\mathrm{d}t \,\mathrm{d}\boldsymbol{\pi} < \infty \qquad \text{for every } \Psi\text{-plan } \boldsymbol{\pi}.$$

Now we use the fact that $\boldsymbol{\pi}|_A$ is still a Ψ-plan so that we can localize the inequality to get

$$|f(\gamma_1) - f(\gamma_0)| \leq \int_\gamma g < \infty \qquad \text{for } \Phi\text{-q.e. curve } \gamma,$$

[3]It is sufficient to check this inequality when $\boldsymbol{\pi}$ is supported on curves contained in a bounded set. Therefore, up to restricting to a smaller set, we can assume that X is bounded and so $\mathfrak{m}(X) < \infty$. In this case we note that $f_n \stackrel{*}{\rightharpoonup} f$ in $L^\Phi(X, \mathfrak{m})$ iff $f_n \rightharpoonup f$ in $L^1(X, \mathfrak{m})$ and $\|f_n\|_{(\Phi),\mathfrak{m}}$ is equibounded (this is because $L^\infty \cap L^\Psi$ is strongly dense in L^Ψ). In particular we have that if $\mu \leq C\mathfrak{m}$ then weak-$*$ convergence in $L^\Phi(X, \mathfrak{m})$ implies weak-$*$ convergence in $L^\Phi(X, \mu)$; we use this observation with $\mathfrak{m}$ and $(\mathrm{e}_t)_\sharp \boldsymbol{\pi}$

and so, by definition, we have that g is a Φ-weak upper gradient for f and thus we have

$$\||\nabla f|_{w,\Phi}\|_{(\Phi)} \le \|g\|_{(\Phi)}. \tag{3.4.11}$$

Using the equivalence theorem along with (3.4.9), (3.4.10), (3.4.11) we conclude $\||\nabla f|_{w,\Phi}\|_{(\Phi),\mathfrak{m}} = \mathcal{F}_v^{\Phi}(f)$ and that the function g with minimal norm that satisfies *(ii)* coincides with $|\nabla f|_{w,\Phi}$. Eventually, it is easy to see that (iv) implies (i), while if $\mathfrak{m}$ is finite, using Lemma 3.4.3 we conclude that (iii) (with $g = |\nabla f|_{w,\Phi}$) implies (iv) with the same g.

□

We have now defined a distinguished object as a gradient. Now we can show the strong locality property and chain rule for this gradient. We follow Proposition 4.8 in [11], but with the Φ-weak gradient definition:

Proposition 3.4.9 (Locality and chain rule) *If $f \in L^1(X, \mathfrak{m})$ has a Φ-weak upper gradient, the following properties hold:*

(a) for any Lipschitz function h on an interval J containing the image of f we have that $h(f)$ has a Φ-weak upper gradient and $|\nabla h(f)|_{w,\Phi} = |h'(f)||\nabla f|_{w,\Phi}$.

(b) for any $\mathscr{L}^1$-negligible Borel set $N \subset \mathbb{R}$ it holds $|\nabla f|_{w,\Phi} = 0$ $\mathfrak{m}$-a.e. on $f^{-1}(N)$;

(c) $|\nabla f|_{w,\Phi} = |\nabla g|_{w,\Phi}$ $\mathfrak{m}$-a.e. on $\{f = g\}$ for every $g \in L^1(X, \mathfrak{m})$ that has a Φ-weak upper gradient.

Proof. Let us first prove $|\nabla h(f)|_{w,\Phi} \le |h'(f)||\nabla f|_{w,\Phi}$ for $h \in C^1$. Recall by Proposition 3.4.6 we know that $f \circ \gamma \in W^{1,1}(0,1)$, for Φ-a.e. curve γ; for those γ we have that $h(f \circ \gamma) \in W^{1,1}(0,1)$ and in particular its weak derivative is $h'(f \circ \gamma) \cdot (f \circ \gamma)'$. Now multiplying (3.4.4) with $g = |\nabla f|_{w,\Phi}$ by $|h'(f(\gamma_t))|$ and integrating we get

$$|h(f(\gamma_1)) - h(f(\gamma_0))| \le \int_0^1 |h(f(\gamma))'|(t)\,\mathrm{d}t \le \int_\gamma |h'(f)||\nabla f|_{w,\Phi}\,\mathrm{d}s \qquad \text{for } \Phi\text{-a.e. } \gamma;$$

so by definition we have that $|h'(f)||\nabla f|_{w,\Phi}$ is a Φ-weak upper gradient for $h(f)$, and in particular by the point wise minimality property of the weak gradient we get

$$|\nabla h(f)|_{w,\Phi} \le |h'(f)||\nabla f|_{w,\Phi}. \tag{3.4.12}$$

(b) First assume that N is compact. Then there exists open sets $A_n \subset \mathbb{R}$ such that $A_n \downarrow N$ and $\mathscr{L}^1(A_1) < \infty$. Also, let $k_n : \mathbb{R} \to [0,1]$ be continuous function satisfying $\chi_N \le k_n \le \chi_{A_n}$, and define

$$\begin{cases} h_n(0) = 0 \\ h_n'(x) = 1 - k_n(x) \end{cases}$$

The sequence (h_n) uniformly converges to the identity map, and each h_n is 1-Lipschitz and C^1. Therefore $h_n(f)$ converge to f in L^1. Taking into account that $h_n' = 0$ on N and (3.4.12) we deduce

$$\begin{aligned} \||\nabla f|_{w,\Phi}\|_{(\Phi)} = \mathcal{F}_v^{\Phi}(f) &\le \liminf_{n\to\infty} \mathcal{F}_v^{\Phi}(h_n(f)) \le \liminf_{n\to\infty} \||h_n'(f)|\,|\nabla f|_{w,\Phi}\|_{(\Phi)} \\ &= \liminf_{n\to\infty} \|\chi_{X_N}|h_n'(f)|\,|\nabla f|_{w,\Phi}\|_{(\Phi)} \le \|\chi_{X_N}|\nabla f|_{w,\Phi}\|_{(\Phi)}, \end{aligned}$$

where we put $X_N = X \setminus f^{-1}(N)$. It remains to deal with the case when N is not compact. In this case we can consider the finite measure $\mu := f_\sharp \mathfrak{m}$; then there exists an increasing sequence (K_n) of compact subsets of N such that $\mu(K_n) \uparrow \mu(N)$. By the result for the compact set we know that $|\nabla f|_{w,\Phi} = 0$ $\mathfrak{m}$-a.e. on $\cup_n f^{-1}(K_n) = H$, and by definition of push forward we get $\mathfrak{m}(f^{-1}(N) \setminus H)) = 0$.

(a) If h is Lipschitz we know by Rademacher theorem that h' exists $\mathscr{L}^1$-a.e. in $\mathbb{R}$, so we can do the same proof for C^1, paying attention to the fact that $h'(f)|\nabla f|_{w,\Phi}$, and the other expressions where h' is present, are well defined thanks to (b). In order to prove the equality we can suppose h is 1-Lipschitz, and so we have that $(1-h'(f))|\nabla f|_{w,\Phi}$ and $(1+h'(f))|\nabla f|_{w,\Phi}$ are Φ-weak upper gradient of $f - h(f)$ and $f + h(f)$ respectively. Now using the subadditivity of the weak gradient:

$$\begin{aligned} 2|\nabla f|_{w,\Phi} &\leq |\nabla(f - h(f))|_{w,\Phi} + |\nabla(f + h(f))|_{w,\Phi} \\ &\leq \big((1 - h'(f)) + (1 + h'(f))\big)|\nabla f|_{w,\Phi} = 2|\nabla f|_{w,\Phi}, \end{aligned}$$

and it follows that all the inequalities are equalities $\mathfrak{m}$-a.e. in X. In particular we get

$$(1 + h'(f))|\nabla f|_{w,\Phi} = |\nabla f + h(f)|_{w,\Phi} \leq |\nabla f|_{w,\Phi} + |\nabla h(f)|_{w,\Phi} \qquad \mathfrak{m}\text{-a.e.}$$

and so $|\nabla h(f)| \geq h'(f)|\nabla f|$; since this is true also for $-h$ we obtain the conclusion.

(c) Thanks to (b) applied to $N = \{0\}$ at the function $f - g$ we have that $|\nabla(f - g)|_{w,\Phi} = 0$ $\mathfrak{m}$-a.e. on $\{f = g\}$; then the equality follows again by subadditivity and 1-homogeneity

$$|\nabla f|_{w,\Phi} - |\nabla(f - g)|_{w,\Phi} \leq |\nabla g|_{w,\Phi} \leq |\nabla f|_{w,\Phi} + |\nabla(f - g)|_{w,\Phi}.$$

□

3.4.3 The relaxation of the integral functional

In this section we will consider a very classical problem of the calculus of variation, generalized in this metric setting: the relaxation of integral funcionals where the integrand is depending only on the gradient. In general one aims at looking at a general functional

$$I(u, \Omega) = \int_\Omega f(x, u, \nabla u)\,\mathrm{d}x \qquad \Omega \subset \mathbb{R}^n, u \in C_c^\infty(\mathbb{R}^n),$$

and then asks whether the relaxation $\mathscr{I}$ of this functional in some topology admits a representation formula in its domain. For example if one take $f(x, u, p) = p^2$ then the domain $D(\mathscr{I}) = W^{1,2}(\mathbb{R}^n)$ and we have $\mathscr{I}(u, \Omega) = \int_\Omega |\nabla u|^2\,\mathrm{d}x$.

We want to generalize this last example to the metric setting, but with a general growth. We will consider only the integration on all X, to simplify the proof. Let us consider the functional G_Φ:

$$G_\Phi(f) := \begin{cases} \int_X \Phi(\mathrm{lip}_a(f))\,\mathrm{d}\mathfrak{m} & \text{if } f \in \mathrm{Lip}_0(X, \mathsf{d}) \\ +\infty & \text{otherwise.} \end{cases}$$

Then let $\mathscr{G}^\Phi$ be the semicontinuous relaxation of $G_\Phi(f)$ with respect to the L^1 convergence:

$$\mathscr{G}^\Phi(f) = \inf\left\{\liminf_{n\to\infty} \int_X \Phi(\mathrm{lip}_a(f_n))\,\mathrm{d}\mathfrak{m} \;:\; f_n \in \mathrm{Lip}_0(X, \mathsf{d}),\ f_n \to f \text{ in } L^1(X, \mathfrak{m})\right\}. \tag{3.4.13}$$

We want to find an explicit representation formula for $\mathscr{G}^\Phi$ and the obvious claim is that $\mathscr{G}^\Phi(f) = \int_X \Phi(|\nabla f|_{w,\Phi})\,\mathrm{d}\mathfrak{m}$, whenever this makes sense. This is obvious in the L^p case, since in this case $G_\Phi = F_\Phi^p$ and so we conclude using Theorem 3.4.8, in particular $\mathcal{F}_v^\Phi(f) = \||\nabla f|_{w,p}\|_p$. In the general Orlicz case, there is not a clear relation between G_Φ and F_Φ, apart from some inequalities, but we still can find that the relaxed quantities are represented by the same gradient. Of course, since we are talking of gradients, we have to require Ψ doubling; this will be the sufficient hypothesis to conclude.

Theorem 3.4.10 (Representation of $\mathscr{G}^\Phi$) *Let Φ be an N-function such that Ψ is doubling, and let $(X, \mathsf{d}, \mathfrak{m})$ be a separable complete space, where bounded sets have finite measure. Then we have*

$$\mathscr{G}^\Phi(f) = \begin{cases} \int_X \Phi(|\nabla f|_{w,\Phi})\,\mathrm{d}\mathfrak{m} & \text{if } f \in BL^{1,\Phi}(X,\mathsf{d},\mathfrak{m}) \\ +\infty & \text{otherwise.} \end{cases}$$

Proof. The proof is based on the duality formula

$$N_\Phi(g) := \int_X \Phi(g)\mathrm{d}\mathfrak{m} = \sup\left\{\int_X fg\,\mathrm{d}\mathfrak{m} - \int_X \Psi(f)\,\mathrm{d}\mathfrak{m} \ : \ f \in L^\Psi(X,\mathfrak{m})\right\}; \tag{3.4.14}$$

one inequality is trivial by Young inequality, for the other one it is sufficient to take

$$f_n(x) = \begin{cases} \varphi(g(x)) & \text{if } |x| \le n \text{ and } |g(x)| \le n \\ 0 & \text{otherwise.} \end{cases}$$

Since the right hand side of (3.4.14) is a supremum of continuous functionals for the weak-$*$ convergence in L^Φ, we have that N_Φ is l.s.c. with respect to this topology. Now let us consider a sequence $(f_n) \subseteq \mathrm{Lip}_0(X,\mathsf{d})$ that realizes the infimum in (3.4.13). Thanks to Lemma 3.4.3, up to subsequences we have that $f_n \to f$ in L^1 and $\mathrm{lip}_a(f_n) \overset{*}{\rightharpoonup} g$ for some $g \in L^\Phi(X,\mathfrak{m})$. Taking the lower semicontinuity of N_Φ into account and using that $g \ge |\nabla f|_{w,\Phi}$ (since g is a Φ-weak upper gradient, thanks to Theorem 3.4.8), we get

$$\mathscr{G}^\Phi(f) = \liminf_n N_\Phi(\mathrm{lip}_a(f_n)) \ge N_\Phi(g) \ge N_\Phi(|\nabla f|_{w,\Phi}).$$

This readily implies that

$$D(\mathscr{G}) \subseteq \left\{f \in BL^{1,\Phi}(X,\mathsf{d},\mathfrak{m}) \ : \ \int_X \Phi(|\nabla f|_{w,\Phi})\,\mathrm{d}\mathfrak{m} < \infty\right\}. \tag{3.4.15}$$

Now it remains to prove the other inequality: let us consider first functions $f \in BL^{1,\Phi}$ such that

$$N_\Phi(C|\nabla f|_{w,\Phi}) < \infty \qquad \text{for some } C > 1. \tag{3.4.16}$$

Let us take a sequence $(f_n) \subset \mathrm{Lip}_0(X,\mathsf{d})$ and $(g_n) \subset L^\Phi(X,\mathfrak{m})$ that satisfy (iv) in Theorem 3.4.8 with $g = |\nabla f|_{w,\Phi}$; then, using the convexity of Φ and taking $\varepsilon = \|g_n - |\nabla f|_{w,\Phi}\|_{\Phi,\mathfrak{m}}$, we have

$$\begin{aligned}\int_X \Phi(g_n)\,\mathrm{d}\mathfrak{m} &= \int_X \Phi\left((1-\varepsilon)\cdot\frac{|\nabla f|_{w,\Phi}}{1-\varepsilon} + \varepsilon\cdot\frac{g_n - |\nabla f|_{w,\Phi}}{\varepsilon}\right)\mathrm{d}\mathfrak{m} \\ &\le (1-\varepsilon)\int_X \Phi\big(\frac{|\nabla f|_{w,\Phi}}{1-\varepsilon}\big)\,\mathrm{d}\mathfrak{m} + \varepsilon\int_X \Phi\left(\frac{g_n - |\nabla f|_{w,\Phi}}{\varepsilon}\right)\mathrm{d}\mathfrak{m} \\ &\le \varepsilon + (1-\varepsilon)\int_X \Phi\left(\frac{|\nabla f|_{w,\Phi}}{1-\varepsilon}\right)\mathrm{d}\mathfrak{m}\end{aligned}$$

taking $n \to \infty$ we have that $\varepsilon \to 0$ and so, thanks to (3.4.16) and dominated convergence we have

$$\mathscr{G}^\Phi(f) \le \liminf_n \int_X \Phi(\mathrm{lip}_a(f_n))\,\mathrm{d}\mathfrak{m} \le \liminf_n \int_X \Phi(g_n)\,\mathrm{d}\mathfrak{m} \le N_\Phi(|\nabla f|_{w,\Phi});$$

Now, in order to remove the technical assumption (3.4.16) it is sufficient to notice (see (3.4.15)) that whenever $f \in D(\mathscr{G}^\Phi)$ we have that the function ρf satisfies (3.4.16) for every $0 < \rho < 1$, and so, since $\rho f \to f$ in L^1 when $\rho \to 1$, by the L^1-lower semcontinuity of $\mathscr{G}^\Phi$ and monotone convergence we get

$$\mathscr{G}^\Phi(f) \le \lim_{\rho\to 1^-} \mathscr{G}^\Phi(\rho f) \le \lim_{\rho\to 1^-} N_\Phi(\rho|\nabla f|_{w,\Phi}) = N_\Phi(|\nabla f|_{w,\Phi}),$$

where we used also the obvious property that $|\nabla(\rho f)|_{w,\Phi} = \rho|\nabla f|_{w,\Phi}$. □

3.5 Other possible definitions

For completeness, we compare the spaces $H^{1,\Phi}_v$ and $BL^{1,\Phi}$ with other spaces, namely $H^{1,\Phi}_c$ and $N^{1,\Phi}$, that are generalizations of those arising respectively in [25] and [75]. We recall briefly their definition:

Definition 3.5.1 ($H^{1,\Phi}_c$ and Cheeger Φ-relaxed energy) *For every $f \in L^1(X, \mathfrak{m})$ let us define*

$$\mathcal{F}^\Phi_c(f) := \inf\left\{\liminf_{n\to\infty} \|g_n\|_{(\Phi),\mathfrak{m}} \;:\; f_n \to f \textit{ in } L^1(X,\mathfrak{m}), \quad g_n \in UG(f_n)\right\}. \tag{3.5.1}$$

We recall that $UG(f)$ is the set of upper gradients for the function f. Then we define $H^{1,\Phi}_c(X, \mathsf{d}, \mathfrak{m}) = D(\mathcal{F}^\Phi_c)$.

In order to define the Newtonian space $N^{1,\Phi}$, generalization of the one presented in [75], we have to introduce the notion of Φ-modulus, following Section 2.1:

$$\mathrm{Mod}_{(\Phi),\mathfrak{m}}(\Gamma) := \inf\left\{\|f\|_{(\Phi),\mathfrak{m}} \;:\; f \in \mathcal{L}^\Phi_+(X,\mathfrak{m}), \quad \int_\gamma f \ge 1 \quad \text{for all } \gamma \in \Gamma\right\}. \tag{3.5.2}$$

Analogously, we say that a property holds for $\mathrm{Mod}_{(\Phi),\mathfrak{m}}$-a.e. curve γ if the set of curves for which the property fails has null Φ-modulus.

Definition 3.5.2 ($N^{1,\Phi}$ and Φ-upper gradient) *A function $g \in L^\Phi(X, \mathfrak{m})$ is a Φ-upper gradient for a Borel integrable function f if it holds*

$$|f(\gamma_{ini}) - f(\gamma_{fin})| \le \int_\gamma g < \infty \qquad \textit{for } \mathrm{Mod}_{(\Phi),\mathfrak{m}}\textit{-a.e. curve } \gamma. \tag{3.5.3}$$

Then the Newtonian space $N^{1,\Phi}$ is defined as the set of Borel integrable functions f that have a Φ-upper gradient. We can define $\mathcal{F}^\Phi_N(f) = \|g\|_{(\Phi),\mathfrak{m}}$, where g is the Φ-upper gradient of minimal norm.

Remark 3.5.3 *The existence of a minimal Φ-upper gradient is easy to prove thanks to the fact that if g is a Φ-upper gradient then $f \circ \gamma$ is absolutely continuous and $(f\circ\gamma)' \le g(\gamma)|\dot\gamma|$ for $\mathrm{Mod}_{(\Phi),\mathfrak{m}}$-a.e. γ (see [75, Proposition 3.1]). As we did for the Φ-weak upper gradients, this leads to the fact that if g_1, g_2 are Φ-upper gradients for f then also $\min\{g_1, g_2\}$ is a Φ-upper gradient, and so we can find a pointwise minimal object with a property similar to (3.4.6).*

The definition of Newtonian space is a little more subtle than the other ones since it is not invariant under modification in a negligible set. However we will see that there will be equivalence with all the other spaces also in this case, up to the choice of a suitable representative. We shall need a stability lemma for Φ-upper gradients. In the proof we will repeatedly use Proposition 2.1.2; notice that we can prove it for $\mathrm{Mod}_{(\Phi),\mathfrak{m}}$ following *verbatim* the proof with $\mathrm{Mod}_{(\Phi),\mathfrak{m}}$ instead of $\mathrm{Mod}_{p,\mathfrak{m}}$, with the exception of (v), where the reflexivity of L^p is used. However it is easy to see that (v) remains true whenever $\mathrm{Mod}_{(\Phi),\mathfrak{m}}(A_n) = 0$, thanks to the subadditivity (i).

Lemma 3.5.4 *Let $f_n \to f$ in $L^1(X,\mathfrak{m})$ and $g_n \to g$ in $L^\Phi(X,\mathfrak{m})$, where g_n is a Φ-upper gradient for f_n. Then there exists a representative of f that has g as a Φ-upper gradient; in particular if $g \in L^\Phi$ then $f \in N^{1,p}$ and $\mathcal{F}^\Phi_N(f) \leq \|g\|_{(\Phi),\mathfrak{m}}$.*

Proof. Here we follow [46, Lemma 7.8]. Let us denote by Γ_0 the set of curves where the following holds (we will often identify $\gamma \in \mathscr{C}(X)$ with $\mathrm{i}\gamma \in \mathrm{AC}^\infty_c([0,1];X)$, when needed):

$$f_n \circ \gamma \in \mathrm{AC}([0,1]) \quad \text{ and } \quad |(f_n \circ \gamma)'|(t) \leq g_n(\gamma_t)|\dot\gamma_t| \qquad \forall n \in \mathbb{N}.$$

By hypothesis we have that $\mathscr{C}(X) \setminus \Gamma_0$ is $\mathrm{Mod}_{(\Phi),\mathfrak{m}}$-negligible. We will need also the set Γ_g of curves for which $\int_\gamma g < \infty$; by Proposition 2.1.2(ii) this happens for $\mathrm{Mod}_{(\Phi),\mathfrak{m}}$-a.e. curve. Let us notice that, up to subsequences, we can assume that $f_n \to f$ almost everywhere. In particular, letting N be the $\mathfrak{m}$-null set where $\lim_{n\to\infty} f_n$ does not exist, we can consider the representative

$$f(x) = \begin{cases} \lim_{n\to\infty} f_n(x) & \forall x \notin N \\ 0 & \text{otherwise.}\end{cases}$$

It is clear that the set of curves $\Gamma_N = \{\gamma \,:\, \mathscr{L}^1(\gamma^{-1}(N)) > 0\}$ is $\mathrm{Mod}_{(\Phi),\mathfrak{m}}$-negligible in fact, considering $\rho = \infty \cdot \chi_N$ we have $\|\rho\|_{(\Phi)} = 0$ and $\int_\gamma \rho = \infty$ whenever $\gamma \in \Gamma_N$. In particular we have $\mathscr{L}^1(\gamma^{-1}(N)) = 0$ for $\mathrm{Mod}_{(\Phi),\mathfrak{m}}$-a.e. curve and so we have

$$f_n \circ \gamma \to f \circ \gamma \quad \mathscr{L}^1\text{-a.e. in } [0,1] \text{ for } \mathrm{Mod}_{(\Phi),\mathfrak{m}}\text{-a.e. curve } \gamma.$$

Denote Γ_1 the set of curves where this happens.

Thanks to Proposition 2.1.2(iii) applied to $|g_n - g|$ we have that if $g_n \to g$ strongly in $L^\Phi(X,\mathfrak{m})$ then there exists a subsequence such that

$$\int_\gamma |g_n - g| \to 0 \qquad \text{for } \mathrm{Mod}_{(\Phi),\mathfrak{m}}\text{-almost every curve } \gamma; \tag{3.5.4}$$

denote Γ_2 the set of curves where this happens.

Now, thanks to (3.5.4) we have that if $\gamma \in \Gamma_2$ then the functions $h_n(t) = g_n(\gamma_t)|\dot\gamma_t|$ are equi-integrable; whenever we have also $\gamma \in \Gamma_0$, then by hypothesis h_n bounds from above the derivative of the absolutely continuous function $f_n \circ \gamma$, and so we deduce that the sequence $(f_n \circ \gamma)$ is equicontinuous and so if they are converging $\mathscr{L}^1$-a.e. to some function, they are also converging uniformly, and in particular everywhere; this implies that $\gamma^{-1}(N) = \emptyset$ for every $\gamma \in \Gamma_0 \cap \Gamma_1 \cap \Gamma_2$. In particular for $\mathrm{Mod}_{(\Phi),\mathfrak{m}}$-almost every curve γ it happens that

$$|f(\gamma_1) - f(\gamma_0)| = \lim_{n\to\infty} |f_n(\gamma_1) - f_n(\gamma_0)| \leq \lim_{n\to\infty} \int_\gamma g_n = \int_\gamma g. \tag{3.5.5}$$

□

3.5.1 Comparison of $\mathcal{F}_c^\Phi$ with $\mathcal{F}_{BL}^\Phi, \mathcal{F}_v^\Phi$

It is clear that $\|g\|_{(\Phi),\mathfrak{m}} \geq \mathcal{F}_{BL}^\Phi(f)$ whenever $g \in UG(f)$, recalling that (3.2.3) holds for every g upper gradient of f, and then using the estimate (3.2.4). By relaxation we get $\mathcal{F}_c^\Phi \geq \mathcal{F}_{BL}^\Phi$; moreover we have that $\mathrm{lip}_a(f)$ is an upper gradient for $f \in \mathrm{Lip}_0(X, \mathsf{d})$ and so it is clear that $\mathcal{F}_v^\Phi \geq \mathcal{F}_c^\Phi$ and so by the equivalence Theorem 3.3.1 we obtain

$$H_v^{1,\Phi}(X, \mathsf{d}, \mathfrak{m}) = H_c^{1,\Phi}(X, \mathsf{d}, \mathfrak{m}) \qquad \text{and} \qquad \mathcal{F}_v^\Phi(f) = \mathcal{F}_c^\Phi(f) \quad \forall f \in L^1(X, \mathfrak{m}). \tag{3.5.6}$$

3.5.2 Comparison of $N^{1,\Phi}$ with $\mathcal{F}_{BL}^\Phi, \mathcal{F}_v^\Phi$

As for $N^{1,\Phi}$, we have that $\mathrm{Mod}_{(\Phi),\mathfrak{m}}(\Gamma) = 0 \Rightarrow \boldsymbol{\pi}(\mathsf{i}^{-1}(\Gamma)) = 0$ for every Ψ-plan $\boldsymbol{\pi}$. In fact, taking a $\mathrm{Mod}_{(\Phi),\mathfrak{m}}$-null set Γ we have that for every $\varepsilon > 0$ there exists a Borel function ρ such that $\|\rho\|_{(\Phi)} \leq \varepsilon$ and $\int_\gamma \rho \geq 1$ for every $\gamma \in \Gamma$; now, letting $\Gamma_0 = \mathsf{i}^{-1}(\Gamma)$, we have (with the usual estimate (3.2.4)):

$$\boldsymbol{\pi}(\Gamma_0) \leq \iint_\gamma \rho \, \mathrm{d}\boldsymbol{\pi} \leq \varepsilon \cdot C(\boldsymbol{\pi}) \cdot \int \int_0^1 \||\dot{\gamma}_t\|_{\Psi,\boldsymbol{\pi}} \, \mathrm{d}t.$$

Letting $\varepsilon \to 0$ we obtain that $\boldsymbol{\pi}(\Gamma_0) = 0$ for every $\boldsymbol{\pi}$-plan and thus Γ_0 is Φ-negligible. So we have that if $f \in N^{1,\Phi}$ and g is a Φ-upper gradient then $f \in BL^{1,\Phi}$ and g is a Φ-weak upper gradient. In particular we have

$$\mathcal{F}_N^\Phi \geq \mathcal{F}_{BL}^\Phi \qquad \text{and } N^{1,\Phi} \subset BL_1^{1,\Phi}, \tag{3.5.7}$$

where $BL_1^{1,\Phi}$ is the set of $f \in BL^{1,\Phi}$ such that $f \circ \gamma \in W^{1,1}(0,1)$ for Φ-almost every γ.

As we noted in Section 3.4.1, in general we have $BL_1^{1,\Phi} \subsetneq BL^{1,\Phi}$, while they coincide if Ψ is doubling (Theorem 3.4.8). In particular in general we have $N^{1,\Phi} \subsetneq BL^{1,\Phi}$, but, thanks to the following proposition, they coincide when Ψ is doubling.

Proposition 3.5.5 *Let us assume Ψ doubling. Then we have $H_v^{1,\Phi} \subseteq N^{1,\Phi}$ and also $\mathcal{F}_v^\Phi \geq \mathcal{F}_N^\Phi$, meaning that for every $f \in H_v^{1,\Phi}$ there exists a Borel representative $\tilde{f} \in N^{1,\Phi}$, such that $\mathcal{F}_v^\Phi(f) \geq \mathcal{F}_N^\Phi(\tilde{f})$*

Proof. By Theorem 3.4.8(iv) there are sequences $(f_n) \subset \mathrm{Lip}_0(X, \mathsf{d})$, $(g_n) \subset L^\Phi(X, \mathfrak{m})$ such that $f_n \to f$ in L^1, $g_n \geq \mathrm{lip}_a(f_n)$ and $g_n \to g$ strongly in $L^\Phi(X, \mathfrak{m})$. Moreover $\mathcal{F}_v^\Phi(f) = \|g\|_{(\Phi),\mathfrak{m}}$.

Now we can apply Lemma 3.5.4 to the functions $\hat{f}_n$ with upper gradients g_n, obtaining that f has a representative in $N^{1,\Phi}$ with g as Φ-upper gradient. In particular we obtain that $\mathcal{F}_N^\Phi(f) \leq \|g\|_{(\Phi),\mathfrak{m}} = \mathcal{F}_v^\Phi(f)$. □

In the next theorem we collect the results of this section.

Theorem 3.5.6 *Let $(X, \mathsf{d}, \mathfrak{m})$ be a Polish space endowed with a measure finite on bounded sets. Then we have:*

$$H_v^{1,\Phi}(X, \mathsf{d}, \mathfrak{m}) = H_c^{1,\Phi}(X, \mathsf{d}, \mathfrak{m}) = BL^{1,\Phi}(X, \mathsf{d}, \mathfrak{m}) \supseteq N^{1,\Phi}(X, \mathsf{d}, \mathfrak{m});$$

$$\mathcal{F}_v^\Phi(f) = \mathcal{F}_c^\Phi(f) = \mathcal{F}_{BL}^\Phi(f) \leq \mathcal{F}_N^\Phi(f) \qquad \forall f \ \textit{Borel function.}$$

Moreover if Ψ is doubling we have all equalities and also equalities at the level of weak gradients: $|\nabla f|_{N,\Phi} = |\nabla f|_{w,\Phi}$.

Proof. We already observed in (3.5.6) the equivalence with $H_c^{1,\Phi}$. Then we proved in (3.5.7) the inclusion for $N^{1,\Phi}$. For the case in which Ψ is doubling, the other inclusion and the other inequality are proved using Proposition 3.5.5 and of course Theorem 3.3.1.

In order to prove the equality of the weak gradients, it is sufficient to note that in the construction provided in Proposition 3.5.5 we prove indeed that every g in Theorem 3.4.8(iv) is also a Φ-weak gradient; but in that theorem we prove that (iv) is satisfied by $|\nabla f|_{w,\Phi}$, and so we obtain that the minimal Φ-weak upper gradient is a Φ-weak gradient. Moreover we proved also that every Φ-weak gradient is also a Φ-weak upper gradient and so we conclude. $\square$

CHAPTER 4

The spaces BV and $H^{1,1}$

In this chapter we consider more closely the degenerate case $\Phi(t) = t$, corresponding to the definition of the BV space. In this case almost all the proof in the equivalence Theorem 3.3.1 still works but a key point, namely Section 1.6, where the analysis of the Hopf-Lax semigroup is done, doesn't work anymore for $\Phi(t) = t$; in Section 4.2 we will fill this gap providing a little weaker result on length spaces, but still sufficient for the equivalence theorem.

The proof of the equivalence of various different definitions will follow the paper [5]. Here we introduce two more spaces, in the spirit of $H_v^{1,\Phi}$ and $BL^{1,\Phi}$, whose equivalence with the previous ones permits to obtain global approximation by Lipschitz functions with bounded support, and also putting the asymptotic Lipschitz constant in place of the slope. The BV case is very particular since, despite the fact that Ψ is not doubling (it is not even finite on the whole real line), we can give a localized version of the variational energy, that is the *total variation* measure $|Df|$; the definition in the relaxed sense (4.4.3) and its basic properties has been given in [67] under some structural assumption, and then extended to locally finite metric measure spaces in [5].

4.1 BV functions and total variation on Euclidean spaces

We refer to Chapter 3 of [7] for a complete review of this topic, with all the proofs; here we will only overview the main properties needed in this paper.

Given an open set $A \subseteq \mathbb{R}^d$, $f \in L^1(A)$ is said to be of *bounded variation* in A if one of the following three equivalent properties hold:

(a) the distributional derivative Df is a $\mathbb{R}^d$-valued measure with finite total variation in A.

(b) The following quantity, called total variation of f in A, is finite:

$$TV_f(A) := \sup\left\{\int_A f \operatorname{div}\varphi \, \mathrm{d}x \ : \ \varphi \in C_c^1(A;\mathbb{R}^d), \ |\varphi| \le 1\right\}.$$

(c) There exists a sequence $(f_n) \subset C^\infty(A)$ converging fo f in $L^1_{\rm loc}(A)$, with equibounded energies: $\sup_n \int_A |\nabla f_n| \,\mathrm{d}x < \infty$.

The equivalence between (a), (b) and (c) leads to relations between the corresponding quantities involved: in particular we have

$$|Df|(A) = TV_f(A) \le \liminf_{n\to\infty} \int_A |\nabla f_n| \,\mathrm{d}x.$$

By means of standard mollifiers and partitions of unity we can get also the following stronger result: there exists a sequence of functions $f_n \in C^\infty(A)$ convergent to f in $L^1(A)$ and such that $|Df_n|(A) \to |Df|(A)$. In our metric context we simply replace $C^\infty(A)$ by the space of locally Lipschitz functions on A.

Moreover the second definition gives us easily the crucial property that the total variation $|Df|$ of the distributional derivative in open sets is lower semicontinuous with respect to $L^1_{\rm loc}$ convergence:

$$\liminf_{n\to\infty} |Df_n|(A) \ge |Df|(A) \qquad \forall A \subseteq \mathbb{R}^d \text{ open set}, \quad f_n \to f \text{ in } L^1_{\rm loc}(A). \tag{4.1.1}$$

4.2 Hopf-Lax formula and Hamilton-Jacobi equation

Here, as we did in Section 1.6, we want to study some elementary properties of the Hopf-Lax formula in a metric setting, in the degenerate case $\Phi(t) = |t|$ (suitable for the study of the ∞-Wasserstein distances, and the corresponding Kuwada lemma), not covered in the previous discussion because we exploited the properties of N-functions. We are dealing with a very simple convex lower semicontinuous Lagrangian, the Lagrange dual of $\Phi(t) = t$:

$$L(s) = \begin{cases} 0 & \text{if } s \le 1; \\ \infty & \text{if } s > 1. \end{cases} \tag{4.2.1}$$

We will use also, for a finer analysis, correspondingly, the *ascending* slope $|\nabla^+ f|$ and the *descending* slope $|\nabla^- f|$:

$$|\nabla^\pm f|(x) := \overline{\lim_{y\to x}} \frac{(f(y) - f(x))^\pm}{\mathsf{d}(y,x)}. \tag{4.2.2}$$

Let $f : X \to \mathbb{R}$ be a Lipschitz function. We set $Q_0 f(x) = f(x)$ and, for $t > 0$,

$$Q_t f(x) := \inf_{y\in X} \Big\{ f(y) + tL\Big(\frac{\mathsf{d}(x,y)}{t}\Big)\Big\}. \tag{4.2.3}$$

Due to the particular form of our Lagrangian, we get

$$Q_t f(x) := \inf_{\mathsf{d}(x,y)\le t} f(y). \tag{4.2.4}$$

Obviously, these transformations act almost as a semigroup: in fact, the triangle inequality gives

$$Q_s Q_t f(x) = \inf_{\mathsf{d}(y,x)\le s} \Big\{ \inf_{\mathsf{d}(y,z)\le t} f(z)\Big\} \ge \inf_{\mathsf{d}(x,z)\le s+t} f(z) = Q_{s+t} f(x).$$

Moreover, if (X, d) is a length space, we have equality and thus Q_t is a semigroup. In fact, under this assumption, for every z such that $\mathsf{d}(x,z) < s+t$ there exists a constant speed curve

$\gamma : [0,1] \to X$ whose length is less than $s+t$ and such that $\gamma_0 = x$ and $\gamma_1 = z$; in particular there will be a time $\eta := s/(s+t)$ such that $y := \gamma_\eta$ satisfies $\mathsf{d}(x,y) < s$ and $\mathsf{d}(y,z) < t$. It follows that $Q_s Q_t f(x) \le \inf_{\mathsf{d}(x,z)<s+t} f(z)$. In order to conclude, one has to observe that, if f is continuous, then

$$\inf_{\mathsf{d}(x,z)\le r} f(z) = \inf_{\mathsf{d}(x,z)<r} f(z) \qquad \forall r > 0,$$

and this is true because in a length space the closure of the open ball is the closed ball.

Also, it is easy to check that the length space property ensures that the Lipschitz constant does not increase:

$$\mathrm{Lip}(Q_t f) \le \mathrm{Lip}(f). \tag{4.2.5}$$

Now we look at the time derivative, to get information on the Hamilton-Jacobi equation satisfied by $Q_t f(x)$:

Theorem 4.2.1 (Time derivative of $Q_t f$) *Let $x \in X$. The map $t \mapsto Q_t f(x)$ is nonincreasing in $[0,\infty)$ and satisfies:*

$$\frac{\mathrm{d}}{\mathrm{d}t} Q_t f(x) + |\nabla Q_t f(x)| \le 0 \qquad \textit{for a.e. } t > 0. \tag{4.2.6}$$

Moreover, if (X,d) is a length space, the map $t \mapsto Q_t f$ is Lipschitz from $[0,\infty)$ to $C(X)$, with Lipschitz constant $\mathrm{Lip}(f)$.

Proof. The basic inequality, that we will use in the first part of the proof is:

$$Q_s f(y) \le Q_{s'} f(y') \qquad \text{whenever } s \ge s' + \mathsf{d}(y,y'). \tag{4.2.7}$$

It holds because the inequality implies $B(y',s') \subseteq B(y,s)$ and thus it is clear by the very definition of $Q_t f$. Now we take x_i and y_i converging to x such that:

$$\lim_{i\to\infty} \frac{Q_t f(x_i) - Q_t f(x)}{\mathsf{d}(x_i,x)} = -|\nabla^- Q_t f|(x), \qquad \lim_{i\to\infty} \frac{Q_t f(x) - Q_t f(y_i)}{\mathsf{d}(x,y_i)} = -|\nabla^+ Q_t f|(x).$$

Now we consider the inequalities, given by (4.2.7), involving x, x_i, y_i:

$$Q_{t+\mathsf{d}(x_i,x)} f(x) \le Q_t f(x_i), \qquad Q_t f(y_i) \le Q_{t-\mathsf{d}(x,y_i)} f(x)$$

and let us define, for brevity, $s_i = \mathsf{d}(x_i,x)$ and $r_i = \mathsf{d}(x,y_i)$. Then we have

$$\begin{aligned}\liminf_{h\to 0^+} \frac{Q_{t+h} f(x) - Q_t f(x)}{h} &\le \liminf_{i\to\infty} \frac{Q_{t+s_i} f(x) - Q_t f(x)}{s_i} \\ &\le \lim_{i\to\infty} \frac{Q_t f(x_i) - Q_t f(x)}{s_i} = -|\nabla^- Q_t f|(x)\end{aligned}$$

and, similarly,

$$\begin{aligned}\liminf_{h\to 0^-} \frac{Q_{t+h} f(x) - Q_t f(x)}{h} &\le \liminf_{i\to\infty} \frac{Q_t f(x) - Q_{t-r_i} f(x)}{r_i} \\ &\le \lim_{i\to\infty} \frac{Q_t f(x) - Q_t f(y_i)}{r_i} = -|\nabla^+ Q_t f|(x).\end{aligned}$$

Using that $|\nabla f| = \max\{|\nabla^+ f|, |\nabla^- f|\}$, the combination of these inequalities gives

$$\liminf_{h\to 0} \frac{Q_{t+h}f(x) - Q_t f(x)}{h} \leq -|\nabla Q_t f|(x) \qquad \forall x \in X, \ \ \forall t > 0.$$

Since $Q_t f(x)$ is obviously non increasing w.r.t. t, we get that is differentiable almost everywhere and so we get the thesis.

If we suppose that (X, d) is also a length space, using the semigroup property and (4.2.5) we get that

$$Q_s f(x) - Q_t f(x) = Q_s f(x) - Q_{t-s}(Q_s f)(x) \leq (t-s)\,\mathrm{Lip}(Q_s f) \leq (t-s)\,\mathrm{Lip}(f) \quad \forall s \in [0, t],$$

and so the thesis. □

Note that, in case (X, d) is not a length space, it might happen that balls are not connected and, as a consequence, that $t \mapsto Q_t\varphi(x)$ is discontinuous; as an example we can take X the curve in Figure 4.1, with the distance induced as subset of $\mathbb{R}^2$.

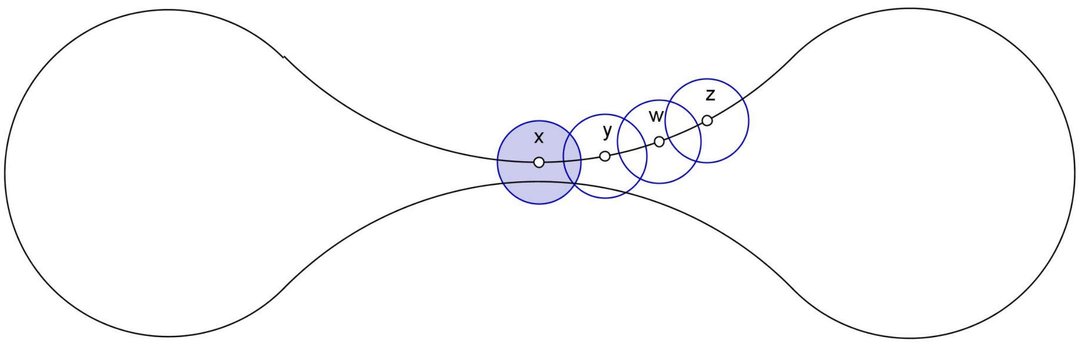

Figure 4.1: *Example of a compact metric space* (X, d) *that is not a length space, having a time discontinuous Hopf-Lax semigroup* Q_t

It is clear that some balls, such as the shaded one centered in x, are disconnected; furthermore if we take a Lipschitz function f equal to 0 in the upper part of the curve and equal to 1 in the lower one, doing an interpolation between two values only in the rightmost and leftmost parts, it is easy to see that $Q_t f(p)$ is discontinuous both in time and space.

Remark 4.2.2 *Unlike the N-function case, here we don't reach the Hamilton-Jacobi inequality for the asymptotic Lipschitz constant, but only for the slope. It is still an open problem whether if* (4.2.6) *holds with the asymptotic Lipschitz constant. However, thanks to Proposition 4.4.1 we can still prove that in* (4.4.3) *we can approximate with* $\mathrm{lip}_a(f_h)$ *in place of* $|\nabla f_h|$.

4.3 The ∞-Wasserstein distance

We already recalled the Ψ-Wasserstein for any Ψ convex lower semicontinuous, even with values in $\mathbb{R} \cup \{+\infty\}$. Thus here we just recall the ∞-Wasserstein distance, that is, the Ψ-Wasserstein distance when $\Psi = L$, where L is defined in (4.2.1). We have

$$W_\infty(\mu,\nu) := \min\left\{\|\mathsf{d}\|_{L^\infty(\gamma)} \,|\gamma \in \Gamma(\mu,\nu)\right\}.$$

It is known (see for instance [24]) that W_∞ is the monotone limit of W_p as p goes to infinity, at least when we are dealing with probability measures; we want to consider also this limit case as a transport problem, in order to have a dual formulation that will be used later on; as it has been already pointed out we need "test distances", that are really transport distances:

$$W_\infty^{(s)}(\mu,\nu) = \min\Big\{\int_{X\times X} sL\Big(\frac{\mathsf{d}(x,y)}{s}\Big)\mathrm{d}\gamma \mid \gamma \in \Gamma(\mu,\nu)\Big\},$$

and for them a duality formula holds:

$$W_\infty^{(s)}(\mu,\nu) = \sup_{\varphi\in \mathrm{Lip}_0(X,\mathsf{d})} \int_X Q_s\varphi \mathrm{d}\mu - \int_X \varphi \mathrm{d}\nu. \tag{4.3.1}$$

In this case, being the cost degenerate, we have that $W_\infty^{(s)} \le 0$ if and only if $W_\infty \le s$.

4.4 Four notions of BV function

Let (X,d) be a complete and separable metric space and let $\mathfrak{m}$ be a nonnegative Borel measure in X. In this section we introduce four notions of BV function and, correspondingly, four notions of total variation. Only three of them will be measures, the other one giving only the value of the total variation of the entire space and difficult to localize. We recall that the aim of this chapter is to show that these notions are equivalent.

4.4.1 BV functions in the variational sense

In the same spirit of Definition 3.1.1 we say that a function $f \in L^1(X,\mathfrak{m})$ is said BV in the variational sense if there exists a sequence $(f_n) \subset \mathrm{Lip}_0(X,\mathsf{d})$ converging to f in $L^1(X,\mathfrak{m})$ and with equibounded energies: $\sup_n \|\mathrm{lip}_a(f_n)\|_1 < \infty$. We shall denote this space by $BV_v(X,\mathsf{d},\mathfrak{m})$. We define also the total variation of the entire space

$$|Df|_v(X) = \inf\left\{\liminf_{h\to\infty} \int_X \mathrm{lip}_a(f_h)\,\mathrm{d}\mathfrak{m} :\ (f_h) \subset \mathrm{Lip}_0(X,\mathsf{d}),\ f_h \to f \text{ in } L^1(X,\mathfrak{m})\right\}. \tag{4.4.1}$$

Note that if we consider $\Phi(t) = t$ then we have $\mathcal{F}_v^\Phi(f) = |Df|_v(X)$. This notion can't be localized as in (4.4.3), if we want to be consistent with the Euclidean case; in fact for a general open set $A \subseteq \mathbb{R}^d$ it is necessary to have *locally* Lipschitz functions approximating the function f (take as an example $f(\theta,r) = \theta$, in the case $X = \overline{B}(0,1) \subseteq \mathbb{R}^2$ with the Lebesgue measure, and as an open set $A = \overline{B}(0,1) \setminus (\{0\}\times[0,1])$).

Proposition 4.4.1 *Let $\mathfrak{m}$ be a measure that is finite on bounded sets. Then for every $f \in \mathrm{Lip}_0(X,\mathsf{d})$ we have*

$$|Df|_v(X) \le \int_X |\nabla f|\,\mathrm{d}\mathfrak{m}.$$

Proof. We notice that, thanks to the general theory of Sobolev spaces in metric measure spaces (see for example [4] or Chapter 3), since $|\nabla f|$ is an upper gradient for $f \in \mathrm{Lip}_0(X, \mathsf{d})$, and $f, |\nabla f| \in L^2(X, \mathfrak{m})$, there is a sequence $(f_h) \subseteq \mathrm{Lip}_0(X, \mathsf{d})$ such that $f_h \to f$ in L^2 and $\mathrm{lip}_a(f_h) \to g$ in L^2 with $g \leq |\nabla f|$; thanks to Lemma 3.3.3 we can take f_h to have uniform bounded support, since f has bounded support. In particular, recalling that $\mathfrak{m}$ is finite on bounded sets, we have that $f_h \to f$ in L^1 and $\mathrm{lip}_a(f_h) \to g$ in L^1 and so this gives that $|Df|_v(X) \leq \int_X g \,\mathrm{d}\mathfrak{m} \leq \int_X |\nabla f| \,\mathrm{d}\mathfrak{m}$ and so we obtain the thesis. □

Moreover, thanks to the lower semicontinuity of the total variation, we obtain from Proposition 4.4.1 an equivalent formulation:

$$|Df|_v(X) = \inf \left\{ \liminf_{h \to \infty} \int_X |\nabla f_h| \,\mathrm{d}\mathfrak{m} : \ f_h \in \mathrm{Lip}_0(X, \mathsf{d}), \ f_h \to f \text{ in } L^1(X, \mathfrak{m}) \right\}. \tag{4.4.2}$$

4.4.2 BV functions in the relaxed sense

We can define a slightly bigger space, requiring that the approximating functions are only *locally* Lipschitz. We shall denote this space by $BV_*(X, \mathsf{d}, \mathfrak{m})$.
We already noticed that this definition coincides with the classical one in Euclidean spaces. Associated to this definition is the relaxed total variation $|Df|_*$, defined on open sets $A \subseteq X$ as:

$$|Df|_*(A) := \inf \left\{ \liminf_{h \to \infty} \int_A \mathrm{lip}_a(f_h) \,\mathrm{d}\mathfrak{m} : \ (f_h) \subset \mathrm{Lip}_{\mathrm{loc}}(A), \ f_h \to f \text{ in } L^1(A) \right\}. \tag{4.4.3}$$

Here "locally Lipschitz in an open set A" means that for all $x \in A$ there exists $r > 0$ such that $B_r(x) \subseteq A$ and the restriction of f to $B_r(x)$ is Lipschitz.

This definition can be seen as the localized version of the variational one; it is clear that $|Df|_v(X) \geq |Df|_*(X)$, but the converse inequality is not at all obvious.

This definition is slightly stronger than the ones considered in [67] and [5] since in their definition the authors use the slope instead of the asymptotic Lipschitz constant. They prove that in their context the set function $A \mapsto |Df|_s(A)$ is the restriction to open sets of a finite Borel measure (the subscript s stands for slope). We follow their proof in order to prove the same for $|Df|_*$: we investigate more closely the properties of this set function in the following lemma. We will write $A \Subset B$ whenever A, B are open sets and $\mathsf{d}(A, X \setminus B) > 0$ (in particular, $A \Subset B$ implies $A \subseteq B$). We say that A_1 and A_2 are well separated if $\mathrm{dist}(A_1, A_2) > 0$.

Lemma 4.4.2 *Let $\mathcal{A}(X)$ be the class of open subsets of X, $u \in L^1(X, \mathfrak{m})$ and let $|Du|_* : \mathcal{A}(X) \to [0, \infty]$ be defined as in (4.4.3), with the convention $|Du|_*(\emptyset) = 0$. Then, $|Du|_*$ satisfies the following properties:*

(i) $|Du|_*(A_1) \leq |Du|_*(A_2)$ *whenever* $A_1 \subseteq A_2$*;*

(ii) $|Du|_*(A_1 \cup A_2) \leq |Du|_*(A_1) + |Du|_*(A_2)$*, with equality if A_1 and A_2 are well separated;*

(iii) If A_n are open and $A_n \subseteq A_{n+1}$ it holds

$$\lim_{n \to \infty} |Du|_*(A_n) = |Du|_*\Big(\bigcup_n A_n\Big). \tag{4.4.4}$$

In particular the formula

$$|Du|_*(B) := \inf\{|Du|_*(A) \ : \ A \subseteq X \text{ open}, \ B \subseteq A\}$$

provides a σ-subadditive extension of $|Du|_*$ whose additive sets, in the sense of Carathéodory, contain $\mathscr{B}(X)$. If follows that $|Du|_* : \mathscr{B}(X) \to [0,\infty]$ is a σ-additive Borel measure.

Proof. The verifications of monotonicity and the additivity on well separated sets are standard. Since we will use (iii) in the proof of the first statement of (ii), we prove (iii) first, denoting $A := \cup_n A_n$. It is sufficient to prove that $\sup |Du|_*(A_n) \geq |Du|_*(A)$ because the converse inequality is trivial by monotonicity, so we can assume that $\sup_n |Du|_*(A_n) < \infty$.

First, we reduce ourselves to the case when A_n satisfy the additional condition

$$\operatorname{dist}(\overline{A}_n, X \setminus A_{n+1}) > 0 \qquad \forall n \in \mathbb{N}. \tag{4.4.5}$$

In order to realize that the restriction to this case is possible, suffices to consider the sets

$$A'_n := \left\{x \in X \ : \ \operatorname{dist}(x, X \setminus A_n) \geq \frac{1}{n}\right\}$$

which satisfy (4.4.5), are contained in A_n and whose union is still equal to A.

In particular, if we call

$$\begin{cases} C_1 = A_2 \\ C_k = A_k \setminus \overline{A_{k-2}} & \text{if } k \geq 2, \end{cases}$$

it is clear that the families $\{C_{3k+1}\}$, $\{C_{3k+2}\}$, $\{C_{3k+3}\}$ are well separated, hence $\sum_j |Du|_*(C_{3j+i}) < \infty$ for all $i \in \{1,2,3\}$. It follows that for any $\varepsilon > 0$ we can find an integer $\bar{k}$ such that

$$\sum_{n=\bar{k}}^{\infty} |Du|_*(C_n) \leq \varepsilon. \tag{4.4.6}$$

Now, to prove (4.4.4) we build a sequence $(u_m) \subseteq \mathrm{Lip}_{\mathrm{loc}}(A)$ such that $u_m \to u$ in $L^1(A, \mathfrak{m})$ and

$$|Du|_*(A_{\bar{k}}) + 2\varepsilon \geq \liminf_{m\to\infty} \int_A \mathrm{lip}_a(u_m)\, \mathrm{d}\mathfrak{m}.$$

In order to do so, we fix m and set $D_h = C_{h+\bar{k}}$, $B_h = A_{h+\bar{k}}$ if $h \geq 1$, $D_0 = B_0 = A_{\bar{k}}$. Then we choose $\psi_{k,h} \in \mathrm{Lip}_{\mathrm{loc}}(D_h)$ in such a way that

$$\int_{D_h} \mathrm{lip}_a(\psi_{k,h})\, \mathrm{d}\mathfrak{m} \leq |Du|_*(D_h) + \frac{1}{m2^k}. \tag{4.4.7}$$

We are going to use Lemma 4.4.3 below with $M = B_h$, $N = D_{h+1}$, so we denote by c_h and $H_h \Subset B_h \cap D_{h+1}$ the constants and the domains given by the lemma. It is then easy to find sufficiently large integers $k(h) \geq h$ satisfying

$$c_h \int_{\overline{H}_h} |\psi_{k(h),h} - u|\, \mathrm{d}\mathfrak{m} \leq \frac{\varepsilon}{2\cdot 2^h} \qquad \text{and} \qquad c_h \int_{\overline{H}_h} |\psi_{k(h+1),h+1} - u|\, \mathrm{d}\mathfrak{m} \leq \frac{\varepsilon}{2\cdot 2^h}. \tag{4.4.8}$$

This is possible because $\overline{H}_h$ is contained in $B_h \cap D_{h+1}$ which, in turn, is contained in D_h. In addition, possibly increasing $k(h)$, we can also have:

$$\int_{D_h} |\psi_{k(h),h} - u|\, \mathrm{d}\mathfrak{m} \leq \frac{1}{m2^h}. \tag{4.4.9}$$

Now we define by induction on h functions $u_{m,h} \in \mathrm{Lip}_{\mathrm{loc}}(B_h)$ for $h \geq 0$: we set $u_{m,0} = \psi_{k(0),0}$ and, given $u_{m,h}$, we build $u_{m,h+1}$ in such a way that:

$$\begin{cases} u_{m,h+1} \equiv u_{m,h} & \text{on } B_{h-1} \\ u_{m,h+1} \equiv \psi_{k(h+1),h+1} & \text{on } B_{h+1} \setminus \overline{B_h}, \end{cases} \tag{4.4.10}$$

$$\|u_{m,h} - u\|_{L^1(B_h)} \leq \frac{1}{m}\left(1 - \frac{1}{2^h}\right), \tag{4.4.11}$$

$$\int_{B_{h+1}} \mathrm{lip}_a(u_{m,h+1})\,\mathrm{d}\mathfrak{m} \leq \int_{B_h} \mathrm{lip}_a(u_{m,h})\,\mathrm{d}\mathfrak{m} + \int_{D_{h+1}} \mathrm{lip}_a(\psi_{k(h+1),h+1})\,\mathrm{d}\mathfrak{m} + \frac{\varepsilon}{2^h}. \tag{4.4.12}$$

Once we have this we are done because we can construct $u_m(x) = u_{m,h}(x)$ if $x \in B_{h-1}$, then it is clear that u_m is well defined thanks to the first equation in (4.4.10) and locally Lipschitz in A. In addition $\|u_m - u\|_{L^1(A)} \leq 1/m$ thanks to (4.4.11) and the monotone convergence theorem and, iterating (4.4.12) and using (4.4.7) and $k(h) \geq h$, we get

$$\begin{aligned} \int_A \mathrm{lip}_a(u_m)\,\mathrm{d}\mathfrak{m} &= \lim_{h\to\infty} \int_{B_h} \mathrm{lip}_a(u_{m,h+1})\,\mathrm{d}\mathfrak{m} \leq \lim_{h\to\infty} \int_{B_{h+1}} \mathrm{lip}_a(u_{m,h+1})\,\mathrm{d}\mathfrak{m} \\ &\leq \sum_{i=0}^{\infty} |Du|_*(D_i) + \frac{2}{m} + \varepsilon \leq |Du|_*(A_{\bar{k}}) + 2\varepsilon + \frac{2}{m}. \end{aligned}$$

In order to prove the induction step in the construction of $u_{m,h}$ we use Lemma 4.4.3 with $M = B_h$, $N = D_{h+1}$, $u = u_{m,h}$ and $v = \psi_{k(h+1),h+1}$. So, applying (4.4.13) of the lemma we find a function $w = u_{m,h+1}$ such that

$$\begin{aligned} \int_{B_{h+1}} \mathrm{lip}_a(u_{m,h+1})\,\mathrm{d}\mathfrak{m} \leq \int_{D_{h+1}} \mathrm{lip}_a(\psi_{k(h+1),h+1})\,\mathrm{d}\mathfrak{m} + \int_{B_h} \mathrm{lip}_a(u_{m,h})\,\mathrm{d}\mathfrak{m} \\ + c_h \int_{\overline{H}_h} |\psi_{k(h+1),h+1} - u_{m,h}|\,\mathrm{d}\mathfrak{m}, \end{aligned}$$

$$\begin{cases} u_{m,h+1} \equiv u_{m,h} & \text{on } B_h \setminus D_{h+1} \supseteq B_{h-1} \\ u_{m,h+1} \equiv \psi_{k(h+1),h+1} & \text{on } D_{h+1} \setminus B_h \supseteq B_{h+1} \setminus \overline{B}_h. \end{cases}$$

By the induction assumption, $u_{m,h} \equiv \psi_{k(h),h}$ on $B_h \setminus \overline{B}_{h-1}$ which contains $\overline{H}_h$, and so we can use (4.4.8) to get (4.4.12). Then (4.4.14) of Lemma 4.4.3 with $\sigma = u$ tells us exactly that

$$\int_{B_{h+1}} |u_{m,h+1} - u|\,\mathrm{d}\mathfrak{m} \leq \int_{D_{h+1}} |\psi_{k(h+1),h+1} - u|\,\mathrm{d}\mathfrak{m} + \int_{B_h} |u_{m,h} - u|\,\mathrm{d}\mathfrak{m}$$

and so by (4.4.8) and the induction assumption we get also (4.4.11):

$$\int_{B_{h+1}} |u_{m,h+1} - u|\,\mathrm{d}\mathfrak{m} \leq \frac{1}{m2^{h+1}} + \frac{1}{m}\left(1 - \frac{1}{2^h}\right) = \frac{1}{m}\left(1 - \frac{1}{2^{h+1}}\right).$$

Now we prove (ii). Having already proved (iii), suffices to show that

$$|Du|_*(A_1' \cup A_2') \leq |Du|_*(A_1) + |Du|_*(A_2) \qquad \text{whenever } A_1' \Subset A_1,\ A_2' \Subset A_2.$$

This inequality can be obtained by applying Lemma 4.4.3 to join optimal sequences for A_1 and A_2, with $M = (A_1' \cup A_2') \cap A_1$ and $N = (A_1' \cup A_2') \cap A_2$. $\square$

Lemma 4.4.3 (Joint lemma) *Let M, N be open sets such that $\mathsf{d}(N\setminus M, M\setminus N) > 0$. There exist an open set $H \Subset M \cap N$ and a constant c depending only on M and N such that for every $u \in \mathrm{Lip}_{\mathrm{loc}}(M)$, $v \in \mathrm{Lip}_{\mathrm{loc}}(N)$ we can find $w \in \mathrm{Lip}_{\mathrm{loc}}(M \cup N)$ such that*

$$\int_{M\cup N} \mathrm{lip}_a w \,\mathrm{d}\mathfrak{m} \le \int_M \mathrm{lip}_a u \,\mathrm{d}\mathfrak{m} + \int_N \mathrm{lip}_a v \,\mathrm{d}\mathfrak{m} + c(M,N) \int_{\overline{H}} |u-v| \,\mathrm{d}\mathfrak{m}; \tag{4.4.13}$$

$$w \equiv u \text{ on neighborhood of } M \setminus N, \qquad w \equiv v \text{ on neighborhood of } N \setminus M.$$

Furthermore, for every $\sigma \in L^1(M \cup N)$ we have

$$\int_{M\cup N} |w-\sigma| \,\mathrm{d}\mathfrak{m} \le \int_M |u-\sigma| \,\mathrm{d}\mathfrak{m} + \int_N |v-\sigma| \,\mathrm{d}\mathfrak{m}. \tag{4.4.14}$$

Proof. The assumption on M and N guarantees the existence of a Lipschitz function $\varphi : X \to [0,1]$ such that

$$\varphi(x) = \begin{cases} 1 & \text{on a neighborhood of } M \setminus N \\ 0 & \text{on a neighborhood of } N \setminus M, \end{cases}$$

so that $H := \{0 < \varphi < 1\} \cap (M \cup N)$ will be an open set contained in $M \cap N$ and well separated from both $M \setminus N$ and $N \setminus M$. Setting $\eta := \mathsf{d}(N \setminus M, M \setminus N)$, it is clear that we can have $\mathrm{Lip}(\varphi) \le 3/\eta$; for example we can take

$$\varphi(x) := \frac{3}{\eta} \min\left\{ \left(\mathsf{d}(x, N\setminus M) - \frac{\eta}{3}\right)^+, \frac{\eta}{3} \right\}.$$

Now we consider the function $w = \varphi u + (1-\varphi)v$ and, using the convexity inequality for the asymptotic Lipschitz constant $\mathrm{lip}_a w \le \varphi \mathrm{lip}_a u + (1-\varphi)\mathrm{lip}_a v + \mathrm{lip}_a \varphi \cdot |u-v|$ (see Lemma 1.3.2 for the simple proof regarding also the slope) and the fact that $\varphi \le \chi_M$ and $1-\varphi \le \chi_N$ on $M \cup N$, splitting the integration on the interior of $\{\varphi = 1\}$, the interior of $\{\varphi = 0\}$ and $\overline{H}$ we end up with:

$$\int_{M\cup N} \mathrm{lip}_a w \,\mathrm{d}\mathfrak{m} \le \int_M \mathrm{lip}_a u \,\mathrm{d}\mathfrak{m} + \int_N \mathrm{lip}_a v \,\mathrm{d}\mathfrak{m} + \frac{3}{\eta} \int_{\overline{H}} |u-v| \,\mathrm{d}\mathfrak{m}.$$

To prove (4.4.14) we simply note that $|w-\sigma| \le \varphi|u-\sigma| + (1-\varphi)|v-\sigma|$ on $M \cup N$. □

4.4.3 Weak-BV functions

Before introducing the third definition we introduce some additional notation and terminology.

Definition 4.4.4 *A measure $\boldsymbol{\pi} \in \mathcal{P}(C([0,1];X))$ is said to be an ∞-*test plan *if the following two properties are satisfied:*

(a) $\boldsymbol{\pi}$ is concentrated on $AC^\infty([0,1];X)$ and $\mathrm{Lip}(\gamma)$ belongs to $L^\infty(C([0,1];X),\boldsymbol{\pi})$;

(b) there exists $C = C(\boldsymbol{\pi}) \ge 0$ such that $(e_t)_\sharp \boldsymbol{\pi} \le C\mathfrak{m}$ for each $t \in [0,1]$.

A Borel subset Γ of $C([0,1];X)$ is said to be 1-negligible if $\boldsymbol{\pi}(\Gamma) = 0$ for every ∞-plan $\boldsymbol{\pi}$. A property of continuous curves is said to be true 1-almost everywhere if the set for which it is false is contained in a 1-negligible set.

This definition is the degenerate case $\Psi = L$ of Definition 3.2.1 (again, this is an homogeneous case and so we can let $\boldsymbol{\pi}$ be a probability measure and $C(\boldsymbol{\pi})$ can be also less than 1). We recall now the definition of weak-BV, suggested in [9] and adopted in [5]. This follows the strong Beppo Levi definition: given a function f in $L^1(X, \mathfrak{m})$, we say that f is a weak-BV function, and write $f \in w - BV(X, \mathsf{d}, \mathfrak{m})$, if the following two conditions are fulfilled:

(i) for 1-almost every curve we have that $f \circ \gamma \in BV(0,1)$; we require also a mild regularity at the boundary, namely

$$|f(\gamma_1) - f(\gamma_0)| \leq |D(f \circ \gamma)|(0,1) \qquad \text{for 1-a.e. } \gamma, \tag{4.4.15}$$

where $|D(f \circ \gamma)| \in \mathcal{M}_+((0,1))$ is the total variation measure of the map $f \circ \gamma : [0,1] \to \mathbb{R}$;

(ii) there exists $\mu \in \mathcal{M}_+(X)$ such that

$$\int \gamma_\sharp |D(f \circ \gamma)|(B)\, d\boldsymbol{\pi}(\gamma) \leq C(\boldsymbol{\pi}) \cdot \|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi})} \mu(B) \qquad \forall B \in \mathcal{B}(X). \tag{4.4.16}$$

Associated to this notion, there is also the concept of weak total variation $|Df|_w$, defined as the least measure μ satisfying (4.4.16) for every ∞-test plan $\boldsymbol{\pi}$. Equivalently, $|Df|_w$ is the least upper bound, in the complete and separable lattice $\mathcal{M}_+(X)$, of the family of measures

$$\frac{1}{C(\boldsymbol{\pi})\|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi})}} \int \gamma_\sharp |D(f \circ \gamma)|\, d\boldsymbol{\pi}(\gamma) \tag{4.4.17}$$

as $\boldsymbol{\pi}$ runs in the class of ∞-test plans.

If we fix $t \in (0,1)$ and we consider the rescaling map R_t from $C([0,1], X)$ to $C([0,1], X)$ mapping γ_s to γ_{ts}, we see that the push-forward $\boldsymbol{\pi}_t = (R_t)_\sharp \boldsymbol{\pi}$ is still a ∞-test plan, with $C(\boldsymbol{\pi}_t) \leq C(\boldsymbol{\pi})$. In addition

$$\|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi}_t)} \leq t \|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi})}.$$

By (4.4.15) we get

$$|f(\gamma_t) - f(\gamma_0)| \leq |D(f \circ \gamma)|(0,t) \qquad \text{for } \boldsymbol{\pi}\text{-a.e. } \gamma, \tag{4.4.18}$$

while (4.4.16) with $A = X$ gives

$$\begin{aligned} \int |D(f \circ \gamma)|(0,t)\, \mathrm{d}\boldsymbol{\pi}(\gamma) &= \int |D(f \circ \gamma)|(0,1)\, \mathrm{d}\boldsymbol{\pi}_t(\gamma) \\ &\leq t C(\boldsymbol{\pi}) \|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi})} |Df|_w(X). \end{aligned} \tag{4.4.19}$$

Now we prove that the class BV_* is contained in the class $w - BV$ and that $|Df|_w \leq |Df|_*$ on open sets. The proof of this fact is not difficult, and follow the same lines of the proofs in Theorem 3.4.1.First of all, we state without proof the following elementary lemma:

Lemma 4.4.5 *Assume that g is an upper gradient of f, that $\gamma : [0,1] \to X$ is Lipschitz and that $\int_\gamma g < \infty$. Then $f \circ \gamma \in W^{1,1}(0,1)$ and $|(f \circ \gamma)'(t)| \leq g(\gamma_t)|\dot{\gamma}_t|$ for a.e. $t \in (0,1)$. In particular*

$$|D(f \circ \gamma)|(B) \leq \mathrm{Lip}(\gamma) \int_B g(\gamma_t)\, \mathrm{d}t \qquad \textit{for any Borel set } B \subseteq (0,1).$$

Given an open set $A \subseteq X$, we take a sequence $(f_n) \subset \mathrm{Lip}_{loc}(A)$ such that $f_n \to f$ in $L^1(A, \mathfrak{m})$ and $\int_A \mathrm{lip}_a(f_n) \mathrm{d}\mathfrak{m} \to |Df|_*(A)$ (whose existence is granted by the definition of relaxed total variation), and use the lemma and the fact that $\mathrm{lip}_a(f_n)$ is an upper gradient for f_n to estimate the weak total variation of f_n as follows:

$$\begin{aligned} \int \gamma_\sharp |D(f_n \circ \gamma)|(A) \, \mathrm{d}\boldsymbol{\pi}(\gamma) &= \int |D(f_n \circ \gamma)|(\gamma^{-1}(A)) \, \mathrm{d}\boldsymbol{\pi}(\gamma) \\ &\leq \int \mathrm{Lip}(\gamma) \int_0^1 g_n(\gamma_t) \chi_A(\gamma_t) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{\pi}(\gamma) \\ &\leq \|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi})} \int_0^1 \int_A g_n \, \mathrm{d}(\mathrm{e}_t)_\sharp \boldsymbol{\pi} \, \mathrm{d}t \\ &\leq \|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi})} C(\boldsymbol{\pi}) \int_A g_n \, \mathrm{d}\mathfrak{m}. \end{aligned} \tag{4.4.20}$$

We now introduce a lemma that permits us, up to a subsequence, to localize the L^1 convergence, so that we can estimate the left hand side of the weak upper gradient inequality.

Lemma 4.4.6 *Let $B \subseteq X$ be a Borel set and let (f_n) be a sequence converging to f in $L^1(B, \mathfrak{m})$. Then, a subsequence of (f_n) converges to f in $L^1(\gamma^{-1}(B), \mathscr{L}^1)$ along 1-almost every curve.*

Proof. We can assume without loss of generality that $B = X$. Possibly extracting a subsequence, we can suppose that

$$\sum_n \|f_n - f\|_{L^1(X,\mathfrak{m})} < \infty.$$

We now fix a ∞-test plan $\boldsymbol{\pi}$ and we show that $\|f_n \circ \gamma - f \circ \gamma\|_{L^1(0,1)} \to 0$ for $\boldsymbol{\pi}$-almost every curve γ. Our choice of the subsequence ensures that the function $g := \sum_n |f_n - f|$ belongs to $L^1(0,1)$. Now, the inequality

$$\int \|g \circ \gamma\|_{L^1(0,1)} \, \mathrm{d}\boldsymbol{\pi}(\gamma) = \iint_0^1 (g \circ \gamma)(t) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{\pi} \leq C(\boldsymbol{\pi}) \int_0^1 \int_X g \, \mathrm{d}\mathfrak{m} < \infty$$

guarantees that $g \circ \gamma$ belongs to $L^1(0,1)$ for $\boldsymbol{\pi}$-a.e. curve γ and thus we can say that $f_n \circ \gamma \to f \circ \gamma$ in $L^1(0,1)$ for $\boldsymbol{\pi}$-a.e. γ. By the arbitrariness of $\boldsymbol{\pi}$, we conclude. $\square$

We can now complete the proof of $|Df|_w \leq |Df|_*$ on open sets, starting from (4.4.20).

Let $A \subseteq X$ be an open set, let $(f_n) \subset \mathrm{Lip}_{loc}(A)$ be a sequence convergent to f in $L^1(A)$ such that

$$\lim_{n\to\infty} \int_A \mathrm{lip}_a(f_n) \, d\mathfrak{m} = |Df|_*(A).$$

Thanks to Lemma 4.4.6 we can find a subsequence $n(s)$ such that $f_{n(s)} \circ \gamma \to f \circ \gamma$ in $L^1(\gamma^{-1}(A))$ along 1-almost every curve γ. By (4.1.1) in the open set $\gamma^{-1}(A)$ we get

$$\gamma_\sharp |D(f \circ \gamma)|(A) \leq \liminf_{s\to\infty} \gamma_\sharp |D(f_{n(s)} \circ \gamma)|(A) \qquad \text{for } \boldsymbol{\pi}\text{-a.e. curve } \gamma.$$

Passing to the limit as $s \to \infty$ in the inequality (4.4.20) with $n = n(s)$, Fatou's lemma gives $\mu_{\boldsymbol{\pi}}(A) \leq |Df|_*(A)$ for all ∞-test plan $\boldsymbol{\pi}$, where $\mu_{\boldsymbol{\pi}}$ is the finite Borel measure in (4.4.17). If

$\pi_1, \dots, \pi_k$ is a finite collection of ∞-test plans, the formula

$$\bigvee_{i=1}^{k} \mu_{\pi_i}(A) = \sup\left\{\sum_{i=1}^{k} \mu_{\pi_i}(A_i) :\ A_1 \subseteq A, \dots, A_k \subseteq A \text{ open, pairwise disjoint}\right\}$$

and the additivity of $|Df|_*$ yield $|Df|_*(A) \geq \vee_1^k \mu_{\pi_i}(A)$ for any open set A. Since this collection is arbitrary, the inequality $|Df|_w(A) \leq |Df|_*(A)$ is proved.

We're not done yet, because we have to prove also the boundary regularity (4.4.15) that is part of our axiomatization of $w - BV$ functions. The inequality would clearly follow if we show that $f \circ \gamma_i$, $i = 0, 1$, is the approximate limit of $f \circ \gamma$ as $t \to i$, namely

$$\lim_{t \downarrow 0} \frac{1}{t} \int_0^t |f(\gamma_s) - f(\gamma_0)| \,\mathrm{d}s = 0, \qquad \lim_{t \downarrow 0} \frac{1}{t} \int_{1-t}^1 |f(\gamma_s) - f(\gamma_1)| \,\mathrm{d}s = 0.$$

This is indeed the context of the next lemma, that we state and prove for $t = 0$ only:

Lemma 4.4.7 (Boundary regularity) *We are given a sequence $(f_n) \subset \mathrm{Lip}_{\mathrm{loc}}(X)$ where $f_n \to f$ in $L^1(X, \mathfrak{m})$ and $\sup_n \int_X \mathrm{lip}_a(f_n) \mathrm{d}\mathfrak{m} < \infty$. Then $t = 0$ is a Lebesgue point for the map $f \circ \gamma : [0, 1] \to \mathbb{R}$ for 1-almost every curve γ.*

Proof. Let us fix an ∞-plan $\boldsymbol{\pi}$, set $C_1 := \sup_n \int_X g_n \mathrm{d}\mathfrak{m}$, $C_2 := C(\boldsymbol{\pi})$ and consider the quantities

$$H_t(\gamma) = \frac{1}{t} \int_0^t |f(\gamma_s) - f(\gamma_0)| \mathrm{d}s.$$

By definition, we know that 0 is a Lebesgue point for $f \circ \gamma$ if $H_t(\gamma) \to 0$ as $t \to 0$. Applying Fatou's lemma we get:

$$\int \liminf_{t \to 0} H_t(\gamma) \mathrm{d}\boldsymbol{\pi} \leq \liminf_{t \to 0} \int H_t(\gamma) \,\mathrm{d}\boldsymbol{\pi}. \tag{4.4.21}$$

We can estimate

$$\int H_t(\gamma) \,\mathrm{d}\boldsymbol{\pi} \leq \int H_t^n(\gamma) \,\mathrm{d}\boldsymbol{\pi} + \frac{1}{t} \iint_0^t \Big(|f_n(\gamma_s) - f(\gamma_s)| + |f_n(\gamma_0) - f(\gamma_0)|\Big) \mathrm{d}s \,\mathrm{d}\boldsymbol{\pi},$$

where $H_t^n(\gamma) = \frac{1}{t} \int_0^t |f_n(\gamma_s) - f_n(\gamma_0)| \mathrm{d}s$. We now treat separately the two terms on the right: first let's note that

$$\begin{aligned}
\int H_t^n(\gamma) \mathrm{d}\boldsymbol{\pi} &= \frac{1}{t} \iint_0^t |f_n(\gamma_s) - f_n(\gamma_0)| \mathrm{d}s \,\mathrm{d}\boldsymbol{\pi} \leq \frac{1}{t} \iint_0^t \int_0^s \mathrm{lip}_a(f_n, \gamma_r) \mathrm{d}r \,\mathrm{d}s \,\mathrm{d}\boldsymbol{\pi} \\
&\leq \frac{1}{t} \iint_0^t \int_0^t \mathrm{lip}_a(f_n, \gamma_r) \mathrm{d}r \,\mathrm{d}s \,\mathrm{d}\boldsymbol{\pi} = \iint_0^t \mathrm{lip}_a(f_n, \gamma_r) \mathrm{d}r \,\mathrm{d}\boldsymbol{\pi} \\
&\leq C_2 \int_0^t \int_X \mathrm{lip}_a(f_n, x) \,\mathrm{d}\mathfrak{m}(x) \,\mathrm{d}\mathfrak{m} \,\mathrm{d}t \leq t C_1 C_2.
\end{aligned}$$

For the second term:

$$\begin{aligned}
&\frac{1}{t} \iint_0^t \Big(|f_n(\gamma_s) - f(\gamma_s)| + |f_n(\gamma_0) - f(\gamma_0)|\Big) \mathrm{d}s \,\mathrm{d}\boldsymbol{\pi} \\
&= \frac{1}{t} \iint_0^t |f_n(\gamma_s) - f(\gamma_s)| \mathrm{d}s \,\mathrm{d}\boldsymbol{\pi} + \int |f_n(\gamma_0) - f(\gamma_0)| \,\mathrm{d}\boldsymbol{\pi} \\
&\leq \frac{1}{t} \int_0^t \int_X |f_n - f| \cdot C_2 \,\mathrm{d}\mathfrak{m} \,\mathrm{d}s + \int_X |f_n - f| \cdot C_2 \,\mathrm{d}\mathfrak{m} \\
&\leq 2C_2 \cdot \|f_n - f\|_{L^1(X, \mathfrak{m})}.
\end{aligned}$$

summing up we get that, choosing n so large that $\|f_n - f\|_{L^1} \le t$,

$$\int H_t(\gamma)\mathrm{d}\boldsymbol{\pi} \le tC_2(C_1 + 2).$$

Now by (4.4.21) we conclude that $\int (\liminf_{t\to 0} H_t)\, \mathrm{d}\boldsymbol{\pi} = 0$ and so, thanks to the arbitrariness of $\boldsymbol{\pi}$, we can say that 0 is a Lebesgue point for 1-almost every curve (here we use also that $f \circ \gamma$ is a BV function for 1-a.e. curve, and in this case we know that H_t has a limit). □

4.4.4 BV_{BL} functions

Also here we give a weaker version of the $w - BV$ definition; we follow the (weak) Beppo Levi definition: we say that a function $f \in L^1(X, \mathfrak{m})$ is a BV_{BL} function if there exists a constant $E = E(f)$ such that for every ∞-plan we have:

$$\int f(\gamma_1) - f(\gamma_0)\, \mathrm{d}\boldsymbol{\pi} \le C(\boldsymbol{\pi}) \cdot E(f)\|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi})}. \tag{4.4.22}$$

We call $|Df|_{BL}(X)$ the least constant E such that (4.4.22) holds true for every ∞-plan $\boldsymbol{\pi}$. In the same spirit we can define $|Df|_{BL}(A)$ for open sets A:

$$|Df|_{BL}(A) = \sup \left\{ \frac{\int f(\gamma_1) - f(\gamma_0)}{C(\boldsymbol{\pi}) \cdot \|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi})}} \right\}, \tag{4.4.23}$$

where the supremum is taken among all ∞-plans concentrated on $\mathrm{AC}([0,1]; C)$ for some closed set $C \Subset A$. It is clear that a $w - BV$ function belongs to BV_{BL} and $|Df|_w(A) \ge |Df|_{BL}(A)$ for every open set A.

4.5 Proof of equivalence

In Section 4.4 we discussed the "easy" inclusions $BV_* \subseteq w - BV \subseteq BV_{BL}$, and the corresponding inequalities (localized on open subsets of X)

$$|Df|_{BL} \le |Df|_w \le |Df|_*.$$

Furthermore we proved also that $BV_v \subseteq BV_*$ and the corresponding total variation inequalities (on the whole space):

$$|Df|_{BL}(X) \le |Df|_w(X) \le |Df|_*(X) \le |Df|_v(X).$$

In this section we prove the main result of this chapter namely the equivalence of the four definitions. So, we have to start from a function $f \in BV_{BL}(X, \mathsf{d}, \mathfrak{m})$, and build a sequence of approximating Lipschitz functions with bounded support in such a way that

$$\limsup_{n\to\infty} \int_X \mathrm{lip}_a(f_n)\, \mathrm{d}\mathfrak{m} \le |Df|_{BL}(X). \tag{4.5.1}$$

As in [11] for the case $q = 2$ and [9] for the case $1 < q < \infty$, our main tool in the construction will the gradient flow in $L^2(X, \mathfrak{m})$ of the functional $\mathcal{F}^1_v(f) = |Df|_v(X)$, starting from f_0. We initially assume that (X, d) is a complete and separable length space (this assumption is used to be able to apply the results of Section 4.2 and in Lemma 4.5.2, to apply (4.3.1)) and that $\mathfrak{m}$ is a finite Borel measure, so that the L^2-gradient flow of $\mathcal{F}^1_v$ can be used. The finiteness and length space assumptions will be eventually removed in the proof of the equivalence result.

We state for completeness the proposition and the lemma we need for the proof; both are proved in the general Orlicz case in Section 3.3. The only things that really change is that in the Kuwada lemma we have to use the Hamilton-Jacobi inequality (4.2.6) (proved for the slope) in (3.3.6), and then use that $|Df|_v(X) \leq \int_X |\nabla f|$ for $f \in \mathrm{Lip}_0(X, \mathsf{d})$ (see Proposition 4.4.1) in (3.3.7).

Proposition 4.5.1 *Let $\mu_t = f_t\mathfrak{m}$ be a curve in $AC^\infty([0,1], (\mathcal{M}_+(X), W_\infty))$. Assume that for some $0 < c < C < \infty$ it holds $c \leq f_t \leq C$ $\mathfrak{m}$-a.e. in X for any $t \in [0,1]$, and that $f_0 \in w - BV(X, \mathsf{d}, \mathfrak{m})$. Then for all $\Phi \in C^2([c, C])$ convex it holds*

$$\int \Phi(f_0)\,\mathrm{d}\mathfrak{m} - \int \Phi(f_s)\,\mathrm{d}\mathfrak{m} \leq s\mathrm{Lip}(\Phi')|Df_0|_w(X) \cdot C \cdot \mathrm{Lip}(\mu_t) \qquad \forall s > 0.$$

Lemma 4.5.2 (Kuwada's lemma for $\mathcal{F}^1_v$) *Let $f_0 \in L^2(X, \mathfrak{m})$ and let (f_t) be the gradient flow of $\mathcal{F}^1_v$ starting from f_0. Assume that for some $0 < c < C < \infty$ it holds $c \leq f_0 \leq C$ $\mathfrak{m}$-a.e. in X. Then the curve $t \mapsto \mu_t := f_t\mathfrak{m} \in \mathcal{M}_+(X)$ is absolutely continuous w.r.t. W_∞ and its W_∞ metric derivative satisfies*

$$|\dot\mu_t| \leq \frac{1}{c} \qquad \text{for a.e. } t \in (0, \infty).$$

We can now prove our main theorem:

Theorem 4.5.3 *Let (X, d) be a separable metric space, and let $\mathfrak{m}$ be a nonnegative Borel measure on X, finite on bounded sets. Then we have*

$$BV_v(X, \mathsf{d}, \mathfrak{m}) = BV_*(X, \mathsf{d}, \mathfrak{m}) = w - BV(X, \mathsf{d}, \mathfrak{m}) = BV_{BL}(X, \mathsf{d}, \mathfrak{m}),$$

and moreover for every function $f \in BV_*(X, \mathsf{d}, \mathfrak{m})$ we have $|Df|_* = |Df|_w = |Df|_{BL}$ as measures and $|Df|_*(X) = |Df|_v(X)$.

Proof. Recalling the results of Section 4.4, to conclude the proof we are only left to show that a function in BV_{BL} is also a function of variational total variation and the two definitions of total variations coincides; at the end we will see also the coincidence of the measures $|Df|_*, |Df|_w$ and $|Df|_{BL}$. We first prove that $|Df|_v(X) \leq |Df|_{BL}(X)$ and then that the set functions $|Df|_*$ and $|Df|_{BL}$ agree on all open sets. This yields the coincidence of the three aforementioned measures on the Borel σ-algebra.

We split the proof of the inequality $|Df|_v(X) \leq |Df|_{BL}(X)$ in three parts: we prove it first for bounded functions and finite measures in length spaces, then we remove the boundedness assumption on f and the length space assumption, and eventually the finiteness assumption on $\mathfrak{m}$. Notice that we will follow the same lines of the equivalence proof in Section 3.3.
Let us consider a bounded function $f_0 \in BV_{BL}$ possibly adding a constant (that doesn't change any of the total variations; the unique non trivial being $|Df|_v$, that doesn't change thanks to Proposition 3.1.3) we can suppose also that $C \geq f_0 \geq c > 0$. Let us consider as before the gradient flow f_t in $L^2(X, \mathfrak{m})$, with respect to $\mathcal{F}^1_v$, starting from f_0. Now, let $\Phi(x) = x^2$, so that $\Phi'' \equiv 2$, and let's substitute f_0 with $f_0 + H$; our computation is left unchanged, because we know that $S_t(f_0 + H) = f_t + H$ and so we can say, using the energy

estimate in Proposition 3.1.5 and the Lipschitz estimate for the curve $t \mapsto (f_t + H)\mathfrak{m}$ given by Lemma 4.5.2, combined with Proposition 4.5.1:

$$\begin{aligned} 2\int_0^s |Df_t|_v(X)\,\mathrm{d}t &= 2\int_0^s |D(f_t + H)|_v(X)\,\mathrm{d}t \\ &= \int_X (f_0 + H)^2\,\mathrm{d}\mathfrak{m} - \int_X (f_s + H)^2\,\mathrm{d}\mathfrak{m} \\ &\le 2s \cdot |Df_0|_{BL}(X) \cdot \frac{C + H}{c + H}. \end{aligned}$$

Now, letting $H \to \infty$, we get that

$$\int_0^s |Df_t|_v(X)\,\mathrm{d}t \le s \cdot |Df_0|_{BL}(X).$$

But, knowing that $|Df_t|_v(X) = \mathcal{F}_v^1(f_t)$ is nonincreasing in t we can say

$$s \cdot |Df_s|_v(X) \le \int_0^s |Df_t|_v(X)\,\mathrm{d}t \le s \cdot |Df_0|_{BL}(X)$$

and thus $|Df_s|_v(X) \le |Df_0|_{BL}(X)$. Now we have that $|Df|_v$ is lower semicontinuous and so, letting $s \downarrow 0$, we obtain that $f_0 \in BV_v(X, \mathsf{d}, \mathfrak{m})$ and that $|Df_0|_v(X) \le |Df_0|_{BL}(X)$.

Now, taking any function $g \in BV_{BL}(X, \mathsf{d}, \mathfrak{m})$, defining $g^N = (g \wedge N) \vee (-N)$, we have $g^N \to g$ in L^1 as N goes to infinity; thanks to Lemma 3.3.2 we get

$$|Dg|_w(X) = \limsup |Dg^N|_w(X) = \limsup |Dg^N|_*(X) = |Dg|_*(X).$$

Now, still assuming $\mathfrak{m}$ to be finite, we see how the length space assumption on X can be easily removed. Indeed, it is not difficult to find an isometric embedding of (X, d) into a complete, separable and length metric space (Y, d_Y): for instance one can use the canonical Kuratowski isometric embedding j of (X, d) into ℓ_∞ and then take as Y the closed convex hull of $j(X)$. For notational simplicity, just assume that $X \subseteq Y$ and that d_Y restricted to $X \times X$ coincides with d. Since X is a closed subset of Y, we may also view $\mathfrak{m}$ as a finite Borel measure in Y supported in X. Then, if $f \in BV_{BL}(X, \mathsf{d}, \mathfrak{m})$, we have also $f \in BV_{BL}(Y, \mathsf{d}_Y, \mathfrak{m})$ and $|Df|_{BL,Y}(Y) \le |Df|_{BL}(X)$, because any ∞-test plan $\boldsymbol{\pi}$ in Y is, by the condition $(\mathrm{e}_t)_\sharp\boldsymbol{\pi} \le \mathfrak{m}$, supported on Lipschitz curves with values in X. Then, applying the equivalence result in $(Y, \mathsf{d}_Y, \mathfrak{m})$, we find a sequence of Lipschitz functions wit bounded support $g_n : Y \to \mathbb{R}$ convergent to f in $L^1(Y, \mathfrak{m})$ satisfying

$$\limsup_{n\to\infty} \int_Y \mathrm{lip}_a(g_n)\,\mathrm{d}\mathfrak{m} \le |Df|_{BL,Y}(Y) \le |Df|_{BL}(X).$$

Now, if $f_n = g_n|_X$, from the inequality $\mathrm{lip}_a(f_n) \le \mathrm{lip}_a(g_n)$ on X we obtain $\limsup_n \int_X \mathrm{lip}_a(f_n)\,\mathrm{d}\mathfrak{m} \le |Df|_{BL}(X)$. On the other hand, it is immediate to check that f_n are Lipschitz in X, with bounded support.

In order to prove the theorem also for measures $\mathfrak{m}$ that are finite on bounded sets we proceed as in Section 3.3, namely, we know by Lemma 3.3.2 that we need only to check the equivalence on bounded and integrable functions; then via Lemma 3.3.3 we reduce to bounded function with bounded support. Now again Lemma 3.3.3 let us conclude that for every closed bounded set C (in which we consider the finite measure $\mathfrak{m}_C(B) = \mathfrak{m}(C \cap B)$), we have $|Df|_{v,X}(X) = |Df|_{v,C}(C)$ whenever $\mathrm{supp}(f) \subset B_r \subset B_{r+4} \subset C$. We can apply the equivalence on C and note that $|Df|_{BL,X}(X) \geq |Df|_{BL,C}(C)$ since ∞-test plan in C can be viewed also as ∞-test plans in X, to get

$$|Df|_v(X) = |Df|_{v,C}(C) = |Df|_{BL,C}(C) \leq |Df|_{BL}(X),$$

and so our proof is complete.

Now it is easy to conclude that we have also $|Df|_*(X) = |Df|_w(X) = |Df|_{BL}$. Moreover for every couple of open sets $B \Subset A$, by definition of $|Df|_{BL}(A)$, we have

$$|Df|_{BL}(A) \geq |Df|_{BL,\bar{B}}(\bar{B}) = |Df|_{*,\bar{B}}(\bar{B}) \geq |Df|_{*,\bar{B}}(B) = |Df|_*(B);$$

using this inequality with $B = A_\delta = \{x \ : \ \mathsf{d}(x, A^c) > \delta\}$ and letting $\delta \to 0$ we know that $A_\delta \uparrow A$ and so, using (iii) of Lemma 4.4.2, we obtain $|Df|_{BL}(A) \geq |Df|_*(A)$.

In particular $|Df|_{BL}$, $|Df|_*$ and $|Df|_{BL}$ agree on open sets, hence the two measures $|Df|_*$ and $|Df|_w$ coincide and as a byproduct we obtain also that $|Df|_{BL}$ can be extended to a measure. □

The following example shows that in general the sup representation does not extend to the absolutely continuous parts.

Example 4.5.4 Let $X = \mathbb{R}^2$, let B be the closed unit ball in $\mathbb{R}^2$, d the Euclidean distance and $\mathfrak{m}(C) = \mathscr{L}^2(C) + \mathscr{H}^1(C \cap \partial B)$, for $C \subseteq X$ Borel. If f is the characteristic function of B, the inequality $\mathfrak{m} \geq \mathscr{L}^2$ gives the inequality between measures $|Df| \leq |Df|_w$. We claim that the two measures coincide. To see this, suffices to show that $|Df|_w(\mathbb{R}^2) \leq 2\pi$ and this inequality follows easily by considering the sequence of functions (each one constant in a neighbourhood of ∂B) $f_n(x) = \varphi_n(|x|)$ with

$$\varphi_n(t) := \begin{cases} 1 & \text{if } t \leq 1 + \frac{1}{n}; \\ 1 - n\big(t - 1 - \frac{1}{n}\big) & \text{if } 1 + \frac{1}{n} < t \leq 1 + \frac{2}{n}; \\ 0 & \text{if } t > 1 + \frac{2}{n}. \end{cases}$$

Since $|Df|(C) = \mathscr{H}^1(C \cap \partial B)$, it follows that $|Df|_w$ is absolutely continuous w.r.t. $\mathfrak{m}$; on the other hand, since f is a characteristic function the same is true for the maps $f \circ \gamma$, so that $|D^a(f \circ \gamma)| = 0$ whenever $f \circ \gamma$ has bounded variation.

We conclude this section with the following corollary to Theorem 4.5.3, dealing with the degenerate case $L^1 = BV$; similar results could be stated also at the level of the Sobolev spaces $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$ and the corresponding test plans of [9].

Corollary 4.5.5 *$BV_v(X, \mathsf{d}, \mathfrak{m})$ coincides with $L^1(X, \mathfrak{m})$ if and only if $(X, \mathsf{d}, \mathfrak{m})$ has a ∞-test plan concentrated on nonconstant rectifiable curves. In addition, (X, d) contains one nonconstant rectifiable curve if and only there exists a finite Borel measure $\mathfrak{m}$ in (X, d) satisfying $BV_v(X, \mathsf{d}, \mathfrak{m}) \neq L^1(X, \mathfrak{m})$.*

Proof. In the first statement, the "only if" part is trivial, since absence of ∞-test plans implies that all L^1 functions are BV_w, and therefore BV_v. In order to prove the converse, we notice that for a given countable dense set $D \subset X$, a curve γ is constant iff $t \mapsto \mathsf{d}(\gamma, x)$ is constant for all $x \in D$. Hence, we can find $x \in D$ and a ∞-test plan $\boldsymbol{\pi}$ such that $\mathsf{d}(\gamma, x)$ is nonconstant in a set with $\boldsymbol{\pi}$-positive measure. The composition

$$f(y) := w(\mathsf{d}(y, x)),$$

where $w : [0, \infty) \to [0, 1]$ is a continuous and nowhere differentiable function, provides a function in $L^1 \setminus BV_w = L^1 \setminus BV_v$.

For the second statement, absence of nonconstant rectifiable curves forces the absence of nontrivial ∞-test plans whatever $\mathfrak{m}$ is and, for the reasons explained above, the coincidence $L^1 = BV_v$. On the other hand, existence of a nonconstant rectifiable curve in (X, d) implies existence of a nonconstant injective curve $\gamma : [0, 1] \to X$ with constant speed. If $u \in L^1(0, 1) \setminus BV(0, 1)$, then it is easily seen that $u \circ \gamma^{-1}$ (arbitrarily defined on $X \setminus \gamma([0, 1])$) belongs to $L^1 \setminus BV_w$ provided we choose $\mathfrak{m} := \gamma_\sharp \mathscr{L}^1$, where $\mathscr{L}^1$ is the restriction of Lebesgue measure to $[0, 1]$. $\square$

We end this section giving a useful characterization of the total variation as weak limit of asymptotic Lipschitz constants.

Proposition 4.5.6 *Let $f \in BV(X, \mathsf{d}, \mathfrak{m})$. Then there exists a sequence $(f_n) \subset \mathrm{Lip}_0(X, \mathsf{d})$ such that $f_n \to f$ in $L^1(X, \mathfrak{m})$ and $\mathrm{lip}_a(f_n)\mathfrak{m} \rightharpoonup |Df|_*$ in duality with $C_b(X)$. In particular for every function $g \in C_b(X)$ we have*

$$\lim_{n\to\infty} \int_X g(x) \mathrm{lip}_a(f_n, x) \, \mathrm{d}\mathfrak{m} = \int_X g(x) \, \mathrm{d}|Df|_*$$

Proof. Let $(f_n) \subset \mathrm{Lip}_0(X, \mathsf{d})$ any optimal sequence in (4.4.1), i.e. such that

$$|Df|_*(X) = \lim_{n\to\infty} \int_X \mathrm{lip}_a(f_n) \, \mathrm{d}\mathfrak{m}. \tag{4.5.2}$$

We have clearly that $f_n \to f$ also in $L^1(A, \mathfrak{m})$ and $f_n \in \mathrm{Lip}_{\mathrm{loc}}(A, \mathsf{d})$ and so by definition (4.4.3) we have also

$$|Df|_*(A) \le \liminf_{n\to\infty} \int_A \mathrm{lip}_a(f_n) \, \mathrm{d}\mathfrak{m}; \tag{4.5.3}$$

now standard measure theory gives us the thesis, in fact if we have $\mu(A) \le \liminf_n \mu_n(A)$ for every open set A and $\mu(X) = \lim_n \mu_n(X)$ then $\mu_n \rightharpoonup \mu$. For the sake of completeness we show also this fact. It is clear that we have also $\mu(C) \ge \limsup_n \mu_n(C)$ for every closed set C.

For every nonnegative μ-integrable function g we have

$$\int_X g(x) \, \mathrm{d}\mu(x) = \int_0^\infty \mu\{g \ge t\} \, \mathrm{d}x = \int_0^\infty \mu\{g > t\} \, \mathrm{d}x;$$

this formula is easy to prove when g is simple and then it follows by approximation. In particular it holds when g is continuous and bounded. In this case, using that $\{g > t\}$ is an open set, we can employ Fatou lemma to obtain

$$\int_X g \, \mathrm{d}\mu = \int_0^\infty \mu\{g > t\} \, \mathrm{d}x \le \liminf_n \int_0^\infty \mu_n\{g > t\} \, \mathrm{d}x = \liminf_n \int_X g \, \mathrm{d}\mu_n;$$

on the other side we also have

$$\int_X g \, \mathrm{d}\mu = \int_0^\infty \mu\{g \ge t\} \, \mathrm{d}x \ge \limsup_n \int_0^\infty \mu_n\{g \ge t\} \, \mathrm{d}x = \limsup_n \int_X g \, \mathrm{d}\mu_n.$$

and so we have $\lim_n \int_X g \, \mathrm{d}\mu_n = \int_X g \, \mathrm{d}\mu$. Since g was an arbitrary nonnegative continuous bounded function, we get the thesis. $\square$

4.6 Possible definitions for $W^{1,1}(X, \mathsf{d}, \mathfrak{m})$

In this section we discuss potential definitions of the space $W^{1,1}$. Here the picture is far from being complete, since at least three definitions are available and we are presently not able to prove their equivalence, unlike for BV. For simplicity, here we assume that $(X, \mathsf{d}, \mathfrak{m})$ is a compact metric space and that $\mathfrak{m}$ is a probability measure. Recall that $BV_v(X, \mathsf{d}, \mathfrak{m})$ denotes the BV space defined by relaxation of the asymptotic Lipschitz constant of Lipschitz functions, while $w - BV(X, \mathsf{d}, \mathfrak{m})$ is the BV space defined with the BV property along curves.

It is immediate to look at the subset of $w - BV(X, \mathsf{d}, \mathfrak{m})$ consisting of functions $f \in L^1(X, \mathsf{d}, \mathfrak{m})$ such that $|Df|_w \ll \mathfrak{m}$. However this is not satisfactory since Example 4.5.4 gives an example of a characteristic function with non trivial gradient that belongs to this subset. For this reason we will add the condition of being absolutely continuous along 1-almost every curve.

Definition 4.6.1 ($W^{1,1}_{BL}$ space) *A function $f \in w - BV(X, \mathsf{d}, \mathfrak{m})$ is said to belong to $W^{1,1}_{BL}(X, \mathsf{d}, \mathfrak{m})$ if the following conditions are satisfied:*

(i) $f \circ \gamma$ belongs to $W^{1,1}(0, 1)$ for 1-almost every curve γ;

(ii) $|Df|_{BL} \ll \mathfrak{m}$.

But we can also provide a definition in the spirit of the weak upper gradient definition:

Definition 4.6.2 (1-weak upper gradient) *A function $g \in L^1(X, \mathfrak{m})$ is said a 1-weak upper gradient for $f \in L^1(X, \mathfrak{m})$ if*

$$|f(\gamma_1) - f(\gamma_0)| \leq \int_\gamma g < \infty \qquad \text{for 1-a.e. } \gamma. \tag{4.6.1}$$

On the other hand, also the construction leading to $BV_v(X, \mathsf{d}, \mathfrak{m})$ (or to the relaxed Orlicz-Sobolev spaces) can be adapted to provide a different definition of $W^{1,1}$:

Definition 4.6.3 (1-relaxed slope) *Let $f \in L^1(X, \mathfrak{m})$. We say that a nonnegative function $g \in L^1(X, \mathfrak{m})$ is a 1-relaxed slope of f if there exist Lipschitz functions with bounded support f_n converging to f in $L^1(X, \mathfrak{m})$ such that $\mathrm{lip}_a(f_n) \rightharpoonup h$ weakly in $L^1(X, \mathfrak{m})$, with $g \geq h$ $\mathfrak{m}$-a.e. in X.*

Then, we may define $w - W^{1,1}(X, \mathsf{d}, \mathfrak{m})$ and $H^{1,1}(X, \mathsf{d}, \mathfrak{m})$ as the space of functions in $L^1(X, \mathsf{d}, \mathfrak{m})$ having a 1-weak upper gradient and a 1-relaxed slope, respectively. It is not difficult to show, using Mazur's lemma, that an equivalent definition of 1-relaxed slope g involves sequences f_n such that $\mathrm{lip}_a(f_n) \leq h_n$, with $h_n \to h$ strongly in $L^1(X, \mathfrak{m})$ and $h \leq g$. Then, this gives that $|Df|_w \leq h\mathfrak{m}$ for all $f \in H^{1,1}(X, \mathsf{d}, \mathfrak{m})$, so that

$$H^{1,1}(X, \mathsf{d}, \mathfrak{m}) \subseteq w - W^{1,1}(X, \mathsf{d}, \mathfrak{m}).$$

Finally, also a fourth intermediate definition of $W^{1,1}(X, \mathsf{d}, \mathfrak{m})$ could be considered, in the spirit of [57], [75], very similar to $w - W^{1,1}$.

Definition 4.6.4 (1-upper gradient) *A Borel nonnegative function $g \in L^1(X, \mathsf{d}, \mathfrak{m})$ is said to be a 1-upper gradient of $f \in L^1(X, \mathsf{d}, \mathfrak{m})$ if there exists a function $\hat{f}$ that coincides $\mathfrak{m}$-almost everywhere with f such that*

$$|\hat{f}(\gamma_1) - \hat{f}(\gamma_0)| \leq \int_\gamma g \, \mathrm{d}s \qquad \textit{for } \mathrm{Mod}_1\textit{-a.e. } \gamma.$$

Recall that

$$\mathrm{Mod}_1(\Gamma) := \inf \left\{ \int_X \rho \, \mathrm{d}\mathfrak{m} : \ \rho \geq 0, \ \int_\gamma \rho \geq 1 \ \ \forall \gamma \in \Gamma \right\}.$$

Since Mod_1-negligible set of curves parametrized on $[0, 1]$ are easily seen to be 1-negligible (it suffices to integrate with respect to any ∞-test plan $\boldsymbol{\pi}$ the inequality $\int_0^1 \rho(\gamma_t)|\dot{\gamma}_t| \geq 1$) we see that the space $W^{1,1}_S(X, \mathsf{d}, \mathfrak{m})$ of functions having 1-upper gradient is contained in $w - W^{1,1}(X, \mathsf{d}, \mathfrak{m})$, while the arguments of [75] provide the inclusion $H^{1,1}(X, \mathsf{d}, \mathfrak{m}) \subseteq W^{1,1}_S(X, \mathsf{d}, \mathfrak{m})$; moreover it is clear that $w - W^{1,1} \subseteq W^{1,1}_{BL}$ Summing up, we have

$$H^{1,1}(X, \mathsf{d}, \mathfrak{m}) \subseteq W^{1,1}_S(X, \mathsf{d}, \mathfrak{m}) \subseteq w - W^{1,1}(X, \mathsf{d}, \mathfrak{m}) \subseteq W^{1,1}_{BL}(X, \mathsf{d}, \mathfrak{m})$$

and we don't know wether equalities hold. We only know that in the last inclusion there can be discrepancy on the gradient itself; in [12], [44] is shown an example where $f \in w - BV$ but we have $|Df|_{BL} < |\nabla f|_w \mathfrak{m}$. It is worthwhile to remark (see [12] for example) that if the measure is doubling and satisfies a $(1, 1)$-Poincaré inequality then there is coincidence of spaces (in particular $H^{1,1} = W^{1,1}_{BL}$), but there is still discrepancy of the gradients, despite there exists a constant $C > 1$ such that

$$|Df|_{BL} \leq |\nabla f|_w \mathfrak{m} \leq C|Df|_{BL}.$$

However we have also equality of the gradients for example if the space is $RCD(K, \infty)$ (see [40, Remark 3.5])

A fifth space could be added to the list, considering general integrable functions f_n and replacing the asymptotic Lipschitz constants $\mathrm{lip}_a(f_n)$ with upper gradients g_n in Definition 4.6.3. However, since 1-upper gradients are characterized as strong L^1 limits of upper gradients, this space is easily seen to coincide with $W^{1,1}_S(X, \mathsf{d}, \mathfrak{m})$.

4.6.1 Comparison between $H^{1,1}$ and $H^{1,\Phi}$

It is rather easy to see that for every N-function Φ, $H^{1,\Phi}_v(X, \mathsf{d}, \mathfrak{m}) \subseteq H^{1,1}(X, \mathsf{d}, \mathfrak{m})$. In fact, given $f \in H^{1,\Phi}_v$, by definition there exists a sequence $(f_n) \subset \mathrm{Lip}_0(X, \mathsf{d})$, converging to f in $L^1(X, \mathfrak{m})$ and with $\sup_n \|\mathrm{lip}_a(f_n)\|_{(\Phi),\mathfrak{m}} = C < \infty$. But then we have that

$$\sup_n \int_X \Phi\Big(\frac{\mathrm{lip}_a(f_n)}{2C}\Big) \, \mathrm{d}\mathfrak{m} \leq 1,$$

and so $\{\mathrm{lip}_a(f_n)/2C\}$, is a family of equi-integrable functions. Thus, thanks to Dunford-Pettis theorem, there is a subsequence $\mathrm{lip}_a(f_{n_k})$ weakly converging (in $L^1(X, \mathfrak{m})$) to some $g \in L^1(X, \mathfrak{m})$, that hence will be a 1-relaxed gradient for f.

It is interesting in particular to see that if $f \in H^{1,\Phi}_v$ we have that $f \circ \gamma$ is BV for Φ-almost every curve, but since $H^{1,\Phi}_v \subseteq H^{1,1} \subseteq w - W^{1,1}$ we have that $f \circ \gamma$ is $W^{1,1}$ for 1-almost every curve.

CHAPTER 5

Reflexivity and discrete approximation of the gradient

In this chapter we focus on the Sobolev Spaces theory in the homogeneous case $\Phi(t) = t^q/q$. This case was already studied in [9], where they find the equivalence of the definitions, as we did in Chapter 3; however they use slightly different definitions (which are equivalent to ours), in particular they relax the functional F_Φ with respect to the L^q topology. One can prove the following, using the uniform convexity of the norm:

Proposition 5.0.1 *Whenever $\mathfrak{m}$ is finite on bounded sets, if $f \in L^q(X, \mathfrak{m})$ has a q-weak upper gradient then there exist Lipschitz functions f_n with bounded support satisfying*

$$\lim_{n\to\infty} \int_X |f_n - f|^q \, d\mathfrak{m} + \int_X \big|\mathrm{lip}_a(f_n) - |\nabla f|_q\big|^q \, d\mathfrak{m} = 0. \tag{5.0.1}$$

We will denote by $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$ the Banach space of functions $f \in L^q(X, \mathfrak{m})$ having a q-weak upper gradient, endowed with the norm

$$\|f\|^q_{W^{1,q}} = \|f\|^q_{L^q} + \||\nabla f|_{*,q}\|^q_{L^q}.$$

By a general property of normed spaces, in order to prove completeness, it suffices to show that any absolutely convergent series in $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$ is convergent; if f_n satisfy $\sum_n \|f_n\|^q_{W^{1,q}} < \infty$, the completeness of $L^q(X, \mathfrak{m})$ yields that $f := \sum_n f_n$ and $g := \sum_n |\nabla f_n|_{*,q}$ converge in $L^q(X, \mathfrak{m})$, and the finite subadditivity of the relaxed gradient together with the lower semicontinuity of $\mathbf{C}_q$ give $f \in W^{1,q}(X, \mathsf{d}, \mathfrak{m})$ and $\int_X |\nabla f|^q_{*,q} \, d\mathfrak{m} \le \|g\|^q_{L^q} \le (\sum_i \||\nabla f_i|_{*,q}\|_{L^q})^q$. A similar argument gives that

$$\left(\int_X |\nabla (f - \sum_{i=1}^N f_i)|^q_{*,q} \, d\mathfrak{m}\right)^{1/q} \le \sum_{i=N+1}^{\infty} \||\nabla f_i|_{*,q}\|_{L^q},$$

hence $\sum_n f_n$ converges in $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$.

5.1 Reflexivity of $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$, $1 < q < \infty$

In this section we prove that the Sobolev spaces $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$ are reflexive when $1 < q < \infty$, (X, d) is doubling (hence also separable), and $\mathfrak{m}$ is finite on bounded sets. Our strategy is to build, by a finite difference scheme, a family of functionals which provides a discrete approximation of the relaxed energy, called in this framework, the Cheeger energy. The definition of the approximate functionals relies on the existence of nice partitions of doubling metric spaces.

Lemma 5.1.1 *For every $\delta > 0$ there exist $\ell_\delta \in \mathbb{N} \cup \{\infty\}$ and pairs set-point (A_i^δ, z_i^δ), $0 \le i < \ell_\delta$, where $A_i^\delta \subset X$ are Borel sets and $z_i^\delta \in X$, satisfying:*

(i) the sets A_i^δ, $0 \le i < \ell_\delta$, are a partition of X and $\mathsf{d}(z_i^\delta, z_j^\delta) > \delta$ whenever $i \ne j$;

(ii) A_i^δ are comparable to balls centered at z_i^δ, namely

$$B\left(z_i^\delta, \frac{\delta}{3}\right) \subset A_i^\delta \subset B\left(z_i^\delta, \frac{5}{4}\delta\right).$$

Proof. Let us fix once for all a countable dense set $\{x_k\}_{k\in\mathbb{N}}$. Then, starting from $z_0^\delta = x_0$, we proceed in this way:

- for $i \ge 1$, set recursively
$$B_i = X \setminus \bigcup_{j<i} \overline{B}(z_j^\delta, \delta);$$
- if $B_i = \emptyset$ for some $i \ge 1$, then the procedure stops. Otherwise, take $z_i^\delta = x_{k_i}$ where
$$k_i = \min\{k \in \mathbb{N} \ : \ x_k \in B_i\}.$$

We claim that for every $\varepsilon > 0$ we have that

$$\bigcup_{i=0}^{\infty} B(z_i^\delta, \delta + \varepsilon) = X.$$

To show this it is sufficient to note that for every $x \in X$ we have a point x_j such that $\mathsf{d}(x_j, x) < \varepsilon$; then either $x_j = z_i^\delta$ for some i or $x_j \in \overline{B}(z_i^\delta, \delta)$ for some i. In both cases we get

$$\forall x \in X \ \exists i \in \mathbb{N} \qquad \text{such that } \mathsf{d}(z_i^\delta, x) < \delta + \varepsilon. \tag{5.1.1}$$

Now we define the sets A_i^δ similarly to a Voronoi diagram constructed from the starting point z_i^δ: for $i \in \mathbb{N}$ we set

$$B_i^\delta = \Big\{x \in X \ : \ \mathsf{d}(x, z_i^\delta) \le \mathsf{d}(x, z_j^\delta) + \varepsilon \quad \forall j\Big\}.$$

It is clear that B_i^δ are Borel sets whose union is the whole of X; we turn them into a Borel partition by setting

$$A_0^\delta = B_0^\delta, \qquad A_j^\delta := B_j^\delta \setminus \bigcup_{i<j} B_i^\delta, \quad j > 0.$$

We can also give an equivalent definition: $x \in A_k^\delta$ iff

$$k = \min I_x \qquad \text{where} \qquad I_x = \left\{ i \in \mathbb{N} \,:\, \mathsf{d}(x, z_i^\delta) \leq \mathsf{d}(x, z_j^\delta) + \varepsilon \quad \forall j \in \mathbb{N} \right\}.$$

In other words, we are minimizing the quantity $\mathsf{d}(x, z_i^\delta)$ and among those indeces i who are minimizing up to ε we take the least one i_x. By this quasi minimality and (5.1.1) we obtain $\mathsf{d}(x, z_{i_x}^\delta) \leq \inf_{i\in\mathbb{N}} \mathsf{d}(x, z_i^\delta) + \varepsilon < \delta + 2\varepsilon$. Furthermore if $\mathsf{d}(x, z_i^\delta) < \delta/2 - \varepsilon/2$ then $I_x = \{i\}$. Indeed, suppose there is another $j \in I_x$ with $j \neq i$, then $\mathsf{d}(z_j^\delta, x) \leq \mathsf{d}(z_i^\delta, x) + \varepsilon \leq \delta/2 + \varepsilon/2$ and so

$$\delta < \mathsf{d}(z_i^\delta, z_j^\delta) \leq \mathsf{d}(z_i^\delta, x) + \mathsf{d}(z_j^\delta, x) \leq \delta.$$

We just showed that

$$B\left(z_i^\delta, \frac{\delta}{2} - \frac{\varepsilon}{2}\right) \subset A_i^\delta \subset B(z_i^\delta, \delta + 2\varepsilon).$$

The dual definition gives us that A_i^δ are a partition of X, and (ii) is satisfied choosing $\varepsilon = \delta/8$. $\square$

Note that this construction is quite simpler if X is locally compact, which is always the case if (X, d) is doubling and complete. In this case we can choose $\varepsilon = 0$.

We remark that partitions with additional properties have also been studied in the literature. For example, in [27] *dyadic* partitions of a doubling metric measure space are constructed.

Definition 5.1.2 (Dyadic partition) *A dyadic partition is made by a sequence $(\ell_h) \subset \mathbb{N} \cup \{\infty\}$ and by collections of disjoint sets (called cubes) $\Delta^h = \{A_i^h\}_{1\leq i<\ell(h)}$ such that for every $h \in \mathbb{N}$ the following properties hold:*

- $\mathfrak{m}\big(X \setminus \bigcup_i A_i^h\big) = 0$;
- *for every $i \in \{1, \ldots, \ell_{h+1}\}$ there exists a unique $j \in \{1, \ldots, \ell_h\}$ such that $A_i^{h+1} \subset A_j^h$;*
- *for every $i \in \{1, \ldots \ell_h\}$ there exists $z_i^h \in X$ such that $B(z_i^h, a_0\delta^h) \subset A_i^h \subset B(z_i^h, a_1\delta^h)$ for some positive constants δ, a_0, a_1 independent of i and h.*

In [27] existence of dyadic decompositions is proved, with δ, a_1 and a_0 depending on the constant $\tilde{c}_D$ in (1.9.1). Although some more properties of the partition might give additional information on the functionals that we are going to construct, for the sake of simplicity we just work with the partition given by Lemma 5.1.1.

In order to define our discrete gradients we give more terminology. We say that A_i^δ is a *neighbor* of A_j^δ, and we denote by $A_i^\delta \sim A_j^\delta$, if their distance is less than δ. In particular $A_i^\delta \sim A_j^\delta$ implies that $\mathsf{d}(z_i^\delta, z_j^\delta) < 4\delta$: indeed, if $\tilde{z}_i^\delta \in A_i^\delta$ and $\tilde{z}_j^\delta \in A_j^\delta$ satisfy $\mathsf{d}(\tilde{z}_i^\delta, \tilde{z}_j^\delta) < \delta'$ we have

$$\mathsf{d}(z_i^\delta, z_j^\delta) \leq \mathsf{d}(z_i^\delta, \tilde{z}_i^\delta) + \mathsf{d}(\tilde{z}_i^\delta, \tilde{z}_j^\delta) + \mathsf{d}(\tilde{z}_j^\delta, z_j^\delta) \leq \frac{10}{4}\delta + \delta'$$

and letting $\delta' \downarrow \delta$ we get

$$\mathsf{d}(z_i^\delta, z_j^\delta) \leq \frac{14}{4}\delta < 4\delta.$$

This leads us to the first important property of doubling spaces:

In a c_D-doubling metric space (X, d), every A_i^δ has at most c_D^3 neighbors. (5.1.2)

Indeed, we can cover $B(z_i^\delta, 4\delta)$ with c_D^3 balls with radius $\delta/2$ but each of them, by the condition $\mathsf{d}(z_i^\delta, z_j^\delta) > \delta$, can contain only one of the z_j^δ's.

Now we fix $\delta \in (0,1)$ and we consider a partition A_i^δ of $\operatorname{supp}\mathfrak{m}$ on scale δ. For every $u \in L^q(X, \mathfrak{m})$ we define the average $u_{\delta,i}$ of u in each cell of the partition by $⨍_{A_i^\delta} u \, \mathrm{d}\mathfrak{m}$. We denote by $\mathcal{PC}_\delta(X)$, which depends on the chosen decomposition as well, the set of functions $u \in L^q(X, \mathfrak{m})$ constant on each cell of the partition at scale δ, namely

$$u(x) = u_{\delta,i} \qquad \text{for } \mathfrak{m}\text{-a.e. } x \in A_i^\delta.$$

We define a linear projection functional $\mathcal{P}_\delta : L^q(X, \mathfrak{m}) \to \mathcal{PC}_\delta(X)$ by $\mathcal{P}_\delta u(x) = u_{\delta,i}$ for every $x \in A_i^\delta$.

The proof of the following lemma is elementary.

Lemma 5.1.3 *$\mathcal{P}_\delta$ are contractions in $L^q(X, \mathfrak{m})$ and $\mathcal{P}_\delta u \to u$ in $L^q(X, \mathfrak{m})$ as $\delta \downarrow 0$ for all $u \in L^q(X, \mathfrak{m})$.*

Indeed, the contractivity of $\mathcal{P}_\delta$ is a simple consequence of Jensen's inequality and it suffices to check the convergence of $\mathcal{P}_\delta$ as $\delta \downarrow 0$ on a dense subset of $L^q(X, \mathfrak{m})$. Since $\mathfrak{m}$ is finite on bounded sets, suffices to consider bounded continuous functions with bounded support. Since bounded closed sets are compact, by the doubling property, it follows that any such function u is uniformly continuous, so that $\mathcal{P}_\delta u \to u$ pointwise as $\delta \downarrow 0$. Then, we can use the dominated convergence theorem to conclude.

We now define an approximate gradient as follows: it is constant on the cell A_i^δ for every δ, $i \in \mathbb{N}$ and it takes the value

$$|\mathcal{D}_\delta u|^q(x) := \frac{1}{\delta^q} \sum_{A_j^\delta \sim A_i^\delta} |u_{\delta,i} - u_{\delta,j}|^q \qquad \forall x \in A_i^\delta.$$

We can accordingly define the functional $\mathcal{F}_{\delta,q} : L^q(X, \mathfrak{m}) \to [0, \infty]$ by

$$\mathcal{F}_{\delta,q}(u) := \int_X |\mathcal{D}_\delta u|^q(x) \, \mathrm{d}\mathfrak{m}(x). \tag{5.1.3}$$

Now, using the weak gradients, we define a functional $\mathrm{Ch} : L^q(X, \mathfrak{m}) \to [0, \infty]$ that we call *Cheeger* energy, formally similar to $\mathscr{G}^\Phi$. Namely, we set

$$\mathrm{Ch}_q(u) := \begin{cases} \int_X |\nabla u|_{w,q}^q \, \mathrm{d}\mathfrak{m} & \text{if } u \text{ has a } q\text{-weak upper gradient} \\ +\infty & \text{otherwise.} \end{cases}$$

At this level of generality, we cannot expect that the functionals $\mathcal{F}_{\delta,q}$ Γ-converge as $\delta \downarrow 0$. However, since $L^q(X, \mathfrak{m})$ is a complete and separable metric space, from the compactness property of Γ-convergence stated in Proposition 1.8.2 we obtain that the functionals $\mathcal{F}_{\delta,q}$ have Γ-limit points as $\delta \downarrow 0$.

Theorem 5.1.4 *Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric measure space with $(\operatorname{supp}\mathfrak{m}, \mathsf{d})$ complete and doubling, $\mathfrak{m}$ finite on bounded sets. Let $\mathcal{F}_q$ be a Γ-limit point of $\mathcal{F}_{\delta,q}$ as $\delta \downarrow 0$, namely*

$$\mathcal{F}_q := \Gamma - \lim_{k \to \infty} \mathcal{F}_{\delta_k, q},$$

for some infinitesimal sequence (δ_k), where the Γ-limit is computed with respect to the $L^q(X, \mathfrak{m})$ distance. Then:

(a) $\mathcal{F}_q$ *is equivalent to the Cheeger energy* Ch_q*, namely there exists* $\eta=\eta(q,c_D)$ *such that*

$$\frac{1}{\eta}\,\mathrm{Ch}_q(u)\le\mathcal{F}_q(u)\le\eta\,\mathrm{Ch}_q(u)\qquad\forall u\in L^q(X,\mathfrak{m}).\tag{5.1.4}$$

(b) The norm on $W^{1,q}(X,\mathsf{d},\mathfrak{m})$ *defined by*

$$\big(\|u\|_q^q+\mathcal{F}_q(u)\big)^{1/q}\qquad\forall u\in W^{1,q}(X,\mathsf{d},\mathfrak{m})\tag{5.1.5}$$

is uniformly convex. Moreover, the seminorm $\mathcal{F}_2^{1/2}$ *is Hilbertian, namely*

$$\mathcal{F}_2(u+v)+\mathcal{F}_2(u-v)=2\big(\mathcal{F}_2(u)+\mathcal{F}_2(v)\big)\qquad\forall u,\,v\in W^{1,2}(X,\mathsf{d},\mathfrak{m}).\tag{5.1.6}$$

Corollary 5.1.5 (Reflexivity of $W^{1,q}(X,\mathsf{d},\mathfrak{m})$) *Let* $(X,\mathsf{d},\mathfrak{m})$ *be a metric measure space with* $(\operatorname{supp}\mathfrak{m},\mathsf{d})$ *doubling and* $\mathfrak{m}$ *finite on bounded sets. The Sobolev space* $W^{1,q}(X,\mathsf{d},\mathfrak{m})$ *of functions* $u\in L^q(X,\mathfrak{m})$ *with a q-relaxed slope, endowed with the usual norm*

$$\big(\|u\|_q^q+\mathrm{Ch}_q(u)\big)^{1/q}\qquad\forall u\in W^{1,q}(X,\mathsf{d},\mathfrak{m}),\tag{5.1.7}$$

is reflexive.

Proof. Since the Banach norms (5.1.5) and (5.1.7) on $W^{1,q}(X,\mathsf{d},\mathfrak{m})$ are equivalent thanks to (5.1.4) and reflexivity is invariant, we can work with the first norm. The Banach space $W^{1,q}(X,\mathsf{d},\mathfrak{m})$ endowed with the first norm is reflexive by uniform convexity and Milman-Pettis theorem. □

We can also prove, by standard functional-analytic arguments, that reflexivity implies separability.

Proposition 5.1.6 (Separability of $W^{1,q}(X,\mathsf{d},\mathfrak{m})$) *If* $W^{1,q}(X,\mathsf{d},\mathfrak{m})$ *is reflexive and* $\mathfrak{m}$ *is finite on bounded sets, then it is separable and bounded Lipschitz functions with bounded support are dense.*

Proof. The density of Lipschitz functions with bounded support follows via Mazur lemma from the density of this convex set in the weak topology, ensured by Proposition 5.0.1 and reflexivity. In order to prove separability, it suffices to consider for any M a countable and $L^q(X,\mathfrak{m})$-dense subset $\mathcal{D}_M$ of

$$\mathcal{L}_M:=\left\{f\in\mathrm{Lip}(X)\cap L^q(X,\mathfrak{m}):\ \int_X|\nabla f|_{w,q}^q\,\mathrm{d}\mathfrak{m}\le M\right\},$$

stable under convex combinations with rational coefficients. The weak closure of $\mathcal{D}_M$ obviously contains $\mathcal{L}_M$, by reflexivity (because if $f_n\in\mathcal{D}_M$ converge to $f\in\mathcal{L}_M$ in $L^q(X,\mathfrak{m})$, then $f_n\to f$ weakly in $W^{1,q}(X,\mathsf{d},\mathfrak{m})$); being this closure convex, it coincides with the strong closure of $\mathcal{D}_M$. This way we obtain that the closure in the strong topology of $\cup_M\mathcal{D}_M$ contains all Lipschitz functions with bounded support. □

The strategy of the proof of statement (a) in Theorem 5.1.4 consists in proving the estimate from above of $\mathcal{F}_q$ with relaxed gradients and the estimate from below with weak gradients. Then, the equivalence between weak and relaxed gradients provides the result. In the estimate from below it will be useful the discrete version of the q-weak upper gradient property:

Definition 5.1.7 (q-weak upper gradient up to scale ε) *Let $f : X \to \mathbb{R}$. We say that a Borel function $g : X \to [0,\infty)$ is a q-weak upper gradient of f up to scale $\varepsilon \geq 0$ if for q-a.e. curve $\gamma \in AC^p([0,1];X)$ such that*

$$\varepsilon < \int_0^1 |\dot\gamma_t| \, \mathrm{d}t$$

it holds

$$\Big| \int_{\partial\gamma} f \Big| \leq \int_\gamma g < \infty. \tag{5.1.8}$$

Next we consider the stability of these discretized q-weak upper gradients (analogous to the stability result given in [75, Lemma 4.11]).

Theorem 5.1.8 (Stability w.r.t. $\mathfrak{m}$-a.e. convergence) *Assume that f_n are $\mathfrak{m}$-measurable, $\varepsilon_n \geq 0$ and that $g_n \in L^q(X,\mathfrak{m})$ are q-weak upper gradients of f_n up to scale ε_n. Assume furthermore that $f_n(x) \to f(x) \in \mathbb{R}$ for $\mathfrak{m}$-a.e. $x \in X$, $\varepsilon_n \to \varepsilon$ and that (g_n) weakly converges to g in $L^q(X,\mathfrak{m})$. Then g is a q-weak upper gradient of f up to scale ε.*

Proof. Fix a p-test plan $\boldsymbol{\pi}$. We have to show that (5.1.8) holds for $\boldsymbol{\pi}$-a.e. γ with $\int_0^1 |\dot\gamma_t| \, \mathrm{d}t > \varepsilon$. Possibly restricting $\boldsymbol{\pi}$ to a smaller set of curves, we can assume with no loss of generality that

$$\int_0^1 |\dot\gamma_t| \, \mathrm{d}t > \varepsilon' \qquad \text{for } \boldsymbol{\pi}\text{-a.e. } \gamma$$

for some $\varepsilon' > \varepsilon$. We consider in the sequel integers h sufficiently large, such that $\varepsilon_h \leq \varepsilon'$.

By Mazur's lemma we can find convex combinations

$$h_n := \sum_{i=N_h+1}^{N_{h+1}} \alpha_i g_i \qquad \text{with } \alpha_i \geq 0, \ \sum_{i=N_h+1}^{N_{h+1}} \alpha_i = 1, \ N_h \to \infty$$

converging strongly to g in $L^q(X,\mathfrak{m})$. Denoting by $\tilde f_n$ the corresponding convex combinations of f_n, h_n are q-weak upper gradients of $\tilde f_n$ and still $\tilde f_n \to f$ $\mathfrak{m}$-a.e. in X.

Since for every nonnegative Borel function $\varphi : X \to [0,\infty]$ it holds (with $C = C(\boldsymbol{\pi})$)

$$\begin{aligned}
\int \Big(\int_\gamma \varphi \Big) \mathrm{d}\boldsymbol{\pi} &= \int \Big(\int_0^1 \varphi(\gamma_t)|\dot\gamma_t| \, \mathrm{d}t \Big) \mathrm{d}\boldsymbol{\pi} \leq \int \Big(\int_0^1 \varphi^q(\gamma_t) \, \mathrm{d}t \Big)^{1/q} \Big(\int_0^1 |\dot\gamma_t|^p \, \mathrm{d}t \Big)^{1/p} \mathrm{d}\boldsymbol{\pi} \\
&\leq \Big(\int_0^1 \int \varphi^q \, \mathrm{d}(\mathrm{e}_t)_\sharp \boldsymbol{\pi} \, \mathrm{d}t \Big)^{1/q} \Big(\iint_0^1 |\dot\gamma_t|^p \, \mathrm{d}t \, \mathrm{d}\boldsymbol{\pi} \Big)^{1/p} \\
&\leq \Big(C \int \varphi^q \, \mathrm{d}\mathfrak{m} \Big)^{1/q} \Big(\iint_0^1 |\dot\gamma_t|^p \, \mathrm{d}t \, \mathrm{d}\boldsymbol{\pi} \Big)^{1/p},
\end{aligned} \tag{5.1.9}$$

we obtain

$$\int \int_\gamma |h_n - g| \, \mathrm{d}\boldsymbol{\pi} \leq C^{1/q} \Big(\iint_0^1 |\dot\gamma_t|^p \, \mathrm{d}t \, \mathrm{d}\boldsymbol{\pi} \Big)^{1/p} \|h_n - g\|_q \to 0.$$

Hence we can find a subsequence $n(k)$ such that

$$\lim_{k\to\infty} \int_\gamma |h_{n(k)} - g| \to 0 \qquad \text{for } \boldsymbol{\pi}\text{-a.e. } \gamma.$$

Since $\tilde f_n$ converge $\mathfrak{m}$-a.e. to f and the marginals of $\boldsymbol{\pi}$ are absolutely continuous w.r.t. $\mathfrak{m}$ we have also that for $\boldsymbol{\pi}$-a.e. γ it holds $\tilde f_n(\gamma_0) \to f(\gamma_0)$ and $\tilde f_n(\gamma_1) \to f(\gamma_1)$.

If we fix a curve γ satisfying these convergence properties, we can pass to the limit as $k \to \infty$ in the inequalities $|\int_{\partial\gamma} \tilde f_{n(k)}| \leq \int_\gamma h_{n(k)}$ to get $|\int_{\partial\gamma} f| \leq \int_\gamma g$. $\square$

In the following lemma we prove that for every $u \in L^q(X,\mathfrak{m})$ we have that $4|\mathcal{D}_\delta u|$ is a q-weak upper gradient for $\mathcal{P}_\delta u$ up to scale $\delta/2$.

Lemma 5.1.9 *Let $\gamma \in AC^p([0,1];X)$. Then we have that*

$$|\mathcal{P}_\delta u(\gamma_b) - \mathcal{P}_\delta u(\gamma_a)| \le 4\int_a^b |\mathcal{D}_\delta u|(\gamma_t)|\dot\gamma_t|\,\mathrm{d}t \qquad \text{for all } a<b \text{ s.t. } \int_a^b |\dot\gamma_t|\,\mathrm{d}t > \delta/2. \tag{5.1.10}$$

In particular $4|\mathcal{D}_\delta u|$ is a q-weak upper gradient of $\mathcal{P}_\delta u$ up to scale $\delta/2$.

Proof. It is enough to prove the inequality under the more restrictive assumption that

$$\frac{\delta}{2} \le \int_a^b |\dot\gamma_t|\,\mathrm{d}t \le \delta, \tag{5.1.11}$$

because then we can slice every interval (a,b) that is longer than $\delta/2$ into subintervals that satisfy (5.1.11), and we get (5.1.8) by adding the inequalities for subintervals and using triangular inequality.

Now we prove (5.1.8) for every a, $b \in [0,1]$ such that (5.1.11) holds. Take any time $t \in [a,b]$; by assumption, it is clear that $\mathsf{d}(\gamma_t,\gamma_a) \le \delta$ and $\mathsf{d}(\gamma_t,\gamma_b) \le \delta$, so that the cells relative to γ_a and γ_b are both neighbors of the one relative to γ_t. By definition then we have:

$$|\mathcal{D}_\delta u|^q(\gamma_t) \ge \frac{1}{\delta^q}\left(|\mathcal{P}_\delta u(\gamma_b) - \mathcal{P}_\delta u(\gamma_t)|^q + |\mathcal{P}_\delta u(\gamma_t) - \mathcal{P}_\delta u(\gamma_a)|^q\right) \ge \frac{1}{2^{q-1}\delta^q}|\mathcal{P}_\delta u(\gamma_b) - \mathcal{P}_\delta u(\gamma_a)|^q.$$

Taking the q-th root and integrating in t we get

$$\int_a^b |\mathcal{D}_\delta u|(\gamma_t)|\dot\gamma_t|\,\mathrm{d}t \ge \frac{|\mathcal{P}_\delta u(\gamma_b) - \mathcal{P}_\delta u(\gamma_a)|}{2^{1-1/q}\delta}\int_a^b |\dot\gamma_t|\,\mathrm{d}t \ge \frac{1}{2}|\mathcal{P}_\delta u(\gamma_b) - \mathcal{P}_\delta u(\gamma_a)|,$$

which proves (5.1.10). □

We can now prove Theorem 5.1.4.

Proof of the first inequality in (5.1.4). We prove that there exists a constant $\eta_1 = \eta_1(c_D)$ such that

$$\mathcal{F}_q(u) \le \eta_1 \int_X |\nabla f|_{*,q}^q\,\mathrm{d}\mathfrak{m} \qquad \forall u \in L^q(X,\mathfrak{m}). \tag{5.1.12}$$

Let $u : X \to \mathbb{R}$ be a Lipschitz function with bounded support. We prove that

$$|\mathcal{D}_\delta u|^q(x) \le 6^q c_D^3(\mathrm{Lip}(u,B(x,6\delta)))^q. \tag{5.1.13}$$

Indeed, let us consider i, $j \in [1,\ell_\delta)\cap\mathbb{N}$ such that A_i^δ and A_j^δ are neighbors. For every $x \in A_i^\delta$, $y \in A_j^\delta$ we have that $\mathsf{d}(x,y) \le \mathrm{diam}(A_i^\delta) + \mathrm{diam}(A_i^\delta) + \mathsf{d}(A_i^\delta, A_j^\delta) \le (10/4+10/4+1)\delta = 6\delta$ and that $y \in B(z_i^\delta, 19\delta/4) \subset B(z_i^\delta, 5\delta)$. Hence

$$\frac{|u_{\delta,i} - u_{\delta,j}|}{\delta} \le \frac{1}{\delta\mathfrak{m}(A_i^\delta)\mathfrak{m}(A_j^\delta)}\int_{A_i^\delta\times A_j^\delta} |u(x)-u(y)|\,\mathrm{d}\mathfrak{m}(x)\,\mathrm{d}\mathfrak{m}(y) \le 6\,\mathrm{Lip}(u,B(z_i^\delta,5\delta)).$$

Thanks to the fact that the number of neighbors of A_i^h does not exceed c_D^3 (see (5.1.2)) we obtain

$$|\mathcal{D}_\delta u|^q(x) \le 6^q c_D^3(\mathrm{Lip}(u,B(x,6\delta)))^q \qquad \forall x \in \operatorname{supp}\mathfrak{m},$$

which proves (5.1.13).

Integrating on X we obtain that

$$\mathcal{F}_{\delta,q}(u) \leq 6^q c_D^3 \int_X (\mathrm{Lip}(u, B(x, 6\delta)))^q \,\mathrm{d}\mathfrak{m}(x).$$

Choosing $\delta = \delta_k$, letting $k \to \infty$ and applying the dominated convergence theorem on the right-hand side as well as the definition of asymptotic Lipschitz constant (3) we get

$$\mathcal{F}_q(u) \leq \liminf_{k\to\infty} \mathcal{F}_{\delta_k,q}(u) \leq 6^q c_D^3 \int_X \mathrm{lip}_a^q(u, x) \,\mathrm{d}\mathfrak{m}(x).$$

By approximation, Proposition 5.0.1 yields (5.1.12) with $\eta_1 = 6^q c_D^3$.

Proof of the second inequality in (5.1.4). We consider a sequence (u_k) which converges to u in $L^q(X, \mathfrak{m})$ with $\liminf_k \mathcal{F}_{\delta_k,q}(u_k)$ finite. We prove that u has a q-weak upper gradient and that

$$\frac{1}{4^q} \int_X |\nabla u|_{w,q}^q \,\mathrm{d}\mathfrak{m} \leq \liminf_k \mathcal{F}_{\delta_k,q}(u_k). \tag{5.1.14}$$

Then, (5.1.4) will follow easily from (5.1.12), (5.1.14), Definition 1.8.1b and the coincidence of weak and relaxed gradients.

Without loss of generality we assume that the right-hand side is finite and, up to a subsequence not relabeled, we assume that the lim inf is a limit. Hence, the sequence $f_k := |\mathcal{D}_{\delta_k} u_k|$ is bounded in $L^q(X, \mathfrak{m})$ and, by weak compactness, there exist $g \in L^q(X, \mathfrak{m})$ and a subsequence $k(h)$ such that $f_{k(h)} \rightharpoonup g$ weakly in $L^q(X, \mathfrak{m})$. By the lower semicontinuity of the q-norm with respect to the weak convergence, we have that

$$\int_X g^q \,\mathrm{d}\mathfrak{m} \leq \liminf_{h\to\infty} \int_X f_{k(h)}^q \,\mathrm{d}\mathfrak{m} = \lim_{k\to\infty} \mathcal{F}_{\delta_k,q}(u_k). \tag{5.1.15}$$

We can now apply Theorem 5.1.8 to the functions $\bar{u}_h = \mathcal{P}_{\delta_{k(h)}}(u_{k(h)})$, which converge to u in $L^q(X, \mathfrak{m})$ thanks to Lemma 5.1.3, and to the functions $g_h = 4 f_{k(h)}$ which are q-weak upper gradients of $\bar{u}_h$ up to scale $\delta_{k(h)}/2$, thanks to Lemma 5.1.9. We obtain that $4g$ is a weak upper gradient of u, hence $g \geq |\nabla u|_{w,q}/4$ $\mathfrak{m}$-a.e. in X. Therefore (5.1.15) gives

$$\frac{1}{4^q} \int_X |\nabla u|_{w,q}^q \,\mathrm{d}\mathfrak{m} \leq \int_X g^q \,\mathrm{d}\mathfrak{m} \leq \lim_{k\to\infty} \mathcal{F}_{\delta_k,q}(u_k).$$

Proof of statement (b). Let $\mathcal{N}_{q,\delta} : L^q(X, \mathfrak{m}) \to [0, \infty]$ be the positively 1-homogeneous function

$$\mathcal{N}_{q,\delta}(u) = \big(\|\mathcal{P}_\delta u\|_q^q + \mathcal{F}_\delta(u)\big)^{1/q} \qquad \forall u \in L^q(X, \mathfrak{m}).$$

For $q \geq 2$ we prove that $\mathcal{N}_{q,\delta}$ satisfies the first Clarkson inequality [53]

$$\mathcal{N}_{q,\delta}^q\left(\frac{u+v}{2}\right) + \mathcal{N}_{q,\delta}^q\left(\frac{u-v}{2}\right) \leq \frac{1}{2}\big(\mathcal{N}_{q,\delta}^q(u) + \mathcal{N}_{q,\delta}^q(v)\big) \qquad \forall u,\, v \in L^q(X, \mathfrak{m}). \tag{5.1.16}$$

Indeed, let $X_\delta \subset \mathbb{N} \cup (\mathbb{N} \times \mathbb{N})$ be the (possibly infinite) set

$$X_\delta = \big([1, \ell_\delta) \cap \mathbb{N}\big) \cup \Big\{(i,j) \in \big([1, \ell_\delta) \cap \mathbb{N}\big)^2 : \; A_i^\delta \sim A_j^\delta\Big\}$$

and let $\mathfrak{m}_\delta$ be the counting measure on X_δ. We consider the function $\Phi_{q,\delta} : L^q(X,\mathfrak{m}) \to L^q(X_\delta,\mathfrak{m}_\delta)$ defined by

$$\begin{cases} \Phi_{q,\delta}[u](i) = (\mathfrak{m}(A_i^\delta))^{1/q} u_{\delta,i} & \forall i \in [1,\ell_\delta)\cap\mathbb{N} \\ \Phi_{q,\delta}[u]((i,j)) = (\mathfrak{m}(A_i^\delta))^{1/q} \dfrac{u_{\delta,i}-u_{\delta,j}}{\delta} & \forall (i,j) \in \big([1,\ell_\delta)\cap\mathbb{N}\big)^2 \text{ s.t. } A_i^\delta \sim A_j^\delta. \end{cases}$$

It can be easily seen that $\Phi_{q,\delta}$ is linear and that

$$\|\Phi_{q,\delta}(u)\|_{L^q(X_\delta,\mathfrak{m}_\delta)} = \mathcal{N}_{q,\delta}(u) \qquad \forall u \in L^q(X,\mathfrak{m}). \tag{5.1.17}$$

Writing the first Clarkson inequality in the space $L^q(X_h,\mathfrak{m}_h)$ and using the linearity of $\Phi_{q,\delta}$ we immediately obtain (5.1.16). Let $\omega : (0,1)\to(0,\infty)$ be the increasing and continuous modulus of continuity $\omega(r) = 1-(1-r^q/2^q)^{1/q}$. >From (5.1.16) it follows that for all u, $v \in L^q(X,\mathfrak{m})$ with $\mathcal{N}_{q,\delta}(u)=\mathcal{N}_{q,\delta}(v)=1$ it holds

$$\mathcal{N}_{q,\delta}\left(\frac{u+v}{2}\right) \le 1-\omega(\mathcal{N}_{q,\delta}(u-v)).$$

Hence $\mathcal{N}_{q,\delta}$ are uniformly convex with the same modulus of continuity ω. Thanks to Lemma 1.8.3 we conclude that also the Γ-limit of these norms, namely (5.1.5), is uniformly convex with the same modulus of continuity.

If $q<2$ the proof can be repeated substituting the first Clarkson inequality (5.1.16) with the second one

$$\left[\mathcal{N}_{q,\delta}\left(\frac{u+v}{2}\right)\right]^p + \left[\mathcal{N}_{q,\delta}\left(\frac{u-v}{2}\right)\right]^p \le \left[\frac{1}{2}\big(\mathcal{N}_{q,\delta}(u)\big)^q + \frac{1}{2}\big(\mathcal{N}_{q,\delta}(v)\big)^q\right]^{1/(q-1)}$$

where $u,v\in L^q(X,\mathfrak{m})$ and $p=q/(q-1)$, see [53]. In this case the modulus ω is $1-(1-(r/2)^p)^{1/p}$.

Finally, let us consider the case $q=2$. From the Clarkson inequality we get

$$\mathcal{F}_2\left(\frac{u+v}{2}\right)+\mathcal{F}_2\left(\frac{u-v}{2}\right) \le 2\big(\mathcal{F}_2(u)+\mathcal{F}_2(v)\big). \tag{5.1.18}$$

If we apply the same inequality to $u=(u'+v')/2$ and $v=(u'-v')/2$ we obtain a converse inequality and, since u' and v' are arbitrary, the equality.

We conclude this section providing a counterexample to reflexivity. We denote by ℓ_1 the Banach space of summable sequences $(x_n)_{n\ge 0}$ and by ℓ_∞ the dual space of bounded sequences, with duality $\langle\cdot,\cdot\rangle$ and norm $\|v\|_\infty$. We shall use the factorization $\ell_1 = Y_i + \mathbb{R}e_i$, where e_i, $0\le i<\infty$, are the elements of the canonical basis of ℓ_1. Accordingly, for fixed i we write $x = x_i' + x_i e_i$ and, for $f:\ell_1\to\mathbb{R}$ and $y\in Y_i$, we set

$$f_y(t) := f(y+te_i) \qquad t\in\mathbb{R}.$$

Proposition 5.1.10 *There exist a compact subset X of ℓ_1 and $\mathfrak{m}\in\mathcal{P}(X)$ such that, if d is the distance induced by the inclusion in ℓ_1, the space $W^{1,q}(X,\mathsf{d},\mathfrak{m})$ is not reflexive for all $q\in(1,\infty)$.*

Proof. For $i \geq 0$, we denote by $\mathfrak{m}_i$ the normalized Lebesgue measure in $X_i := [0, 2^{-i}]$ and define X to be the product of the intervals X_i and $\mathfrak{m}$ to be the product measure. Since X is a compact subset of ℓ_1, we shall also view $\mathfrak{m}$ as a probability measure in ℓ_1 concentrated on X.

Setting $f^v(x) := \langle v, x\rangle$, we shall prove that the map $v \mapsto f^v$ provides a linear isometry between ℓ_∞, endowed with the norm

$$|v|_\infty := \left(\int_X |\langle v, x\rangle|^q \, \mathrm{d}\mathfrak{m}(x) + \|v\|_\infty^q\right)^{1/q} \tag{5.1.19}$$

and $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$. Since the norm (5.1.19) is equivalent to the ℓ_∞ norm, it follows that $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$ contains a non-reflexive closed subspace and therefore it is itself non-reflexive.

Since the Lipschitz constant of f^v is $\|v\|_\infty$, it is clear that $\||\nabla f^v|_{w,q}\|_{L^q} \leq \|v\|_\infty$. To prove equality, suffices to show that $\int_X |\nabla f^v|_{w,q}^q \, \mathrm{d}\mathfrak{m} \geq \|v\|_\infty^q$. Therefore we fix an integer $i \geq 0$ and we prove that $\int_X |\nabla f^v|_{w,q}^q \, \mathrm{d}\mathfrak{m} \geq |v_i|^q$.

Fix a sequence (f^n) of Lipschitz functions with bounded support with f^n and $\mathrm{lip}_a(f^n)$ strongly convergent in $L^q(X, \mathfrak{m})$ to f^v and $|\nabla f^v|_{w,q}$ respectively. Possibly refining the sequence, we can assume that

$$\sum_n \|f^n - f^v\|_q^q < \infty. \tag{5.1.20}$$

If we show that

$$\liminf_{n\to\infty} \int_X \mathrm{lip}_a^q(f^n, x) \, \mathrm{d}\mathfrak{m}(x) \geq |v_i|^q \tag{5.1.21}$$

we are done. Denoting $\mathfrak{m} = \tilde{\mathfrak{m}}_i \otimes \mathfrak{m}_i$ the factorization of $\mathfrak{m}$ (with $\tilde{\mathfrak{m}}_i \in \mathcal{P}(Y_i)$), we can use the obvious pointwise inequalities

$$\mathrm{lip}_a(g, y + te_i) \geq \mathrm{lip}_a(g_y, t) \geq |\nabla g_y|(t)$$

and Fatou's lemma, to reduce the proof of (5.1.21) to the one-dimensional statement

$$\liminf_{n\to\infty} \int_{X_i} |\nabla f_y^n|^q(t) \, \mathrm{d}\mathfrak{m}_i(t) \geq |v_i|^q \qquad \text{for } \tilde{\mathfrak{m}}_i\text{-a.e. } y \in Y_i. \tag{5.1.22}$$

Since (5.1.20) yields

$$\int_{Y_i} \sum_n \|f_y^n - f_y^v\|_{L^q(X_i, \mathfrak{m}_i)}^q \, \mathrm{d}\tilde{\mathfrak{m}}_i(y) = \sum_n \|f^n - f^v\|_{L^q(X, \mathfrak{m})}^q < \infty,$$

we have that $f_y^n \to f_y^v$ in $L^q(X_i, \mathfrak{m}_i) = L^q(X_i, 2^i \mathscr{L}^1)$ for $\tilde{\mathfrak{m}}_i$-a.e. $y \in Y_i$. We have also $|\nabla f_y^v|(t) = |v_i|$ for any $t \in X_i$, therefore (5.1.22) is a consequence of the well-known lower semicontinuity in $L^q(X_i, \mathscr{L}^1)$ of $g \mapsto \int_{X_i} |g'(t)|^q \, d\mathscr{L}^1(t)$ for Lipschitz functions defined on the real line (notice also that in this context we can replace the slope with the modulus of derivative, wherever it exists). □

5.2 Lower semicontinuity of the slope of Lipschitz functions

Let us recall, first, the formulation of the Poincaré inequality in metric measure spaces.

Definition 5.2.1 *The metric measure space* $(X, \mathsf{d}, \mathfrak{m})$ *supports a weak* $(1, q)$*-Poincaré inequality if there exist constants* τ, $\Lambda > 0$ *such that for every* $u \in W^{1,q}(X, \mathsf{d}, \mathfrak{m})$ *and for every* $x \in \operatorname{supp} \mathfrak{m}$, $r > 0$ *the following holds:*

$$\fint_{B(x,r)} \left| u - \fint_{B(x,r)} u \right| \mathrm{d}\mathfrak{m} \le \tau\, r \left(\fint_{B(x,\Lambda r)} |\nabla u|_{w,q}^q \,\mathrm{d}\mathfrak{m} \right)^{1/q}. \tag{5.2.1}$$

Many different and equivalent formulations of (5.2.1) are possible: for instance we may replace in the right hand side $|\nabla u|_{w,q}^q$ with $|\nabla u|^q$, requiring the validity of the inequality for Lipschitz functions only. The equivalence of the two formulations has been first proved in [50], but one can also use the equivalence of weak and relaxed gradients to establish it. Other formulations involve the median, or replace the left hand side by

$$\inf_{m \in \mathbb{R}} \fint_{B(x,r)} |u - m| \,\mathrm{d}\mathfrak{m}.$$

The following lemma contains the fundamental estimate to prove our result.

Lemma 5.2.2 *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be a doubling metric measure space which supports a weak* $(1, q)$*-Poincaré inequality with constants* τ, Λ. *Let* $u \in W^{1,q}(X, \mathsf{d}, \mathfrak{m})$ *and let* $g = |\nabla u|_{w,q}^q$. *There exists a constant* $C > 0$ *depending only on the doubling constant* $\tilde{c}_D$ *and* τ *such that*

$$|u(x) - u(y)| \le C\mathsf{d}(x, y)\big(\big(M_q^{2\Lambda \mathsf{d}(x,y)} g(x)\big)^{1/q} + \big(M_q^{2\Lambda \mathsf{d}(x,y)} g(y)\big)^{1/q}\big), \tag{5.2.2}$$

for every Lebesgue points x, $y \in X$ *of (a representative of)* u.

Proof. The main estimate in the proof is the following. Denoting by $u_{z,r}$ the mean value of u on $B(z, r)$, for every $s > 0$, x, $y \in X$ such that $B(x, s) \subset B(y, 2s)$ we have that

$$|u_{x,s} - u_{y,2s}| \le C_0(\tilde{c}_D, \tau)s\big(M_q^{2\Lambda s} g(y)\big)^{1/q}. \tag{5.2.3}$$

Since $\mathfrak{m}$ is doubling and the space supports $(1, q)$-Poincaré inequality, from (1.9.2) we have that

$$\begin{aligned} |u_{x,s} - u_{y,2s}| &\le \fint_{B(x,s)} |u - u_{y,2s}| \,\mathrm{d}\mathfrak{m} \le \beta 2^\alpha \fint_{B(y,2s)} |u - u_{y,2s}| \,\mathrm{d}\mathfrak{m} \\ &\le 2^{1+\alpha}\beta\tau s \left(\fint_{B(y,2\Lambda s)} g^q \,\mathrm{d}\mathfrak{m} \right)^{1/q} \end{aligned}$$

and we obtain (5.2.3) with $C_0 = 2^{1+\alpha}\beta\tau$.

For every $r > 0$ let $s_n = 2^{-n} r$ for every $n \ge 1$. If x is a Lebesgue point for u then $u_{x,s_n} \to u(x)$ as $n \to \infty$. Hence, applying (5.2.3) to $x = y$ and $s_n = 2^{-n} r$, summing on $n \ge 1$ and remarking that $M_q^{2\Lambda s_n} g \le M_q^{\Lambda r} g$, we get

$$|u_{x,r} - u(x)| \le \sum_{n=0}^{\infty} |u_{x,s_n} - u_{x,2s_n}| \le \sum_{n=0}^{\infty} C_0 s_n \big(M_q^{\Lambda r} g(x)\big)^{1/q} = C_0 r \big(M_q^{\Lambda r} g(x)\big)^{1/q}. \tag{5.2.4}$$

For every $r > 0$, x, y Lebesgue points of u such that $B(x,r) \subset B(y,2r)$, we can use the triangle inequality, (5.2.3) and (5.2.4) to get

$$\begin{aligned}|u(x)-u(y)| &\le |u(x)-u_{x,r}| + |u_{x,r}-u_{y,2r}| + |u_{y,2r}-u(y)| \\ &\le C_0 r\big(M_q^{\Lambda r} g(x)\big)^{1/q} + C_0 r\big(M_q^{2\Lambda r} g(y)\big)^{1/q} + C_0 r\big(M_q^{\Lambda r} g(y)\big)^{1/q}.\end{aligned}$$

Taking $r = \mathsf{d}(x,y)$ (which obviously implies $B(x,r) \subset B(y,2r)$) and since $M_q^{\varepsilon} f(x)$ is nondecreasing in ε we obtain (5.2.2) with $C = 2C_0$. □

Proposition 5.2.3 *Let $(X, \mathsf{d}, \mathfrak{m})$ be a doubling metric measure space, supporting a weak $(1,q)$-Poincaré inequality with constants τ, Λ and with* $\operatorname{supp}\mathfrak{m} = X$ *There exists a constant $C > 0$ depending only on the doubling constant $\tilde{c}_D$ and τ such that*

$$|\nabla u| \le C\,|\nabla u|_{w,q} \quad \mathfrak{m}\text{-a.e. in } X \tag{5.2.5}$$

for any Lipschitz function u with bounded support.

Proof. We set $g = |\nabla u|_{w,q}^q$; we note that g is bounded and with bounded support, thus $M_q^{\varepsilon} g$ converges to g in $L^q(X,\mathfrak{m})$ as $\varepsilon \to 0$. Let us fix $\lambda > 0$ and a point x where (1.9.4) is satisfied by $M_q^{\lambda} g$. Let $y_n \to x$ be such that

$$|\nabla u|(x) = \lim_{n\to\infty} \frac{|u(y_n)-u(x)|}{\mathsf{d}(y_n,x)} \tag{5.2.6}$$

and set $r_n = \mathsf{d}(x,y_n)$, $B_n = B(y_n,\lambda r_n) \subset B(x,2r_n)$. Since (5.2.2) of Lemma 5.2.2 holds for $\mathfrak{m}$-a.e. $y \in B_n$, from the monotonicity of $M_q^{\varepsilon} g$ we get

$$\begin{aligned}|u(x)-u(y_n)| &\le \fint_{B_n} |u(x)-u(y)|\,\mathrm{d}\mathfrak{m}(y) + \lambda r_n \operatorname{Lip}(u,B_n) \\ &\le C r_n\left(\big(M_q^{4\Lambda r_n} g(x)\big)^{1/q} + \fint_{B_n} \big(M_q^{4\Lambda r_n} g(y)\big)^{1/q}\,\mathrm{d}\mathfrak{m}(y)\right) + \lambda r_n L,\end{aligned}$$

where L is the Lipschitz constant of u. For n large enough $B_n \subset B(x,1)$ and $4\Lambda r_n \le \lambda$. Using monotonicity once more we get

$$|u(x)-u(y_n)| \le C r_n\left(M_q^{\lambda} g(x) + \fint_{B_n} \big(M_q^{\lambda} g\big)^{1/q}\,\mathrm{d}\mathfrak{m}\right) + \lambda r_n L \tag{5.2.7}$$

for n large enough. Since $B(y_n,r_n) = B_n \subset B(x,2r_n)$ and since x is a 1-Lebesgue point for $M_q^{\lambda} g$, we apply (1.9.5) of Lemma 1.9.3 to the sets B_n to get

$$\lim_{n\to\infty} \fint_{B_n} M_q^{\lambda} g\,\mathrm{d}\mathfrak{m} = M_q^{\lambda} g(x). \tag{5.2.8}$$

We now divide both sides in (5.2.7) by $r_n = \mathsf{d}(x,y_n)$ and let $n \to \infty$. From (5.2.8) and (5.2.6) we get

$$|\nabla u|(x) \le 2C\big(M_q^{\lambda} g(x)\big)^{1/q} + \lambda L.$$

Since this inequality holds for $\mathfrak{m}$-a.e. x, we can choose an infinitesimal sequence $(\lambda_k) \subset (0,1)$ and use the $\mathfrak{m}$-a.e. convergence of $M_q^{\lambda_k} g$ to g to obtain (5.2.5). □

Theorem 5.2.4 *Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric measure space with $\mathfrak{m}$ doubling, which supports a weak $(1,q)$-Poincaré inequality and satisfies* $\operatorname{supp}\mathfrak{m} = X$*. Then, for any open set $A \subset X$ it holds*

$$u_n,\, u \in \mathrm{Lip}_{\mathrm{loc}}(A),\ u_n \to u \ \text{ in } L^1_{\mathrm{loc}}(A) \quad \Longrightarrow \quad \liminf_{n\to\infty} \int_A |\nabla u_n|^q \,\mathrm{d}\mathfrak{m} \geq \int_A |\nabla u|^q \,\mathrm{d}\mathfrak{m}. \tag{5.2.9}$$

In particular, it holds $|\nabla u| = |\nabla u|_{w,q}$ $\mathfrak{m}$-a.e. in X for all $u \in \mathrm{Lip}_{\mathrm{loc}}(X) \cap L^q$.

Proof. By a simple truncation argument we can assume that all functions u_n are uniformly bounded, since $|\nabla(M \wedge v \vee -M)| \leq |\nabla v|$ and $|\nabla(M \wedge v \vee -M)| \uparrow |\nabla v|$ as $M \to \infty$. Possibly extracting a subsequence we can also assume that the lim inf in the right-hand side of (5.2.9) is a limit and, without loss of generality, we can also assume that it is finite. Fix a bounded open set B with $\mathrm{dist}(B, X \setminus A) > 0$ and let $\psi : X \to [0,1]$ be a cut-off Lipschitz function identically equal to 1 on a neighborhood of B, with support bounded and contained in A. It is clear that the functions $v_n := u_n\psi$ and $v := u\psi$ are globally Lipschitz, $v_n \to v$ in $L^q(X, \mathfrak{m})$ and (v_n) is bounded in $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$.

From the reflexivity of this space proved in Corollary 5.1.5 we have that, possibly extracting a subsequence, (v_n) weakly converges in the Sobolev space to a function w. Using Mazur's lemma, we construct another sequence $(\hat v_n)$ that is converging strongly to w in $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$ and $\hat v_n$ is a finite convex combination of $v_n, v_{n+1}, \ldots$. In particular we get $\hat v_n \to w$ in $L^q(X, \mathfrak{m})$ and this gives $w = v$. Moreover,

$$\int_B |\nabla \hat v_n|^q \,\mathrm{d}\mathfrak{m} \leq \sup_{k \geq n} \int_B |\nabla v_k|^q \,\mathrm{d}\mathfrak{m}.$$

Eventually, from Proposition 5.2.3 applied to the functions $v - \hat v_n$ we get:

$$\begin{aligned}
\left(\int_B |\nabla v|^q \,\mathrm{d}\mathfrak{m}\right)^{1/q} &\leq \liminf_{n\to\infty} \left\{ \left(\int_B |\nabla \hat v_n|^q \,\mathrm{d}\mathfrak{m}\right)^{1/q} + \left(\int_B |\nabla (v - \hat v_n)|^q \,\mathrm{d}\mathfrak{m}\right)^{1/q} \right\} \\
&\leq \limsup_{n\to\infty} \left\{ \left(\int_B |\nabla v_n|^q \,\mathrm{d}\mathfrak{m}\right)^{1/q} \right\} + C \limsup_{n\to\infty} \|v - \hat v_n\|_{W^{1,q}} \\
&= \limsup_{n\to\infty} \left(\int_B |\nabla v_n|^q \,\mathrm{d}\mathfrak{m}\right)^{1/q}.
\end{aligned}$$

Since $v_n \equiv u_n$ and $v \equiv u$ on B we get

$$\int_B |\nabla u|^q \,\mathrm{d}\mathfrak{m} \leq \limsup_{n\to\infty} \int_B |\nabla u_n|^q \,\mathrm{d}\mathfrak{m} \leq \lim_{n\to\infty} \int_A |\nabla u_n|^q \,\mathrm{d}\mathfrak{m}$$

and letting $B \uparrow A$ gives the result. □

Remark 5.2.5 *An important consequence of Theorem 5.2.4 is that the weak gradient $|\nabla f|_{q'}$ does not depend on q' for $q' \geq q$. In fact this is obvious when f is Lipschitz since Jensen inequality gives that $(1,q)$-Poincaré implies $(1,q')$-Poincaré, and so $|\nabla f|_q = |\nabla f| = |\nabla f|_{q'}$. Then we can use Proposition 5.1.6, i.e. the density of Lipschitz functions in $W^{1,q}$ and $W^{1,q'}$, in order to conclude (see for example Corollary A9 in [17]).*

5.3 Discrete gradients for general metric spaces

Here we provide another type of approximation via discrete gradients which doesn't even require the space (X, d) to be doubling; moreover it can be adapted to give a discrete approximation also in the case of Orlicz-Sobolev spaces (at least in the case Ψ doubling).

We slightly change the definition of discrete gradient: instead of taking the sum of the finite differences, that is forbidden due to the fact that the number of terms can not in general be uniformly bounded from above, we simply take the supremum among the finite differences. Let us fix a decomposition A_i^δ of $\operatorname{supp}\mathfrak{m}$ as in Lemma 5.1.1. Let $u \in L^q(X, \mathfrak{m})$ and denote by $u_{\delta,i}$ the mean of u in A_i^δ as before. We consider the discrete gradient

$$|\mathcal{D}_\delta u|_\infty(x) = \frac{1}{\delta} \sup_{A_j^\delta \sim A_i^\delta} \{|u_{\delta,i} - u_{\delta,j}|\} \qquad \forall x \in A_i^\delta.$$

Then we consider the functional $\mathcal{F}_\delta^\infty : L^q(X, \mathfrak{m}) \to [0, \infty]$ given by

$$\mathcal{F}_\delta^\infty(u) := \int_X |\mathcal{D}_\delta(u)|_\infty^q(x) \, \mathrm{d}\mathfrak{m}(x).$$

With these definitions, the following theorem holds.

Theorem 5.3.1 *Let $(X, \mathsf{d}, \mathfrak{m})$ be a Polish metric measure space with $\mathfrak{m}$ finite on bounded sets. Let $\mathcal{F}_q^\infty$ be a Γ-limit point of $\mathcal{F}_{q,\delta}^\infty$ as $\delta \downarrow 0$, namely*

$$\mathcal{F}_q^\infty := \Gamma - \lim_{k \to \infty} \mathcal{F}_{q,\delta_k}^\infty,$$

where $\delta_k \to 0$ and the Γ-limit is computed with respect to the $L^q(X, \mathfrak{m})$-distance. Then the functional $\mathcal{F}_q^\infty$ is equivalent to Cheeger's energy, namely there exists a constant $\eta_\infty = \eta_\infty(q)$ such that

$$\frac{1}{\eta_\infty} \mathrm{Ch}_q(u) \le \mathcal{F}_q^\infty(u) \le \eta_\infty \mathrm{Ch}_q(u) \qquad \forall u \in L^q(X, \mathfrak{m}). \tag{5.3.1}$$

The proof follows closely the one of Theorem 5.1.4. An admissible choice for η_∞ is 6^q.

5.4 Optimality of the Poincaré assumption for the lower semicontinuity of slope.

This is still an open problem. As shown to us by P. Koskela, the doubling assumption, while sufficient to provide reflexivity of the Sobolev spaces $W^{1,q}(X, \mathsf{d}, \mathfrak{m})$, is not sufficient to ensure the lower semicontinuity (11) of slope. Indeed, one can consider for instance the Von Koch snowflake $X \subset \mathbb{R}^2$ endowed with the Euclidean distance. Since X is a self-similar fractal satisfying Hutchinson's open set condition (see for instance [36]), it follows that X is Ahlfors regular of dimension $\alpha = \ln 4 / \ln 3 \in (1, 2)$, namely $0 < \mathscr{H}^\alpha(X) < \infty$, where $\mathscr{H}^\alpha$ denotes α-dimensional Hausdorff measure in $\mathbb{R}^2$. Using self-similarity it is easy to check that $(X, \mathsf{d}, \mathscr{H}^\alpha)$ is doubling. However, since absolutely continuous curves with values in X are constant, the q-weak upper gradient of any Lipschitz function f vanishes. Then, the equivalence of weak and relaxed gradients gives $|\nabla f|_{*,q} = 0$ $\mathscr{H}^\alpha$-a.e. on X. By Proposition 5.0.1 we obtain Lipschitz functions f_n convergent to f in $L^q(X, \mathscr{H}^\alpha)$ and satisfying

$$\lim_{n \to \infty} \int_X \mathrm{lip}_a^q(f_n, x) \, \mathrm{d}\mathscr{H}^\alpha(x) = 0.$$

Since $\mathrm{lip}_a(f_n, \cdot) \geq |\nabla f_n|$, if $|\nabla f|$ is not trivial we obtain a counterexample to (11).

One can easily show that any linear map, say $f(x_1, x_2) = x_1$, has a nontrivial slope on X at least $\mathscr{H}^\alpha$-a.e. in X. Indeed, $|\nabla f|(x) = 0$ for some $x \in X$ implies that the geometric tangent space to X at x, namely all limit points as $X \ni y \to x$ of normalized secant vectors $(y - x)/|y - x|$, is contained in the vertical line $\{x_1 = 0\}$. However, a geometric rectifiability criterion (see for instance [7, Theorem 2.61]) shows that this set of points x is contained in a countable union of Lipschitz curves, and it is therefore σ-finite with respect to $\mathscr{H}^1$ and $\mathscr{H}^\alpha$-negligible.

This proves that doubling is not enough. On the other hand, quantitative assumptions weaker than the Poincaré inequality might still be sufficient to provide the result.

CHAPTER 6

The p-Weak Gradient depends on p

In this section we answer an interesting question in the context of Sobolev Spaces in metric measure spaces: whether the weak gradient $|\nabla u|_{w,p}$ depends on the choice of p. The answer is positive, and actually the example is not very difficult, and arises already in $\mathbb{R}$ with a measure of the form $w\mathscr{L}^1$, for some weight w. The weight we provide easily adapt to the $\mathbb{R}^n$ case.

A partial positive answer is given by an example due to Koskela, reported in [9]: a function f is constructed such that $f \in W^{1,2}$ and $|\nabla f|_2 = 0$, whereas $f \notin W^{1,p}$ for $p > 2$. However it was still an open question whether if $f \in W^{1,p} \cap W^{1,p'}$ can have the gradient depending on the exponent.

We recall also that if some assumption are satisfied in the measurable metric structure, then we can gain independence: in fact we already observed in Remark 5.2.5 that if $(X, \mathsf{d}, \mathfrak{m})$ is doubling and supports a $(1, p)$ Poincaré inequality then the weak gradient is independent of the exponent for $p' \geq p$. Another results of independence is given in [40], where the authors prove an even stronger statement in the case of $RCD(K, \infty)$ spaces: whenever $f \in W^{1,p}$ and $|\nabla f|_p \in L^q$ then $f \in W^{1,q}$ and $|\nabla f|_q = |\nabla f|_p$.

We recall here the definition of weak gradient, in the spirit of $N^{1,p}(X, \mathsf{d}, \mathfrak{m})$. See (2.4.1) for the definition of $\mathrm{Mod}_{p,\mathfrak{m}}$ for families of curves; in the sequel we will consider only rectifiable curves $\gamma : [ini, fin] \to X$ with constant metric speed.

Definition 6.0.1 *Let $(X, d, \mathfrak{m})$ be a metric measure space and $p \geq 1$. A Borel function $g\colon X \to [0, \infty]$ is a p-upper gradient of $f\colon X \to \mathbb{R}$ if*

$$|f(\gamma_{fin}) - f(\gamma_{ini})| \leq \int_\gamma g \, \mathrm{d}s \quad \textit{for } \mathrm{Mod}_{p,\mathfrak{m}}\textit{-a.e. curve } \gamma.$$

If $p > 1$ then the minimal p-upper gradient $|\nabla f|_{\mathfrak{m},p}$ of $f\colon X \to \mathbb{R}$ is the p-upper gradient characterized, up to $\mathfrak{m}$-negligible sets, by the property

$$|\nabla f|_{\mathfrak{m},p} \leq g \qquad \mathfrak{m}\textit{-a.e. in } X \textit{ for every p-upper gradient } g \textit{ of } f.$$

For the remainder of the paper we fix $\alpha > 0$ and denote $\beta = 1/\alpha$. We now describe how the facts about weak gradients in Theorem 1 follow from the assertions about the measure.

Let μ be the measure from Theorem 1 and consider the metric measure space $(\mathbb{R}^n, |\cdot|, \mu)$ so that $p \leq 1+\alpha$ implies $\mathrm{Mod}_{p,\mu}(\Gamma_c) = 0$, where Γ_c is the set of non constant rectifiable curves. In this case the function identically equal to zero is a p-upper gradient for every function; hence $|\nabla f|_{\mathfrak{m},p} = 0$ for any function $f\colon \mathbb{R}^n \to \mathbb{R}$.

Now we recall the notion of a Muckenhoupt A_p-weight on $\mathbb{R}^n$; we only consider the case $p > 1$ though a similar definition may be given for $p = 1$ [48]. If $(X, \mathsf{d}, \mathfrak{m})$ is a metric measure space, with $f\colon X \to \mathbb{R}$ and $A \subseteq X$ Borel measurable such that $\mathfrak{m}(A) > 0$, then we denote $f_A = \fint_A f \, \mathrm{d}\mathfrak{m} = (1/\mathfrak{m}(A)) \int_A f \, \mathrm{d}\mathfrak{m}$ whenever the quotient is well defined. If no measure is specified, integrals over subsets of $\mathbb{R}^n$ are with respect to Lebesgue measure $\mathscr{L}^n$; we also use the notation $\mathscr{L}^n(A) = |A|$.

Definition 6.0.2 *Let $p > 1$. A function $w\colon \mathbb{R}^n \to \mathbb{R}^+$ is a Muckenhoupt A_p-weight if for some constant $C > 0$ and all balls $B \subset \mathbb{R}^n$,*

$$\left(\fint_B w\right)\left(\fint_B w^{1/(1-p)}\right)^{p-1} \leq C. \tag{6.0.1}$$

Muckenhoupt A_p-weights were first introduced in [69] as precisely those weights for which the Hardy maximal function of the associated measure is bounded in L^p. The A_p condition has numerous applications, for example to weighted Sobolev spaces [26] and regularity of the solutions of degenerate elliptic equations [35].

We recall that, by Hölder's inequality, the condition of a weak p-Poincaré inequality becomes weaker as p increases. If a metric measure space equipped with a doubling measure admits a weak p-Poincaré inequality then it admits a differentiable structure [25]; in fact, a Lip-lip inequality suffices in place of a Poincaré inequality [54]. Roughly, a Lip-lip inequality states that at almost every point the variation of a Lipschitz function on small scales is independent of the precise choice of scale.

We use the fact that if w is a Muckenhoupt A_p-weight on $\mathbb{R}^n$ then the measure $\mu = w\mathcal{L}^n$ is p-admissible [48]; this means that μ is doubling and satisfies a weak p-Poincaré inequality. For $n = 1$ the converse holds: if μ is p-admissible then w must be an A_p-weight [19]. However, inequality (6.0.1) seemed easier to check than verifying Poincaré inequality directly.

If a doubling metric measure space admits a weak p-Poincaré inequality then, for Lipschitz functions, the p-upper gradient $|\nabla f|_{\mathfrak{m},p}$ agrees, up to negligible sets, with the slope [25], [54]. Hence, for $p > 1+\alpha$, if μ is the measure in Theorem 1, then the p-weak slope $|\nabla f|_{\mathfrak{m},p}$ of Lipschitz functions $f\colon \mathbb{R}^n \to \mathbb{R}$ on $(\mathbb{R}^n, |\cdot|, \mu)$ is non trivial.

We also note that, since μ is absolutely continuous with respect to Lebesgue measure, the metric measure space $(\mathbb{R}^n, |\cdot|, \mu)$ satisfies a Lip-lip inequality. Further, in any metric measure space $(X, d, \mathfrak{m})$, lower semicontinuity of the map, defined on Lipschitz functions,

$$f \mapsto \int_X |\nabla f|^p \, \mathrm{d}\mathfrak{m}$$

in L^p implies the p-weak gradient agrees with the slope for Lipschitz functions [4]. Hence we observe that a Lip-lip inequality is not sufficient for lower semicontinuity of the integral of the p-th power of the slope; this answers a question raised in [4].

We now give an idea of the construction of the weight w in Theorem 1. Firstly we suppose $n = 1$; one starts with the weight $w_1 \equiv 1$, then repeatedly defines $w_k = \min\{w_{k-1}, g_k\}$ where g_k is a scaled and translated copy of $|x|^\alpha$ centred on some rational q_k. We do this for a dense, non repeating, sequence of rationals $(q_k)_{k=1}^{\infty}$ and define $w = \inf_k w_k$. The function $1/w^s$ is locally integrable for $s < \beta$ but nowhere locally integrable for $s \geq \beta$; this discrepancy allows us to prove the first property in Theorem 1. Further, provided the copies of $|x|^\alpha$ are scaled to be sufficiently thin, each stage in the construction increases the left hand side of inequality (6.0.1) only a small amount; this allows us to prove the second property in Theorem 1. To prove Theorem 1 for general n we define $\widehat{w}(x_1, \dots, x_n) = \min\{w(x_1), \dots, w(x_n)\}$ on $\mathbb{R}^n$. Then $\widehat{w}$ has the same integrability properties as w (but now with respect to $\mathcal{L}^n$), which gives the first property, and the lattice property of A_p-weights [56] allows us to extend the second property from w to $\widehat{w}$.

6.1 Construction of the weight

Fix a sequence $\varepsilon_k > 0$ such that $\prod_{k=1}^{\infty}(1 + \varepsilon_k) < \infty$ and enumerate the rational numbers by a sequence $(q_k)_{k=1}^{\infty}$ with $q_k \neq q_l$ for $k \neq l$. We inductively define a sequence of continuous weights $w_k \colon \mathbb{R} \to \mathbb{R}^+$; among other properties the weights satisfy $w_k \leq w_{k-1}$ and $w_k(x) > 0$ if $x \notin \{q_l \colon l = 1, \dots, k\}$. Denoting by w the limit of the weights w_k we will verify Theorem 1 for the weight $\widehat{w}$ on $\mathbb{R}^n$ given by $\widehat{w}(x_1, \dots, x_n) = \min\{w(x_1), \dots, w(x_n)\}$.

Let $w_1 \colon \mathbb{R} \to \mathbb{R}^+$ be the function which is constant and equal to 1. Fix $k \in \mathbb{N}$ for which the weight w_{k-1} has been defined; we show how to define w_k. Since w_{k-1} is continuous and $w_{k-1}(q_k) > 0$ (using the properties described in the introduction) we can choose $R_k > 0$ so that

$$w_{k-1}(q_k)/2 \leq w_{k-1}(x) \leq 2w_{k-1}(q_k)$$

for $|x - q_k| \leq 4R_k$.

Fix $r_k > 0$ such that:

$$r_k \leq w_{k-1}(q_k)^{\beta}\varepsilon_k,$$

$$8r_k \leq \varepsilon_k(R_k - r_k)$$

and

$$2r_k(p - \alpha + 1)/(p - 1) \leq \varepsilon_k(R_k - r_k).$$

We let

$$g_k(x) = 2w_{k-1}(q_k)|(x - q_k)/r_k|^{\alpha}$$

for $x \in \mathbb{R}$ and define $w_k \colon \mathbb{R} \to \mathbb{R}^+$ by

$$w_k(x) = \min\{w_{k-1}(x), g_k(x)\}.$$

The function w_k is continuous, $w_k \leq w_{k-1}$ and $w_k > 0$ if $x \notin \{q_l \colon l = 1, \dots, k\}$.

Denote $I_k = (q_k - r_k, q_k + r_k)$ and note that $w_k = w_{k-1}$ outside I_k. We also define $J_k = (q_k - R_k, q_k + R_k)$, $J_k^+ = [q_k + r_k, q_k + R_k)$ and $J_k^- = (q_k - R_k, q_k - r_k)$.

Let $w \colon \mathbb{R} \to \mathbb{R}^+$ be given by $w = \inf_k w_k$. We define a Borel weight $\widehat{w} \colon \mathbb{R}^n \to \mathbb{R}^+$ by:

$$\widehat{w}(x_1, \dots, x_n) = \min\{w(x_1), \dots, w(x_n)\}$$

and let $\mu = \widehat{w}\mathcal{L}^n$.

6.2 The p-modulus on curves is trivial for small p

In this section we show that $p < 1+\alpha$ implies $\mathrm{Mod}_{p,\mu}(\Gamma_c) = 0$, where Γ_c is the family of non constant absolutely continuous curves in $\mathbb{R}^n$. This fact arises from simple integrability properties of $1/w$ on $\mathbb{R}$ which follow from corresponding properties of $1/|x|^\alpha$. Recall that $\beta = 1/\alpha$.

Lemma 6.2.1 *Let $r = e^{\alpha(\alpha+1)}$. The weight $w\colon \mathbb{R} \to \mathbb{R}^+$ has the following integrability properties:*

(1) The function $1/w^s$ is locally Lebesgue integrable if $s < \beta$.

(2) The function $1/(w^\beta |\log(w/r)|^{1+\alpha})$ is locally Lebesgue integrable.

(3) The function $1/w^s$ is nowhere locally Lebesgue integrable if $s \geq \beta$.

(4) The function $1/(w^\beta |\log(w/r)|)$ is nowhere locally Lebesgue integrable.

Proof. Suppose first $s < \beta$ and $N \in \mathbb{N}$. Clearly, for each integer $k > 1$, $w_k = w_{k-1}$ outside I_k implies

$$\int_{-N}^{N} \frac{1}{w_k^s} \leq \int_{-N}^{N} \frac{1}{w_{k-1}^s} + \int_{q_k - r_k}^{q_k + r_k} \frac{1}{w_k^s}.$$

We show the second term is relatively small. Indeed, since

$$w_k(x) \geq \frac{1}{2} w_{k-1}(q_k) |(x - q_k)/r_k|^\alpha$$

for $x \in (q_k - r_k, q_k + r_k)$ and $\alpha s < 1$, we have,

$$\begin{aligned}
\int_{q_k - r_k}^{q_k + r_k} \frac{1}{w_k^s} &\leq \frac{2^s r_k^{\alpha s}}{w_{k-1}(q_k)^s} \int_{q_k - r_k}^{q_k + r_k} \frac{1}{|x - q_k|^{\alpha s}} \\
&\leq C r_k / w_{k-1}(q_k)^s \\
&\leq C r_k / w_{k-1}(q_k)^\beta \\
&\leq C \varepsilon_k.
\end{aligned}$$

Since w_1 was constant (so trivially locally integrable) and ε_k were chosen small we deduce that the sequence $\int_{-N}^{N} 1/w_k^s$ is bounded uniformly in k. By the Monotone Convergence Theorem we obtain that $1/w^s$ is integrable on the interval $[-N, N]$.

For the second assertion a similar estimate is required: first of all the function $\Phi : t \mapsto t(-\log(t^\alpha/r))^{1+\alpha}$ is increasing in $(0,1)$, and thus we can make the estimate

$$\int_{q_k - r_k}^{q_k + r_k} \frac{1}{\Phi(w_k^\beta)} \leq \int_{q_k - r_k}^{q_k + r_k} \frac{1}{\Phi(C_k |\frac{x - q_k}{r_k}|)} = \frac{2 r_k}{C_k} F(C_k)$$

where $C_k = (w_{k-1}(q_k)/2)^\beta$ and F is the primitive of $1/\Phi$ such that $F(0) = 0$. Substituting $r_k \leq C C_k \varepsilon_k$ and using the definition of Φ we obtain

$$\int_{q_k - r_k}^{q_k + r_k} \frac{1}{w_k^\beta |\log(w_k/r)|^{1+\alpha}} \leq \frac{2 r_k}{C_k} F(1) \leq C \varepsilon_k.$$

We now obtain the required integrability as before.

Now suppose $s \geq \beta$ and I is a non empty interval. Then we can find $k \in \mathbb{N}$ for which $q_k \in I$. It follows,

$$\int_I 1/w^s \geq C(k) \int_I 1/|x - q_k|^{\alpha s}$$

and the right hand side is equal to ∞ since $\alpha s \geq 1$. In the same way we have that $w_k^\beta \log(w_k/r) \sim C(k)|x - q_k| \log |x - q_k|$ in a neighborhood of q_k and so the final statement follows. $\square$

Notice the previous lemma implies that w is nonzero outside a set of Lebesgue measure zero. We recall some elementary facts about the modulus which are valid on any metric measure space [46].

Lemma 6.2.2 *Let $(X, d, \mathfrak{m})$ be a metric measure space. The modulus* $\mathrm{Mod}_{p,m}$ *satisfies*

$$\mathrm{Mod}_{p,\mathfrak{m}}(\Gamma_a) \leq \mathrm{Mod}_{p,\mathfrak{m}}(\Gamma_b)$$

if Γ_a and Γ_b are two curve families such that each curve in Γ_a has a subcurve in Γ_b. Further, $\mathrm{Mod}_{p,\mathfrak{m}}(\Gamma) = 0$ *if and only if there is a p-integrable Borel function $g\colon X \to [0, \infty]$ such that $\int_\gamma g \,\mathrm{d}s = \infty$ for each $\gamma \in \Gamma$.*

Now we can deduce the required properties of the p-modulus on $(\mathbb{R}^n, |\cdot|, \mu)$.

Proposition 6.2.3 *Let Γ_c be the family of non constant absolutely continuous curves on $\mathbb{R}^n$ and $p \leq 1 + \alpha$. Then* $\mathrm{Mod}_{p,\mu}(\Gamma_c) = 0$.

Proof. For each $k \in \mathbb{N}$ let Γ_k be the family of non constant absolutely continuous curves with image contained in $[-k, k]^n$. Using Lemma 6.2.2 it suffices to show that $\mathrm{Mod}_{p,\mu}(\Gamma_k) = 0$ for each k.

First suppose $p < 1 + \alpha$; fix $k \in \mathbb{N}$ and recall $\beta = 1/\alpha$. Let $g\colon \mathbb{R}^n \to \mathbb{R}^+$ be equal to $1/\widehat{w}^\beta$ inside $[-k, k]^n$ and identically 0 outside $[-k, k]^n$. Suppose $\gamma = (\gamma_1, \ldots, \gamma_n) \in \Gamma_k$ and fix i such that the image of γ_i contains some non trivial interval $I \subset \mathbb{R}$. Then,

$$\begin{aligned}
\int_\gamma g \,\mathrm{d}s &\geq \int_\gamma 1/w(x_i)^\beta \,\mathrm{d}s \\
&\geq \int_{\gamma_i} 1/w(t)^\beta \,\mathrm{d}s \\
&\geq \int_I 1/w(t)^\beta \,\mathrm{d}t \\
&= \infty
\end{aligned}$$

using Lemma 6.2.1. However,

$$\begin{aligned}
\int_{\mathbb{R}^n} g^p \,\mathrm{d}\mu &= \int_{[-k,k]^n} \widehat{w}^{1-\beta p} \\
&\leq \int_{[-k,k]^n} \sum_{i=1}^n w(x_i)^{1-\beta p} \\
&\leq n(2k)^{n-1} \int_{-k}^k w(t)^{1-\beta p} \,\mathrm{d}t
\end{aligned}$$

which, by Lemma 6.2.1, is finite if $\beta p - 1 < \beta$ or, equivalently, $p < 1 + \alpha$. Hence, by Lemma 6.2.2, $\mathrm{Mod}_{p,\mu}(\Gamma_k) = 0$ and the proposition follows.

In the case $p = 1 + \alpha$ we choose $g = 1/(\widehat{w}^{\beta}|\log(\widehat{w}/r)|)$; the argument is then identical using the analogous statements about integrability from Lemma 6.2.1. $\square$

6.3 The Muckenhoupt A_p condition for large p

We suppose throughout this section that $p > 1 + \alpha$. We first show that w is a Muckenhoupt A_p-weight on $\mathbb{R}$ and then deduce $\widehat{w}$ is an A_p-weight on $\mathbb{R}^n$ using the lattice property of A_p-weights [56]. To verify w is a Muckenhoupt A_p-weight the idea will be that constructing w_k from w_{k-1} can increase the left side of inequality (6.0.1) only very slightly. We use a different argument depending on whether the ball in (6.0.1) is relatively small or relatively large.

It will be important during the proof that $|x|^\alpha$ is a Muckenhoupt A_p-weight on $\mathbb{R}$; this fact is well known (for example see Remark 4 [19]; this is also valid in $\mathbb{R}^n$ provided $p > 1 + n\alpha$) but we prefer to provide here a self-contained proof.

Lemma 6.3.1 *The function $g(x) = |x|^\alpha$ on $\mathbb{R}$ is an A_p-weight.*

Proof. Let $I = [a, b]$ be an interval. Denote $I^+ = I \cap [0, +\infty)$ and $I^- = I \cap (-\infty, 0]$. Without loss of generality we can assume that $|I^+| \geq |I^-|$; in this case we have that $I^- \subseteq -I^+$ and so, using that g is an even function, we have

$$\left(\fint_I g\right)\left(\fint_I g^{1/(1-p)}\right)^{p-1} \leq 2^p \left(\fint_{I^+} g\right)\left(\fint_{I^+} g^{1/(1-p)}\right)^{p-1}.$$

Hence it is sufficient to prove (6.0.1) only for intervals $I = [a, b]$ such that $0 \leq a < b$.
We distinguish two cases:

- $2a \geq b$. In this case, given the monotonicity of g we can estimate each of the factors in the left hand side of (6.0.1) with the values of the integrand at the endpoint: in particular we can estimate it from above by $g(b)/g(a) \leq 2^\alpha$.

- $2a < b$. In this case we have that $1/(b-a) \leq 2/b$ and so

$$\fint_a^b x^\alpha \,\mathrm{d}x \leq \frac{1}{b-a}\int_0^b x^\alpha \,\mathrm{d}x \leq \frac{2b^\alpha}{\alpha+1};$$

$$\fint_a^b x^{\alpha/(1-p)} \,\mathrm{d}x \leq \frac{1}{b-a}\int_0^b x^{\alpha/(1-p)} \,\mathrm{d}x \leq \frac{2b^{\alpha/(1-p)}}{\alpha/(1-p)+1}.$$

  These two inequalities together give us precisely (6.0.1), with C depending only on α and p.

$\square$

The following Lemma will be used to estimate (6.0.1) for relatively small intervals; the idea will be that early stages in the construction play no role on small scales.

Lemma 6.3.2 *Suppose $q \in \mathbb{R}$, $R > 0$ and $f\colon (q-R, q+R) \to \mathbb{R}^+$ is Borel with $L/2 \le f \le 2L$ for some $L > 0$.*

Let $0 < r < R$ and $g(x) = 2L|(x-q)/r|^\alpha$ for $x \in \mathbb{R}$.

Define $h\colon (q-R, q+R) \to \mathbb{R}^+$ by

$$h(x) = \min\{f(x), g(x)\}.$$

Then for any interval $I \subset (q-R, q+R)$ we have,

$$\left(⨍_I h\right)\left(⨍_I h^{1/(1-p)}\right)^{p-1} \le C \tag{6.3.1}$$

where the constant $C > 0$ depends only on α and p.

Proof. Fix an interval $I = (a, b) \subset (q-R, q+R)$; we consider several cases depending on the length and position of I.

Suppose $|b-a| > r/8^\beta$. We have the simple estimate

$$⨍_I h \le ⨍_I f \le 2L. \tag{6.3.2}$$

For the second term in (6.3.1) we use the bounds on f and the fact that $h = f$ outside $(q-r, q+r)$ to see

$$\int_I h^{1/(1-p)} \le \int_{q-r}^{q+r} g^{1/(1-p)} + CL^{1/(1-p)}|I|.$$

Using the fact $p > 1 + \alpha$ and $r < 8^\beta|I|$ we can continue,

$$\begin{aligned}\int_{q-r}^{q+r} g^{1/(1-p)} &= (2L/r^\alpha)^{1/(1-p)} \int_0^r |x|^{\alpha/(1-p)} \\ &\le CL^{1/(1-p)} r \\ &\le CL^{1/(1-p)}|I|.\end{aligned}$$

Thus we obtain

$$\left(⨍_I h^{1/(1-p)}\right)^{p-1} \le CL^{-1}$$

and, by combining this with (6.3.2), we obtain (6.3.1).

Now suppose $|b-a| \le r/8^\beta$ and $I \subset [q-(r/4^\beta), q+(r/4^\beta)]$. Then $h = g$ on I and (6.3.1) follows from Lemma 6.3.1.

Finally suppose $|b-a| \le r/8^\beta$ and I is not strictly contained in the interval $[q-(r/4^\beta), q+(r/4^\beta)]$. This implies that $|x-q| \ge r/4^\beta - r/8^\beta$ for all $x \in I$; it follows that the values of g, and hence the values of h, on I are comparable to L. In this case the validity of (6.3.1) is again clear. □

The next lemma will be used to estimate (6.0.1) for relatively large intervals; the idea is that w_k and w_{k-1} agree except on a relatively small interval.

Lemma 6.3.3 *The following estimates hold:*

$$\int_{I_k} w_k \le \varepsilon_k \int_{J_k^+} w_{k-1},$$

$$\int_{I_k} w_k^{1/(1-p)} \le \varepsilon_k \int_{J_k^+} w_{k-1}^{1/(1-p)}.$$

The same estimate hold also if we put J_k^- instead of J_k^+.

Proof. We only prove them for J_k^+, the other ones being similar. Let $L = w_{k-1}(q_k)$. For the first estimate we note,

$$\int_{I_k} w_k \le 2|I_k| w_{k-1}(q_k) = 4r_k L$$

and

$$\int_{J_k^+} w_{k-1} \ge L/2(R_k - r_k)$$

so the estimate holds since R_k was chosen sufficiently large relative to r_k. The argument for the second estimate is similar: we have, since $p > 1 + \alpha$,

$$\int_{I_k} w_k^{1/(1-p)} \le \int_{-r_k}^{r_k} \left(\left| \frac{x}{r_k} \right|^\alpha L \right)^{1/(1-p)} = 2r_k L^{1/(1-p)} \frac{p-1}{p-1-\alpha},$$

$$\int_{J_k^+} w_{k-1}^{1/(1-p)} \ge (2L)^{1/(1-p)} (R_k - r_k)$$

and again, since R_k are sufficiently large relative to r_k, we get the conclusion. □

We now put together Lemma 6.3.2 and Lemma 6.3.3 to obtain the required control on inequality (6.0.1) for the weights w_k used to construct w.

Lemma 6.3.4 *There exists a constant $C > 0$, depending only on p and α, such that for all intervals I,*

$$\left(\fint_I w_k \right) \left(\fint_I w_k^{1/(1-p)} \right)^{p-1} \le \max \left\{ (1+\varepsilon_k)^p \left(\fint_I w_{k-1} \right) \left(\fint_I w_{k-1}^{1/(1-p)} \right)^{p-1}, C \right\}.$$

Proof. We clearly can assume $I \cap I_k \neq \varnothing$ since $w_k = w_{k-1}$ outside I_k. First suppose $|I| > |J_k|$ so that (without loss of generality) $J_k^+ \subset I$. Using Lemma 6.3.3 we can estimate,

$$\begin{aligned}
\fint_I w_k &= \frac{1}{|I|} \left(\int_{I_k} w_k + \int_{I \setminus I_k} w_k \right) \\
&\le \frac{1}{|I|} \left(\varepsilon_k \int_{J_k^+} w_{k-1} + \int_{I \setminus I_k} w_{k-1} \right) \\
&\le \frac{1}{|I|} \left(\varepsilon_k \int_I w_{k-1} + \int_I w_{k-1} \right) \\
&= (1+\varepsilon_k) \fint_I w_{k-1}.
\end{aligned}$$

One obtains the estimate

$$\left(\fint_I w_k^{1/(1-p)}\right)^{p-1} \le (1+\varepsilon_k)^{p-1}\left(\fint_I w_{k-1}^{1/(1-p)}\right)^{p-1}$$

in exactly the same way. Hence we obtain the desired inequality for this interval I.

Next we suppose $|I| \le |J_k|$ so that $I \subset (q_k - 4R_k, q_k + 4R_k)$. Then, from the construction of w_k, we have

$$w_{k-1}(q_k)/2 \le w_{k-1}(x) \le 2w_{k-1}(q_k)$$

whenever $|x - q_k| \le 4R_k$. By applying Lemma 6.3.2 with $q = q_k$, $R = 4R_k$, $f = w_{k-1}$, $L = w_{k-1}(q_k)$, $r = r_k$ and $g = g_k$ we obtain

$$\left(\fint_I w_k\right)\left(\fint_I w_k^{1/(1-p)}\right)^{p-1} \le C$$

with constant C depending only on p and α. This proves the claimed inequality. □

By iterating Lemma 6.3.4 we can easily show that w is an A_p-weight on $\mathbb{R}$; combining this with the lattice property of A_p-weights will then show that $\widehat{w}$ is an A_p-weight on $\mathbb{R}^n$.

Proposition 6.3.5 *If $p > 1+\alpha$ then $\widehat{w}$ is an A_p-weight on $\mathbb{R}^n$.*

Proof. By repeated application of Lemma 6.3.4 and the fact ε_k can be chosen small we deduce

$$\left(\fint_I w_k\right)\left(\fint_I w_k^{1/(1-p)}\right)^{p-1}$$

is bounded uniformly in k and I. Using the monotone convergence theorem we deduce that

$$\left(\fint_I w\right)\left(\fint_I w^{1/(1-p)}\right)^{p-1}$$

is bounded uniformly in I. This shows that w is an A_p-weight on $\mathbb{R}$.

We now observe that

$$x = (x_1, \ldots, x_n) \mapsto \eta_i(x) := w(x_i)$$

is an A_p-weight on $\mathbb{R}^n$ for each $1 \le i \le n$. Indeed; we may use cubes instead of Euclidean balls in the left hand side of (6.0.1) and then the left hand side of (6.0.1), corresponding to the weight η_i, reduces to the corresponding expression for the weight w on $\mathbb{R}$. Such an expression is obviously bounded since w is an A_p-weight on $\mathbb{R}$.

By [56, Proposition 4.3] the minimum of a finite collection of A_p-weights is again an A_p-weight; hence $\widehat{w} = \min\{\eta_1, \ldots, \eta_n\}$ is an A_p-weight. □

Taken together, Proposition 6.2.3 and Proposition 6.3.5 prove Theorem 1.

6.4 Characterization of the weak gradient on $\mathbb{R}$

Let μ be a locally finite Borel measure on $\mathbb{R}$. We give a characterization of the p-weak gradient for Lipschitz functions defined on $(\mathbb{R}, |\cdot|, \mu)$. The idea is that integrability properties of the absolutely continuous part of μ give information about which intervals (considered as curves) have non trivial p-modulus; these intervals then determine the p-weak gradient. A similar characterization has been found in [18], for measures μ whose absolutely continuous part with respect to Lebesgue measure is bounded by below by a constant, and a weaker result is stated in [21], Theorem 2.6.4, where the author characterize the measures for which the p-weak gradient is $|f'|$ for every $f \in C^\infty$ (which is equivalent to the closability of the Sobolev norm he considers).

It is worth noticing that, at least when $p = 2$, a very similar question has been investigated by some authors in the calculus of variations, posed as a semicontinuity problem; in [38], [66] they found exactly the same answer that we find.

Throughout this section we fix $p > 1$ and let q be the corresponding Hölder conjugate so that $p^{-1} + q^{-1} = 1$. Given a compact interval $I \subset \mathbb{R}$ we define the corresponding curve $\gamma_I \colon I \to \mathbb{R}$ by $\gamma_I(t) = t$. Denote the Lebesgue decomposition of μ by $\mu = \mu_a + \mu_s$. Let $\mu_a = f_a \mathcal{L}^1$ with $f_a \colon \mathbb{R} \to \mathbb{R}$ a Borel function and fix a Lebesgue null set $N \subset \mathbb{R}$ on which μ_s is concentrated.

Lemma 6.4.1 *For any interval $[a, b] \subset \mathbb{R}$ we have $\mathrm{Mod}_{p,\mu}\big(\{\gamma_{[a,b]}\}\big) > 0$ if and only if $f_a^{1/(1-p)}$ is Lebesgue integrable on $[a, b]$.*

Proof. This lemma is an easy corollary of Theorem 5.1 in [6]; however we want to give here a self-contained and more elementary proof since Γ consists of only one curve. If $a = b$ the statement is trivial so we assume $a < b$.

We write an equivalent definition for $\mathrm{Mod}_{p,\mu}$, using the homogeneity of the problem (see [6]):

$$\mathrm{Mod}_{p,\mu}(\{\gamma_{[a,b]}\})^{1/p} = \inf \left\{ \frac{\|g\|_{L^p(\mu)}}{\int_a^b g(x)\,\mathrm{d}x} \right\}, \tag{6.4.1}$$

where the infimum is taken over all Borel functions g which are p-integrable with respect to μ (this set is non empty since μ is locally finite).

Let $g \colon \mathbb{R} \to \mathbb{R}$ be any Borel function. From Hölder's inequality we have

$$\int_a^b g(x)\,\mathrm{d}x \le \left(\int_a^b g^p(x) f_a(x)\,\mathrm{d}x \right)^{1/p} \left(\int_a^b f_a(x)^{1/(1-p)}\,\mathrm{d}x \right)^{1/q}. \tag{6.4.2}$$

Now, if $f_a^{1/(1-p)}$ is $\mathcal{L}^1$ integrable on $[a, b]$, by using inequality (6.4.2) in (6.4.1) we get that

$$\mathrm{Mod}_{p,\mu}(\{\gamma_{[a,b]}\})^{1/p} \ge \inf \left\{ \frac{\|g\|_{L^p(\mu_a)}}{\int_a^b g(x)\,\mathrm{d}x} \right\} \ge \frac{1}{\|f_a^{1/(1-p)}\|_{L^1(\mathcal{L}^1)}^{1/q}} > 0.$$

If otherwise $f_a^{1/(1-p)}$ is not integrable then, letting $f_\varepsilon = \max\{f_a, \varepsilon\}$, we use

$$g(x) = \begin{cases} 0 & \text{if } x \in N \cup (\mathbb{R} \setminus [a, b]) \\ f_\varepsilon^{1/(1-p)}(x) & \text{otherwise} \end{cases}$$

as a test function in (6.4.1) and using $\mu_a \le f_\varepsilon \mathcal{L}^1$ we get

$$\mathrm{Mod}_{p,\mu}(\{\gamma_{[a,b]}\})^{1/p} \le \left(\int_a^b f_\varepsilon^{1/(1-p)}(x)\,\mathrm{d}x \right)^{-1/q}.$$

Letting $\varepsilon \to 0$ we obtain, by monotone convergence, that $\mathrm{Mod}_{p,\mu}(\{\gamma_{[a,b]}\}) = 0$. $\square$

Theorem 6.4.2 *Let*

$$\mathcal{N}_p = \Big\{ x \in \mathbb{R} \text{ such that } f_a^{1/(1-p)} \text{ is integrable on a neighbourhood of } x \Big\}. \tag{6.4.3}$$

Let $f\colon \mathbb{R} \to \mathbb{R}$ be Lipschitz and define, for μ almost every x,

$$|df|_{p,\mu}(x) = \begin{cases} |f'(x)| & \text{if } x \in \mathcal{N}_p \setminus N \\ 0 & \text{otherwise.} \end{cases} \tag{6.4.4}$$

Then $|\nabla f|_{p,\mu}(x) = |df|_{p,\mu}(x)$ for μ-almost every x.

Proof. We first note that equation (6.4.4) makes sense because f' exists $\mathcal{L}^1$-almost everywhere, by Rademacher theorem, and so it exists also μ_a-almost everywhere; hence f' exists μ-almost everywhere in the complement of N. We note that, thanks to Lemma 6.4.1, we have the following equivalent definition for $\mathcal{N}_p$:

$$\mathcal{N}_p = \bigcup_{\varepsilon > 0} \big\{ x \in \mathbb{R} \text{ such that } \mathrm{Mod}_{p,\mu}\big(\{\gamma_{[x-\varepsilon,x+\varepsilon]}\}\big) > 0 \big\}. \tag{6.4.5}$$

Denote by B the set of points where f is not differentiable. Set

$$G_f = \{ g : \mathbb{R} \to [0,\infty) \text{ bounded Borel function } : \ g(x) \geq |f'(x)| \text{ for } \mathcal{L}^1\text{-a.e. } x \in \mathcal{N}_p \}.$$

We will prove that G_f is exactly the set of bounded p-upper gradients for f. This implies the theorem: indeed, $|df|_{p,\mu} \in G_f$ and for any $g \in G_f$ we have that $g(x) \geq |df|_{p,\mu}(x)$ for μ almost every $x \in \mathbb{R}$.

Step 1. g a bounded p-upper gradient $\Longrightarrow g \in G_f$.
Let D_p be the set of Lebesgue points of g with respect to the Lebesgue measure. Since g is a bounded Borel function, we know that $\mathcal{L}^1(D_p^c) = 0$. Now take a point $x \in \mathcal{N}_p \cap D_p \setminus (B \cup N)$. Thus there exists ε such that $\mathrm{Mod}_{p,\mu}\big(\{\gamma_{[x-\varepsilon,x+\varepsilon]}\}\big) > 0$; but then $\mathrm{Mod}_{p,\mu}\big(\{\gamma_{[x-\delta,x+\delta]}\}\big) > 0$ for every $0 < \delta \leq \varepsilon$. This, together with the definition of the p-upper gradient, gives us that

$$|f(x+\delta) - f(x-\delta)| \leq \int_{x-\delta}^{x+\delta} |\nabla f|_{p,\mu}(s)\,\mathrm{d}s,$$

and so, passing to the limit when $\delta \to 0$, we get that $|f'(x)| \leq g(x)$, and so the thesis.

Step 2. $g \in G_f \Longrightarrow g$ is a p-upper gradient.
To prove this implication we first show that

$$\Gamma = \{ \gamma \ : \ \gamma \text{ has end points } a < b,\ (a,b) \cap \mathcal{N}_p^c \neq \emptyset \}$$

is $\mathrm{Mod}_{p,\mu}$-null. Let $\mathcal{B}_p = \mathcal{N}_p^c$. First let $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{B}_p$ be a set of points dense in $\mathcal{B}_p$. From the definition of $\mathcal{N}_p$ we know that for every n there exists a non negative function $f_n \in L^p(\mathbb{R},\mu)$ such that f_n is not locally Lebesgue integrable at x_n, that is:

$$\int_{x_n-\varepsilon}^{x_n+\varepsilon} f_n(s)\,\mathrm{d}s = \infty \qquad \forall \varepsilon > 0. \tag{6.4.6}$$

Now we take $f = \sum_n a_n f_n$ where the a_n are positive real numbers small enough so that f belongs to $L^p(\mathbb{R}, \mu)$. For every curve $\gamma \in \Gamma$ with end points $a < b$ we have that $x_n \in (a, b)$ for some n (since $\{x_n\}_{n\in\mathbb{N}}$ were dense in $\mathcal{B}_p$) and so we have that $[x_n - \varepsilon, x_n + \varepsilon] \subset (a, b)$ for $\varepsilon > 0$ small enough. In particular, using (6.4.6),

$$\int_\gamma f \geq \int_a^b f(s)\mathrm{d}s \geq a_n \int_a^b f_n(s)\mathrm{d}s \geq a_n \int_{x_n-\varepsilon}^{x_n+\varepsilon} f_n(s)\,\mathrm{d}s = \infty$$

and so $\mathrm{Mod}_{p,\mu}(\Gamma) = 0$.
Suppose $g \in G_f$ and $\gamma \notin \Gamma$ has end points $a < b$. Then $(a, b) \subset \mathcal{N}_p$ and hence,

$$|f(a) - f(b)| \leq \int_a^b |f'(x)|\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x \leq \int_\gamma g.$$

Thus the set of curves where the upper gradient property fails is a p negligible set; therefore g is a p-upper gradient of f.

□

Remark 6.4.3 *It seems that one can generalize the observations in section 5 about weak gradients on $\mathbb{R}$ to analogous statements about $\mathbb{R}^n$; the statement here should be that the weak gradient at a point is the restriction to a subspace (depending on the point and the measure) of the ordinary derivative. This generalization involves the equivalent definition of weak gradient from [25] as an integrand whose integral represents the Cheeger energy. The Cheeger energy is a functional obtained by relaxing the integral of the slope using convergence of Lipschitz functions; the paper [22] provides integral representations of many such functionals. Unfortunately, when $n > 1$, apart from peculiar cases, it is not possible to give a concrete description of the subspaces but a rather abstract one. In Section 7.4 we generalize this abstract characterization to Banach spaces.*

CHAPTER 7

A definition via derivations and integration by parts

7.1 Sobolev spaces via derivations

Here $(X, \mathsf{d}, \mathfrak{m})$ will be any complete separable metric measure space, where $\mathfrak{m}$ is a nonnegative Borel measure, finite on bounded sets; in particular we don't assume structural assuptions, namely doubling measure nor a Poincaré inequality are required to hold. In the sequel we will denote by $\mathrm{Lip}_0(X, \mathsf{d})$ the set of Lipschitz functions with bounded support, and with $L^0(X, \mathfrak{m})$ the set of measurable function on X, without integrability assumption.

In this Chapter we will use the notations for the Sobolev spaces as we did for the introduction: in particular $H^{1,p}$ will denote the Sobolev space made up by relaxation, $BL^{1,p}$ will denote the Sobolev space made up by looking at curves and $W^{1,p}$ will be the new one, made up with an integration by parts formula (with derivations).

7.1.1 Derivations

We state precisely what we mean here by derivations:

Definition 7.1.1 *A derivation $\boldsymbol{b}$ is a linear map $\boldsymbol{b} : \mathrm{Lip}_0(X, \mathsf{d}) \to L^0(X, \mathfrak{m})$ such that the following properties hold:*

(i) (Leibniz rule) for every $f, g \in \mathrm{Lip}_0(X, \mathsf{d})$, we have $\boldsymbol{b}(fg) = \boldsymbol{b}(f)g + f\boldsymbol{b}(g)$;

(ii) (Weak locality) There exists some function $g \in L^0(X, \mathfrak{m})$ such that

$$|\boldsymbol{b}(f)|(x) \le g(x) \cdot \mathrm{lip}_a f(x) \qquad \textit{for } \mathfrak{m}\textit{-a.e. } x, \forall f \in \mathrm{Lip}_0(X, \mathsf{d}).$$

The smallest function g with this property is denoted by $|b|$.

From now on, we will refer to the set of derivation as $\mathrm{Der}(X, \mathsf{d}, \mathfrak{m})$ and when we write $\boldsymbol{b} \in L^p$ we mean $|\boldsymbol{b}| \in L^p$. Since the definition of derivation is local on open sets we can extend $\boldsymbol{b}$ to locally Lipschitz functions. In order to get to (1), we need also the definition of divergence, and this is done simply imposing the integration by parts formula: whenever

$\boldsymbol{b} \in L^1_{\mathrm{loc}}$ we define $\operatorname{div} \boldsymbol{b}$ as the operator that maps $\mathrm{Lip}_0(X, \mathsf{d}) \ni f \mapsto -\int_X \boldsymbol{b}(f) \, d\mathfrak{m}$ (whenever this makes sense). We will say $\operatorname{div} \boldsymbol{b} \in L^p$ when this operator has an integral representation via an L^p function: $\operatorname{div} \boldsymbol{b} = h \in L^p$ if

$$-\int_X \boldsymbol{b}(f) \, d\mathfrak{m} = \int_X h \cdot f \, d\mathfrak{m} \qquad \forall f \in \mathrm{Lip}_0(X, \mathsf{d}).$$

It is obvious that if $\operatorname{div} \boldsymbol{b} \in L^p$, then is unique. Now we set

$$\mathrm{Der}^p(X, \mathsf{d}, \mathfrak{m}) = \big\{\boldsymbol{b} \in \mathrm{Der}(X, \mathsf{d}, \mathfrak{m}) \ : \ \boldsymbol{b} \in L^p(X, \mathfrak{m})\big\}$$

$$\mathrm{Der}^{p_1,p_2}(X, \mathsf{d}, \mathfrak{m}) = \big\{\boldsymbol{b} \in \mathrm{Der}(X, \mathsf{d}, \mathfrak{m}) \ : \ \boldsymbol{b} \in L^{p_1}(X, \mathfrak{m}), \operatorname{div} \boldsymbol{b} \in L^{p_2}(X, \mathfrak{m})\big\}$$

We will often drop the dependence on $(X, \mathsf{d}, \mathfrak{m})$ when it is clear. We notice that Der, Der^p and Der^{p_1,p_2} are real vector spaces, the last two being also Banach spaces endowed respectively with the norm $\|\boldsymbol{b}\|_p = \||\boldsymbol{b}|\|_p$ and $\|\boldsymbol{b}\|_{p_1,p_2} = \|\boldsymbol{b}\|_{p_1} + \|\operatorname{div} \boldsymbol{b}\|_{p_2}$. For brevity we will denote $\mathrm{Der}^{\infty,\infty} = \mathrm{Der}_b$ (b stands for bounded). The last space we will consider is $D(\mathrm{div})$, that will be consisting of derivation $\boldsymbol{b}$ such that $|\boldsymbol{b}|, \operatorname{div} \boldsymbol{b} \in L^1_{\mathrm{loc}}(X, \mathfrak{m})$; it is clear that $\mathrm{Der}^{p,q} \subset D(\mathrm{div})$ for all $p, q \in [1, +\infty]$.

In the sequel we will need a simple operation on derivations, namely the multiplication by a scalar function: let $u \in L^0(X, \mathfrak{m})$, then we can consider the derivation $u\boldsymbol{b}$ that acts simply as $u\boldsymbol{b}(f)(x) = u(x) \cdot \boldsymbol{b}(f)(x)$: it is obvious that this is indeed a derivation. We now prove a simple lemma about multiplications:

Lemma 7.1.2 *Let $\boldsymbol{b}$ a derivation; then if $u \in L^0(X, \mathfrak{m})$ we have $|u\boldsymbol{b}| = |\boldsymbol{b}| \cdot |u|$. Moreover, if $u \in \mathrm{Lip}_b(X, \mathsf{d})$ and $\boldsymbol{b} \in \mathrm{Der}^{p_1,p_2}$ we have that $u\boldsymbol{b}$ is a derivation such that*

$$\operatorname{div}(u\boldsymbol{b}) = u \operatorname{div} \boldsymbol{b} + \boldsymbol{b}(u) \qquad \textit{and} \qquad u\boldsymbol{b} \in \mathrm{Der}^{p_1,p_3},$$

where $p_3 = \max\{p_1, p_2\}$; in particular we have that $\mathrm{Der}^{p,p}$ is a $\mathrm{Lip}_b(X, \mathsf{d})$-module.

Proof. Let us prove the first assertion: it is clear that $|u\boldsymbol{b}|(x) \le |\boldsymbol{b}|(x) \cdot |u(x)|$ by the definition; the other inequality is obvious in $\{u = 0\}$. In order to prove the converse inequality also in $\{u \neq 0\}$ we can choose $\boldsymbol{b}_u = u\boldsymbol{b}$ and

$$g(x) = \begin{cases} u^{-1}(x) & \text{if } u(x) \neq 0 \\ 0 & \text{otherwise,} \end{cases}$$

and then we know that $|g\boldsymbol{b}_u| \le |g| \cdot |\boldsymbol{b}_u|$. Noting that $\boldsymbol{b}(f) = g\boldsymbol{b}_u(f)$ in $\{u \neq 0\}$ for every $f \in \mathrm{Lip}_b(X, \mathsf{d})$, we get also $|\boldsymbol{b}| = |g\boldsymbol{b}_u|$ in the same set and so we get

$$|\boldsymbol{b}| = |g\boldsymbol{b}_u| \le |g| \cdot |\boldsymbol{b}_u| \le |g| \cdot |u| \cdot |\boldsymbol{b}| = |\boldsymbol{b}| \qquad \text{in } \{u \neq 0\},$$

in particular we get $|\boldsymbol{b}_u| = |\boldsymbol{b}| \cdot |u|$ in $\{u \neq 0\}$ and thus the thesis.

For the second equality we can use Leibniz rule: let $f \in \mathrm{Lip}_b(X, \mathsf{d})$, and note $\boldsymbol{b}(fu) = u\boldsymbol{b}(f) + f\boldsymbol{b}(u)$

$$\begin{aligned} -\int_X u\boldsymbol{b}(f) \, d\mathfrak{m} &= -\int_X \boldsymbol{b}(fu) \, d\mathfrak{m} + \int_X f\boldsymbol{b}(u) \, d\mathfrak{m} \\ &= \int_X fu \cdot \operatorname{div} \boldsymbol{b} \, d\mathfrak{m} + \int_X f\boldsymbol{b}(u) \, d\mathfrak{m} \\ &= \int_X f \cdot (u \operatorname{div} \boldsymbol{b} + \boldsymbol{b}(u)) \, d\mathfrak{m} \end{aligned}$$

and so, thanks to the arbitrariness of f we get $\operatorname{div}(u\boldsymbol{b}) = u \operatorname{div} \boldsymbol{b} + \boldsymbol{b}(u)$. □

Lemma 7.1.3 (Strong locality in $D(\mathrm{div})$**)** *Let* $\boldsymbol{b} \in D(\mathrm{div})$. *Then for every* $f, g \in \mathrm{Lip}(X, \mathsf{d})$ *we have*

(i) $\boldsymbol{b}(f) = \boldsymbol{b}(g)$ $\mathfrak{m}$*-almost everywhere in* $\{f = g\}$*;*

(ii) $\boldsymbol{b}(f) \le |\boldsymbol{b}| \cdot \mathrm{lip_a}(f|_C)$ $\mathfrak{m}$*-almost everywhere in* C*, for every closed set* C*;*

Proof. In order to prove (i), thanks to the linearity, it is sufficient to consider $g = 0$ and f with support contained in $B = B_r(x_0)$, where we can take $r > 0$ as small as we want; then we can conclude by linearity and weak locality. So we can suppose that both $|\boldsymbol{b}|$ and $\mathrm{div}\, \boldsymbol{b}$ are integrable in B. Now we can consider $\varphi_\varepsilon(x) = (x - \varepsilon)_+ - (x + \varepsilon)_-$; we have φ_ε is a 1-Lipschitz function such that $|\varphi_\varepsilon(x) - x| \le \varepsilon$ and $\varphi(x) = 0$ whenever $|x| \le \varepsilon$. Let $f_\varepsilon = \varphi_\varepsilon(f)$; we have $\boldsymbol{b}(f_\varepsilon)$ is a family of equi-integrable functions and so there is a subsequence converging weakly in L^1 to some function g. Moreover $f_\varepsilon \to f$ uniformly and in particular

$$\int_X \boldsymbol{b}(f_\varepsilon)\, \mathrm{d}\mathfrak{m} - \int_X \boldsymbol{b}(f)\, \mathrm{d}\mathfrak{m} = -\int_X (f_\varepsilon - f) \cdot \mathrm{div}\, \boldsymbol{b}\, \mathrm{d}\mathfrak{m} \to 0; \tag{7.1.1}$$

since this is true also for $\chi \boldsymbol{b}$ whenever $\chi \in \mathrm{Lip}_0(X, \mathsf{d})$, we obtain $\langle \chi, \boldsymbol{b}(f_\varepsilon)\rangle \to \langle \chi, \boldsymbol{b}(f)\rangle$ and so $g = \boldsymbol{b}(f)$. In particular, putting $\rho = \chi_{\{f=0\}}\mathrm{sgn}(\boldsymbol{b}(f))$ and noting that $\mathrm{lip}_a(f_\varepsilon) = 0$ in the set $\{|f| < \varepsilon\}$ we obtain

$$\int_{\{f=0\}} |\boldsymbol{b}(f)|\, \mathrm{d}\mathfrak{m} = \int_X \rho \cdot \boldsymbol{b}(f) = \lim_{\varepsilon \to 0} \int_X \rho \cdot \boldsymbol{b}(f_\varepsilon)\, \mathrm{d}\mathfrak{m} = 0.$$

For (ii) we proceed as follows: for every closed ball $\bar{B}_r(y)$ we consider the McShane extension of the function f restricted to $C \cap \bar{B}_r(y)$ and we call it g^r_y. In particular we have $f = g^r_y$ on $C \cap B_r(y)$ and $\mathrm{Lip}(g^r_y, B_r(y)) = \mathrm{Lip}(f, B_r(y) \cap C) = \mathrm{Lip}(f|_C, B_r(y))$. Applying (i) of this lemma we find that $\boldsymbol{b}(f) = \boldsymbol{b}(g^r_y)$ $\mathfrak{m}$-a.e. on $C \cap \bar{B}_r(y)$; in particular

$$|\boldsymbol{b}(f)|(x) \le |\boldsymbol{b}| \cdot \mathrm{Lip}(f|_C, B_r(y)) \qquad \mathfrak{m}\text{-a.e. on } C \cap B_r(y).$$

Since we have $B_r(y) \subset B_{2r}(x)$ whenever $x \in B_r(y)$, we obtain

$$|\boldsymbol{b}(f)|(x) \le |\boldsymbol{b}| \cdot \mathrm{Lip}(f|_C, B_{2r}(x)) \qquad \mathfrak{m}\text{-a.e. on } C \cap B_r(y);$$

now we can drop the dependance on y and then let $r \to 0$ to get the thesis. □

7.1.2 Definition via derivations

In this whole section we treat the Sobolev spaces $W^{1,p}$ with $1 \le p < +\infty$; the case of the space BV will be treated separately. We state here the main definition of Sobolev space via derivations: we want to follow the definition (1) but in place of the scalar product between the vector field and the weak gradient we assume there is simply a continuous linear map.

Definition 7.1.4 *Let* $f \in L^p(X, \mathsf{d}, \mathfrak{m})$*; then* $f \in W^{1,p}(X, \mathsf{d}, \mathfrak{m})$ *if, setting* $q = p/(p-1)$*, there exists a continuous linear map* $L_f : \mathrm{Der}^{q,q} \to L^1(X, \mathfrak{m})$ *satisfying*

$$\int_X L_f(\boldsymbol{b})\, \mathrm{d}\mathfrak{m} = -\int_X f\, \mathrm{div}\, \boldsymbol{b}\, \mathrm{d}\mathfrak{m} \qquad \text{for all } \boldsymbol{b} \in \mathrm{Der}^{q,q}, \tag{7.1.2}$$

such that $L_f(h\boldsymbol{b}) = hL_f(\boldsymbol{b})$ *for every* $h \in \mathrm{Lip}_b, \boldsymbol{b} \in \mathrm{Der}^{q,q}$. *When* $p = 1$ *we have to assume also that* L_f *can be extended to an* L^∞*-linear map in* $\mathrm{Der}^\infty_b := L^\infty \cdot \mathrm{Der}_b$.

Since from the definition it is not obvious, we prove that $L_f(\boldsymbol{b})$ is uniquely defined whenever $f \in W^{1,p}$ and $\boldsymbol{b} \in \mathrm{Der}^{q,q}$:

Remark 7.1.5 (Well posedness of L_f) *Let us fix $\boldsymbol{b} \in \mathrm{Der}^{q,q}$, $f \in W^{1,p}$; let L_f and $\tilde{L}_f$ be two different linear maps given in the definition on $W^{1,p}$. Let $h \in \mathrm{Lip}_b(X, \mathsf{d})$: using Lemma 7.1.2 we have $h\boldsymbol{b} \in \mathrm{Der}^{q,q}$ and so we can use (7.1.2) and the L^∞-linearity to get*

$$\int_X hL_f(\boldsymbol{b})\,\mathrm{d}\mathfrak{m} = \int_X L_f(h\boldsymbol{b}) = -\int_X f\,\mathrm{div}(h\boldsymbol{b})\,\mathrm{d}\mathfrak{m},$$

and the same is true for $\tilde{L}_f$. In particular, since the right hand side does not depend on L_f, we have $\int_X hL_f(\boldsymbol{b}) = \int_X h\tilde{L}_f(\boldsymbol{b})$, and thanks to the arbitrariness of $h \in \mathrm{Lip}_b(X, \mathsf{d})$ we conclude that $L_f(\boldsymbol{b}) = \tilde{L}_f(\boldsymbol{b})$ $\mathfrak{m}$-a.e. We will call this common value $\boldsymbol{b}(f)$, since it extends $\boldsymbol{b}$ on Lipschitz functions. The same result is true also for $p = 1$ and $\boldsymbol{b} \in \mathrm{Der}_b^\infty$.

Now we can give the definition of weak gradient, in some sense dual to the definition of $|\boldsymbol{b}|$:

Theorem 7.1.6 *Let $f \in W^{1,p}(X, \mathsf{d}, \mathfrak{m})$; then there exists a function $g_f \in L^p(X, \mathsf{d}, \mathfrak{m})$ such that*

$$|\boldsymbol{b}(f)| \le g_f \cdot |\boldsymbol{b}| \quad \mathfrak{m}\text{-a.e. in } X \qquad \forall \boldsymbol{b} \in \mathrm{Der}^{q,q}. \tag{7.1.3}$$

Definition 7.1.7 (p-weak gradient) *Let $f \in W^{1,p}(X, \mathsf{d}, \mathfrak{m})$. The least function g_f (in the $\mathfrak{m}$-a.e. sense) that realizes (7.1.3) is denoted with $|\nabla f|_p$, the p-weak gradient of f*

Proof. [of Theorem 7.1.6] We reduce to prove the existence of a weak gradient in the integral sense; then thanks to Lip_b-linearity we can prove the theorem. In fact if we find a function $g \in L^p(X, \mathsf{d}, \mathfrak{m})$ such that

$$\int_X \boldsymbol{b}(f)\,\mathrm{d}\mathfrak{m} \le \int_X g|\boldsymbol{b}|\,\mathrm{d}\mathfrak{m} \qquad \forall \boldsymbol{b} \in \mathrm{Der}^{q,q}, \tag{7.1.4}$$

then, choosing $\boldsymbol{b}_h = h\boldsymbol{b}$ with $h \in \mathrm{Lip}_b(X, \mathsf{d})$, we can localize the inequality thus obtaining $\boldsymbol{b}(f) \le g|\boldsymbol{b}|$; using this inequality also with the derivation $-\boldsymbol{b}$ we get (7.1.3).

So, we're given a function $f \in W^{1,p}$ and we want to find $g \in L^p$ satisfying (7.1.4); let us note that, by definition, there exists a constant $C = \|L_f\|$ such that for every $\boldsymbol{b} \in \mathrm{Der}^{q,q}$

$$\int \boldsymbol{b}(f)\,\mathrm{d}\mathfrak{m} \le \|L_f(\boldsymbol{b})\|_1 \le C\|\boldsymbol{b}\|_q \tag{7.1.5}$$

Let us consider two functionals in the Banach space $Y = L^q(X, \mathsf{d}, \mathfrak{m})$:

$$\Psi_2(h) = C\|h\|_{L^q(\mathfrak{m})} \tag{7.1.6}$$

$$\Psi_1(h) = \sup\left\{\int_X \boldsymbol{b}(f)\,\mathrm{d}\mathfrak{m} : |\boldsymbol{b}| \le h\ ,\ \boldsymbol{b} \in \mathrm{Der}^{q,q}\right\} \tag{7.1.7}$$

where the supremum of the empty set is meant to be $-\infty$. Equation (7.1.5) guarantees that

$$\Psi_1(h) \le \Psi_2(h) \qquad \forall h \in Y. \tag{7.1.8}$$

Moreover Ψ_2 is convex and continuous while we claim that Ψ_1 is concave: it is clearly positive 1-homogeneus and it is sufficient to show that

$$\Psi_1(h_1 + h_2) \geq \Psi_1(h_1) + \Psi_1(h_2).$$

We can assume that $\Psi_1(h_i) > -\infty$ for $i = 1, 2$ because otherwise the inequality is trivial. In this case for everi $\varepsilon > 0$ we can pick two derivations $\boldsymbol{b}_i \in \mathrm{Der}^q$ such that

$$\int_X \boldsymbol{b}_1(f)\,\mathrm{d}\mathfrak{m} \geq \Psi_1(h_1) - \varepsilon \qquad |\boldsymbol{b}_1| \leq h_1$$

$$\int_X \boldsymbol{b}_2(f)\,\mathrm{d}\mathfrak{m} \geq \Psi_1(h_2) - \varepsilon \qquad |\boldsymbol{b}_2| \leq h_2$$

and so we can consider $\boldsymbol{b}_1 + \boldsymbol{b}_2$ that still belongs to $\mathrm{Der}^{q,q}$ and clearly $|\boldsymbol{b}_1 + \boldsymbol{b}_2| \leq |\boldsymbol{b}_1| + |\boldsymbol{b}_2| \leq (h_1 + h_2)$ and so

$$\Psi_1(h_1 + h_2) \geq \int_X L_f(\boldsymbol{b}_1 + \boldsymbol{b}_2)\,\mathrm{d}\mathfrak{m} \geq \Psi_1(h_1) + \Psi_1(h_2) - 2\varepsilon,$$

and we get the desired inequality letting $\varepsilon \to 0$. By Hahn-Banach theorem we can find a continuous linear functional L on $L^q(X, \mathsf{d}, \mathfrak{m})$ such that

$$\Psi_1(h) \leq L(h) \leq \Psi_2(h).$$

Case $p > 1$. We know that $(L^q)^* = L^p$ and so we can find $g \in L^p$ such that $L(h) = \int_X gh\,\mathrm{d}\mathfrak{m}$. This proves the existence and moreover we have that $L(h) \leq \Psi_2(h) = C\|h\|_q$ for every $h \in Y$ and so we have also that $\|g\|_p \leq C$.

Case $p = 1$, X compact. In this case (notice that here we have to put Der^∞_b in place of $\mathrm{Der}^{q,q}$ in (7.1.7)) if we restrict $L : C_b(X) \to \mathbb{R}$ we can see it as a positive linear such that $L(h) \leq C\|h\|_\infty$ and so, thanks to the compactness of X, it can be represented as a finite measure, i.e. there exists $\mu \in \mathcal{M}_+(X)$ such that $L(h) = \int_X h\,\mathrm{d}\mu$ for every $h \in C_0(X)$ and $\mu(X) \leq C$. Now let us fix $\boldsymbol{b} \in \mathrm{Der}^\infty_b$ and let

$$h_\varepsilon(x) = \begin{cases} \frac{1}{|\boldsymbol{b}|} & \text{if } |\boldsymbol{b}|(x) \geq \varepsilon \\ \varepsilon^{-1} & \text{otherwise} \end{cases}$$

in such a way that $|h_\varepsilon \boldsymbol{b}| \leq 1$ with equality in $\{|\boldsymbol{b}| \geq \varepsilon\}$. Now let us consider for every $h \in C_0(X)$ the derivation $h \cdot h_\varepsilon \cdot \boldsymbol{b}$; we know that $|h \cdot h_\varepsilon \cdot \boldsymbol{b}| \leq |h|$ and so we can use (7.1.7) and the L^∞-linearity to infer that

$$\int_X h h_\varepsilon \boldsymbol{b}(f)\,\mathrm{d}\mathfrak{m} \leq \int_X |h|\,\mathrm{d}\mu \qquad \forall h \in C_0(X);$$

this permits us to localize the inequality to $h_\varepsilon \boldsymbol{b}(f)\mathfrak{m} \leq \mu$. Now we have a family of measures $\mathcal{F} = \{h_\varepsilon \boldsymbol{b}(f)\mathfrak{m} \ : \ \forall \boldsymbol{b} \in \mathrm{Der}^\infty_b, \forall \varepsilon > 0\}$ such that $\nu \leq \mu$ whenever $\nu \in \mathcal{F}$. Now we can consider the supremum of the measures in $\mathcal{F}$, defined as

$$\mu_{\mathcal{F}}(A) = \sup\Big\{\sum_{i=1}^N \nu_i(A_i) \ : \ \nu_i \in \mathcal{F}, \bigcup A_i \subseteq A, \ A_i \text{ disjoint } \Big\};$$

it is readily seen that this is in fact a measure, and it is the least measure ρ such that $\nu \leq \rho$ for every $\rho \in \mathcal{F}$. The existence is clear thanks to the fact that $\nu \leq \mu$, and in particular we have that $\mu_{\mathcal{F}} \leq \mu$; moreover, since for every $\nu \in \mathcal{F}$ we have that $\nu << \mathfrak{m}$, also the supremum inherits this property, in particular we have $\mu_{\mathcal{F}} = g\mathfrak{m}$ for some $g \in L^1(\mathfrak{m})$. In particular, again fixing $\boldsymbol{b} \in \mathrm{Der}_b^\infty$, we have that

$$h_\varepsilon \boldsymbol{b}(f) \leq g \quad \mathfrak{m}\text{-a.e.} \qquad \forall \varepsilon > 0; \tag{7.1.9}$$

in particular, we can divide (7.1.9) by h_ε to obtain

$$\begin{cases} \boldsymbol{b}(f) \leq g|\boldsymbol{b}| & \mathfrak{m}\text{-a.e. in } \{|\boldsymbol{b}| \geq \varepsilon\} \\ \boldsymbol{b}(f) \leq g\varepsilon & \mathfrak{m}\text{-a.e. in } \{|\boldsymbol{b}| < \varepsilon\}. \end{cases} \tag{7.1.10}$$

Since ε is arbitrary we obtain $\boldsymbol{b}(f) \leq g|\boldsymbol{b}|$ for $\mathfrak{m}$-almost every $x \in X$, that is the thesis; also in this case $p = 1$ we have $\|g\|_1 \leq \mu(X) \leq \|L_f\|$.

Case $p = 1$, X general. In order to remove the compactness assumption, for every compact non negligible set $K \subseteq X$ let us consider the two functionals in the Banach space $Y_K = L^\infty(K, \mathsf{d}, \mathfrak{m})$:

$$\Psi_2(h) = C\|h\|_{L^\infty(K,\mathfrak{m})} \tag{7.1.11}$$

$$\Psi_1(h) = \sup\left\{\int_K \boldsymbol{b}(f)\,\mathrm{d}\mathfrak{m} : |\boldsymbol{b}| \leq h \quad \mathfrak{m}\text{-a.e. on } K \ , \ \boldsymbol{b} \in \mathrm{Der}_b^\infty\right\}. \tag{7.1.12}$$

Now we can argue precisely as before to obtain $g_K \in L^1(K, \mathfrak{m})$ such that $\|g_K\|_1 \leq \|L_f\|$

$$\boldsymbol{b}(f) \leq g_K|\boldsymbol{b}| \quad \mathfrak{m}\text{-a.e. on } K \qquad \forall \boldsymbol{b} \in \mathrm{Der}^q. \tag{7.1.13}$$

Now for every increasing sequence of compact sets K_n, let us consider $g(x) = \inf_{K_n \ni x} g_{K_n}(x)$. Denoting $Y := \bigcup_n K_n$, it is easy to note that $g \in L^1(Y, \mathfrak{m})$, since $\|g\|_{L^1(Y,\mathfrak{m})} = \sup_n \|g\|_{L^1(K_n,\mathfrak{m})} \leq \sup_n \|g_{K_n}\|_{L^1(K_n,\mathfrak{m})} \leq \|L_f\|$, and we have that

$$\boldsymbol{b}(f) \leq g|\boldsymbol{b}| \quad \mathfrak{m}\text{-a.e. on } Y \qquad \forall \boldsymbol{b} \in \mathrm{Der}^q;$$

so, in order to conlcude, it is sufficient to find a sequence K_n such that $\mathfrak{m}(X \setminus \bigcup_n K_n) = 0$, but this can be done thanks to the hypothesis of $\mathfrak{m}$ finite on bounded sets (so we can find $\theta > 0$ such that $\theta\mathfrak{m}$ is finite and then apply Prokhorov theorem to $\theta\mathfrak{m}$).

□

7.2 Equivalence with other definitions

In this section we want to prove, when $p > 1$, that Definition 7.1.4 is equivalent to the other ones $H_v^{1,p}$ and $BL^{1,p}$, given in [9]. As a byproduct we obtain the equivalence also with other definitions of Sobolev Spaces, for example the one given in [25]; the approach is similar to $H_v^{1,p}$ but the relaxation is made with general L^p functions, and the asymptotic Lipschitz constant is replaced by upper gradients, or the one given in [75], similar to $BL^{1,p}$ but with a slightly stronger notion of negligibility of set of curves.

We will prove that $H_v^{1,p} \subseteq W^{1,p} \subseteq BL^{1,p}$ and that the following inequality is true for the weak gradients:

$$|\nabla f|_{p,v} \leq |\nabla f|_p \leq |\nabla f|_{p,BL} \qquad \mathfrak{m}\text{-a.e. in } X$$

Then for $p > 1$, using the equivalence $H^{1,p}_v = BL^{1,p}$ and $|\nabla f|_{p,BL} = |\nabla f|_{p,v}$ in [9] will let us conclude; also the coincidence with other definitions can be found in [9]. For $p = 1$ the equivalence is still an open question.

Let us recall briefly the definitions of $H^{1,p}_v$ (in the stronger version given in [4]) and $BL^{1,p}$:

Definition 7.2.1 (Relaxed Sobolev Space) *A function $f \in L^p(X, \mathfrak{m})$ belongs to $H^{1,p}_v(X, \mathsf{d}, \mathfrak{m})$ if and only if there exists a sequence $(f_n) \subset \mathrm{Lip}_0(X, \mathsf{d})$ and a function $g \in L^p(X, \mathfrak{m})$ such that*

$$\lim_{n\to\infty} \|f_n - f\|_p + \|\mathrm{lip}_a(f_n) - g\|_p = 0.$$

The function g with minimal L^p norm that has this property will be denoted with $|\nabla f|_{p,v}$

In order to define the space $BL^{1,p}$ we recall the Definition 2.6.1 of test plans and the subsequent definition of the BL space:

Definition 7.2.2 (Weak Sobolev Space) *A function $f \in L^p(X, \mathfrak{m})$ belongs to $BL^{1,p}(X, \mathsf{d}, \mathfrak{m})$ if there exists a function $g \in L^p(X, \mathfrak{m})$ that is a p-weak upper gradient of f, i.e. it is such that*

$$\left|\int_{\partial\gamma} f\right| \leq \int_\gamma g < \infty \qquad \text{for } p\text{-a.e. } \gamma. \tag{7.2.1}$$

The minimal p-weak upper gradient (in the pointwise sense) will be denoted by $|\nabla f|_{p,BL}$.

7.2.1 $H^{1,p}_v \subseteq W^{1,p}$

Let $f \in H^{1,p}_v$. Then we have a sequence of Lipschitz functions such that $f_n \xrightarrow{p} f$ and $\mathrm{Lip}_a(f_n) \xrightarrow{p} |\nabla f|_{p,v}$. Then by the strong locality property of derivation and the definition of divergence we know that for every $\boldsymbol{b} \in \mathrm{Der}^{q,q}$

$$\left|\int_X f_n \cdot \mathrm{div}\, \boldsymbol{b} \,\mathrm{d}\mathfrak{m}\right| = \left|\int_X \boldsymbol{b}(f_n) \,\mathrm{d}\mathfrak{m}\right| \leq \int_X |\boldsymbol{b}| \cdot \mathrm{lip}_a(f_n) \,\mathrm{d}\mathfrak{m},$$

and so, taking limits, we have that

$$\left|\int_X f \cdot \mathrm{div}\, \boldsymbol{b} \,\mathrm{d}\mathfrak{m}\right| \leq \int_X |\boldsymbol{b}| \cdot |\nabla f|_{p,v} \,\mathrm{d}\mathfrak{m} \qquad \forall \boldsymbol{b} \in \mathrm{Der}^{q,q} \tag{7.2.2}$$

Now we have to construct the linear functional $L_f : \mathrm{Der}^{q,q} \to L^1$. So, fix $\boldsymbol{b} \in \mathrm{Der}^{q,q}$ and let $\mu_{\boldsymbol{b}} = |\boldsymbol{b}| \cdot |\nabla f|_{p,v}\mathfrak{m}$. Notice that $\mu_{\boldsymbol{b}}$ is a finite measure. Now let $R^{\boldsymbol{b}} : \mathrm{Lip}_b(X, \mathsf{d}) \to \mathbb{R}$ be the linear functional defined by

$$R^{\boldsymbol{b}}(h) = -\int_X f \cdot \mathrm{div}(h\boldsymbol{b}) \,\mathrm{d}\mathfrak{m};$$

notice that, thanks to (7.2.2), $|R^{\boldsymbol{b}}(h)| \leq C\|h\|_\infty$, where $C = \mu_{\boldsymbol{b}}(X)$ and so it can be extended to a continuous linear functional on $C_b(X)$; since $|R^{\boldsymbol{b}}(h)| \leq \int_X |h| \,\mathrm{d}\mu_{\boldsymbol{b}}$, we have that $R^{\boldsymbol{b}}(h)$ can be represented as an integral with respect to a signed measure $\mathfrak{m}_{\boldsymbol{b}}$, whose total variation is less then $\mu_{\boldsymbol{b}}$, but since $\mu_{\boldsymbol{b}}$ is absolutely continuous with respect to $\mathfrak{m}$, so it is $\mathfrak{m}_{\boldsymbol{b}}$; if we denote by $L_f(\boldsymbol{b})$ the density of $\mathfrak{m}_{\boldsymbol{b}}$ relative to $\mathfrak{m}$, we have

$$-\int_X f \cdot \mathrm{div}(h\boldsymbol{b}) \,\mathrm{d}\mathfrak{m} = \int h \cdot L_f(\boldsymbol{b}) \,\mathrm{d}\mathfrak{m} \qquad \forall\, h \in \mathrm{Lip}_b(X, \mathsf{d}) \tag{7.2.3}$$

$$|L_f(\boldsymbol{b})| \leq |\boldsymbol{b}| \cdot |\nabla f|_{p,v} \qquad \mathfrak{m}\text{-almost everywhere} \tag{7.2.4}$$

Now we have to check the Lip_b-linearity, but this is easy since for every $h \in \mathrm{Lip}_b$ by definition we have $R^{h\boldsymbol{b}}(h_1) = R^{\boldsymbol{b}}(h \cdot h_1)$, for every bounded Lipschitz function h_1; in particular

$$\int_X h_1 \cdot L_f(h\boldsymbol{b})\,\mathrm{d}\ mm = \int_X h_1 \cdot hL_f(\boldsymbol{b})\,\mathrm{d}\mathfrak{m} \qquad \forall h_1 \in \mathrm{Lip}_b(X),$$

and so $L_f(h\boldsymbol{b}) = hL_f(\boldsymbol{b})$.

7.2.2 $W^{1,p} \subseteq BL^{1,p}$

The crucial observation is that every q-plan induce a derivation:

Proposition 7.2.3 *Let $\boldsymbol{\pi}$ be a q-plan. For every function $f \in \mathrm{Lip}_b(X, \mathsf{d})$ let us consider $\boldsymbol{b}_{\boldsymbol{\pi}}(f)$, the function such that:*

$$\int_X g \cdot \boldsymbol{b}_{\boldsymbol{\pi}}(f)\,\mathrm{d}\mathfrak{m} = \int_{AC} \int_0^1 g(\gamma_t) \frac{d(f \circ \gamma)}{ds}(t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi}(\gamma) \qquad \forall g \in L^p. \tag{7.2.5}$$

Then we have that $\boldsymbol{b}_{\boldsymbol{\pi}} \in \mathrm{Der}^{q,q}$ and moreover

$$\int_X g \cdot |\boldsymbol{b}_{\boldsymbol{\pi}}|\,\mathrm{d}\mathfrak{m} \le \iint_\gamma g\,\mathrm{d}s\,\mathrm{d}\boldsymbol{\pi}(\gamma) \qquad \forall g \in L^p, g \ge 0; \tag{7.2.6}$$

$$\int_X f \cdot \mathrm{div}(\boldsymbol{b}_{\boldsymbol{\pi}})\,\mathrm{d}\mathfrak{m} = \int_{AC} (f(\gamma_1) - f(\gamma_2))\,\mathrm{d}\boldsymbol{\pi}(\gamma) \qquad \forall f \in L^p. \tag{7.2.7}$$

Proof. We first fix $f \in \mathrm{Lip}_b(X, \mathsf{d})$ and notice that the right hand side in (7.2.5) is well defined thanks to Rademacher theorem. Then the Leibniz rule is easy to check thanks to its validity in the right hand side of (7.2.5). In order to find a good candidate for $|\boldsymbol{b}_{\boldsymbol{\pi}}|$, we estimate $\frac{d(f\circ\gamma)}{ds} \le \mathrm{lip}_a(f)(\gamma_t)|\dot\gamma_t|$ and so, for every nonnegative $g \in L^p$ we have

$$\int_0^1 g(\gamma_t) \frac{d(f \circ \gamma)}{ds}(t)\,\mathrm{d}t \le \int_0^1 g(\gamma_t)\mathrm{lip}_a(f)(\gamma_t)|\dot\gamma_t|\,\mathrm{d}t;$$

integrating with respect to $\boldsymbol{\pi}$ and using Fubini theorem we get

$$\int_X g \cdot \boldsymbol{b}_{\boldsymbol{\pi}}(f)\,\mathrm{d}\mathfrak{m} \le \int_X g \cdot \mathrm{lip}_a(f)\,\mathrm{d}\mu_{\boldsymbol{\pi}}, \tag{7.2.8}$$

where $\mu_{\boldsymbol{\pi}} = \int_0^1 (\mathrm{e}_t)_\sharp(\|\dot\gamma_t|\boldsymbol{\pi})\,\mathrm{d}t$ is the *barycenter* of $\boldsymbol{\pi}$, and it is such that

$$\int_X g\,\mathrm{d}\mu_{\boldsymbol{\pi}} = \iint_\gamma g\,\mathrm{d}s\,\mathrm{d}\boldsymbol{\pi}. \tag{7.2.9}$$

In particular we can use Hölder's inequality to estimate the behavior of $\mu_{\boldsymbol{\pi}}$:

$$\begin{aligned}
\int_X g\,\mathrm{d}\mu_{\boldsymbol{\pi}} &= \int_{AC} \int_0^1 g(\gamma_t)|\dot\gamma_t|\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi} \\
&\le \left(\iint |g(\gamma_t)|^p\,\mathrm{d}t\mathrm{d}\boldsymbol{\pi}\right)^{1/p} \left(\iint |\dot\gamma_t|^q\,\mathrm{d}t\mathrm{d}\boldsymbol{\pi}\right)^{1/q} \\
&\le C(\boldsymbol{\pi})^{1/p} \cdot \|g\|_{L^p(\mathfrak{m})} \cdot \|E_q(\gamma)\|_{L^q(\boldsymbol{\pi})},
\end{aligned}$$

and so, by duality argument, we obtain that $\mu_{\boldsymbol{\pi}} = h\mathfrak{m}$ with $h \in L^q(X, \mathfrak{m})$; using this representation in (7.2.8) we obtain

$$\int_X g \cdot \boldsymbol{b}_{\boldsymbol{\pi}}(f)\, d\mathfrak{m} \leq \int_X g \cdot \mathrm{lip}_a(f) h\, d\mathfrak{m} \qquad \forall\, g \in L^q,\ g \geq 0.$$

So we deduce that $|\boldsymbol{b}_{\boldsymbol{\pi}}| \leq h$ and in particular $\boldsymbol{b}_{\boldsymbol{\pi}} \in L^q$ and (7.2.6) is true thanks to (7.2.9).

It remains to prove the last equality: by definition of divergence we have, for $f \in \mathrm{Lip}_0(X, \mathsf{d})$

$$\int f \cdot \operatorname{div} \boldsymbol{b}_{\boldsymbol{\pi}}\, d\mathfrak{m} = \int_{AC} \int_0^1 \frac{d(f \circ \gamma)}{ds}(t)\, dt\, d\boldsymbol{\pi}(\gamma) = \int (f(\gamma_1) - f(\gamma_0))\, d\boldsymbol{\pi}, \tag{7.2.10}$$

thanks to the fact that the fundamental theorem of calculus holds for Lipschitz functions. By definition of q-plan we have also that $(\mathrm{e}_t)_\sharp \boldsymbol{\pi} = f_t \mathfrak{m}$ where $f_t \leq C(\boldsymbol{\pi})$ for every $t \in [0, 1]$; since $\boldsymbol{\pi}$ is a probability measure we have $\int f_t\, d\mathfrak{m} = 1$ and so $f_t \in L^1 \cap L^\infty$ and in particular $f_t \in L^q$ and so $\operatorname{div} \boldsymbol{b}_{\boldsymbol{\pi}} = (f_1 - f_0) \in L^q$. This enables us to extend (7.2.10) to $f \in L^p$ and so we proved also (7.2.7). □

Lemma 7.2.4 *Let $f \in W^{1,p}(X, \mathsf{d}, \mathfrak{m})$. Then $|\nabla f|_w$ is a p-weak upper gradient for f.*

Proof. By Proposition 7.2.3 we know that to every q-plan $\boldsymbol{\pi}$ we can associate a derivation $\boldsymbol{b}_{\boldsymbol{\pi}} \in \mathrm{Der}^{q,q}$; we use this derivation in the definition of $W^{1,p}$ and, using also Theorem 7.1.6, we obtain

$$-\int_X f \cdot \operatorname{div} \boldsymbol{b}_{\boldsymbol{\pi}}\, d\mathfrak{m} \leq \int |\nabla f|_w \cdot |\boldsymbol{b}_{\boldsymbol{\pi}}|\, d\mathfrak{m};$$

Now, using (7.2.6) and (7.2.7), we obtain precisely

$$\int_{AC} (f(\gamma_0) - f(\gamma_1)) d\boldsymbol{\pi} \leq \int_{AC} \int_\gamma |\nabla f|_w\, ds\, d\boldsymbol{\pi}. \qquad \forall \boldsymbol{\pi}\ q\text{-plan} \tag{7.2.11}$$

We can "localize" this inequality using the fact that for every Borel set $A \subseteq C([0, 1]; X)$ such that $\boldsymbol{\pi}(A) \neq 0$, we have that $\boldsymbol{\pi}_A = \frac{1}{\boldsymbol{\pi}(A)} \boldsymbol{\pi}|_A$ is still a q-plan and so we can infer that

$$\int_A (f(\gamma_0) - f(\gamma_1)) d\boldsymbol{\pi} \leq \int_A \int_\gamma |\nabla f|_w\, ds\, d\boldsymbol{\pi}. \qquad \forall A \subset C([0, 1]; X), \tag{7.2.12}$$

and so $f(\gamma_0) - f(\gamma_1) \leq \int_\gamma |\nabla f|_w$ for $\boldsymbol{\pi}$-almost every curve. Applying the same conclusion to $-f$ we get that the upper gradient property is true for $\boldsymbol{\pi}$-almost every curve. Since $\boldsymbol{\pi}$ was an arbitrary q-plan, by definition we have

$$|f(\gamma_0) - f(\gamma_1)| \leq \int_\gamma |\nabla f|_w\, ds \qquad \text{for } p\text{-almost every curve } \gamma$$

and so $|\nabla f|_w$ is a p-weak upper gradient. □

Theorem 7.2.5 *[H=W=BL] Let $(X, \mathsf{d}, \mathfrak{m})$ be a separable complete metric space, endowed with a measure $\mathfrak{m}$ that is finite on bounded sets. Then we have*

$$H_v^{1,p}(X, \mathsf{d}, \mathfrak{m}) = W^{1,p}(X, \mathsf{d}, \mathfrak{m}) = BL^{1,p}(X, \mathsf{d}, \mathfrak{m});$$

moreover we have $|\nabla f|_{p,v} = |\nabla f|_{p,w} = |\nabla f|_{p,BL}$ for every $f \in W^{1,p}$.

Proof. It is sufficient to use Lemma 7.2.4 and the results of Section 7.2.1 to obtain $H_v^{1,p} \subseteq W^{1,p} \subseteq BL^{1,p}$ and $|\nabla f|_v \geq |\nabla f|_w \geq |\nabla f|_{BL}$. Then we use the equivalence theorem in [9]: given $f \in BL^{1,p}$, we have that $f \in H_v^{1,p}$ and $|\nabla f|_v \leq |\nabla f|_{BL}$. This let us conclude. □

7.3 BV space via derivations

From now on we will denote $\int_X \mathrm{d}\mu = \mu(X)$ whenever $\mu \in \mathcal{M}(X)$.

Definition 7.3.1 *Let $f \in L^1(X, \mathsf{d}, \mathfrak{m})$; we say $f \in BV(X, \mathsf{d}, \mathfrak{m})$ if there exists a continuous linear map $L_f : \mathrm{Der}_b \to \mathcal{M}(X)$ satisfying*

$$\int_X \mathrm{d}L_f(\boldsymbol{b}) = -\int_X f \operatorname{div} \boldsymbol{b}\, d\mathfrak{m} \qquad \forall\ \boldsymbol{b} \in \mathrm{Der}_b, \tag{7.3.1}$$

such that $L_f(h\boldsymbol{b}) = hL_f(\boldsymbol{b})$ for every $h \in \mathrm{Lip}_b(X, \mathsf{d}), \boldsymbol{b} \in \mathrm{Der}_b$.

As in the $W^{1,p}$ case, we prove that $L_f(\boldsymbol{b})$ is uniquely defined whenever $f \in BV$ and $\boldsymbol{b} \in \mathrm{Der}_b$:

Remark 7.3.2 (Well posedness of L_f) *Let us fix $\boldsymbol{b} \in \mathrm{Der}_b$, $f \in BV$; let L_f and $\tilde{L}_f$ be two different linear maps given in the definition on BV. Let $h \in \mathrm{Lip}_b(X, \mathsf{d})$: using Lemma 7.1.2 we have $h\boldsymbol{b} \in \mathrm{Der}_b$ and so we can use (7.3.1) and the Lip_b-linearity to get*

$$\int_X h\, \mathrm{d}L_f(\boldsymbol{b}) = \int_X \mathrm{d}L_f(h\boldsymbol{b}) = -\int_X f \operatorname{div}(h\boldsymbol{b})\, \mathrm{d}\mathfrak{m},$$

and the same is true for $\tilde{L}_f$. In particular $\int_X h\, \mathrm{d}L_f(\boldsymbol{b}) = \int_X h\, \mathrm{d}\tilde{L}_f(\boldsymbol{b})$, and thanks to the arbitrariness of $h \in \mathrm{Lip}_b(X, \mathsf{d})$ we conclude that $L_f(\boldsymbol{b}) = \tilde{L}_f(\boldsymbol{b})$. We will call this common value $Df(\boldsymbol{b})$.

Now we can give the definition of total variation:

Theorem 7.3.3 *Let $f \in BV(X, \mathsf{d}, \mathfrak{m})$; then there exists a finite measure $\nu \in \mathcal{M}_+(X)$ such that, for every Borel set $A \subseteq X$,*

$$\int_A \mathrm{d}Df(\boldsymbol{b}) \leq \int_A |\boldsymbol{b}|^*\, \mathrm{d}\nu \qquad \forall \boldsymbol{b} \in \mathrm{Der}_b, \tag{7.3.2}$$

where g^* denotes the upper semicontinuous envelope of g. The least measure that realizes this inequality is denoted with $|Df|$, the weak total variation of f. Moreover

$$|Df|(X) = \sup\{|Df(\boldsymbol{b})(X)|\ :\ |\boldsymbol{b}| \leq 1, \boldsymbol{b} \in \mathrm{Der}_b\}. \tag{7.3.3}$$

Proof. We argue similarly to Theorem 7.1.6: by hypothesis we have that $f \in BV$ and so there exists a Lip_b-linear map $Df : \mathrm{Der}_b \to \mathcal{M}(X)$ such that $Df(\boldsymbol{b})(X) \leq C\|\boldsymbol{b}\|_{L^\infty}$, where we can take $C = \sup\{|Df(\boldsymbol{b})(X)|\ :\ |\boldsymbol{b}| \leq 1, \boldsymbol{b} \in \mathrm{Der}_b\}$. Note that if $|\boldsymbol{b}| \leq h$ where $h \in C_b(X)$ then we have that

$$\int_K \mathrm{d}Df(\boldsymbol{b}) \leq C \sup_{x \in K} h(x) \qquad \forall K \subseteq X \text{ compact}; \tag{7.3.4}$$

in fact, denoting with $\rho_n = \min\{1 - n\mathsf{d}(x, K)\}$, we have that $\rho_n \to \chi_K$ pointwise and $0 \leq \rho_n \leq 1$ so, by dominated convergence theorem,

$$\int_K \mathrm{d}Df(\boldsymbol{b}) = \lim_{n\to\infty} \int_X \rho_n\, \mathrm{d}Df(\boldsymbol{b}) \leq C \lim_{n\to\infty} \|\rho_n \boldsymbol{b}\|_\infty \leq C \lim_n \sup_{x\in X} \rho_n(x)h(x) = C \sup_{x \in K} h(x),$$

where the last equality holds thanks to the compactness of K. Now, for every compact set $K \subseteq X$ and consider two functionals in the Banach space $Y = C_b(K)$:

$$\Psi_2(h) = C\|h\|_\infty \tag{7.3.5}$$

$$\Psi_1(h) = \sup\left\{\int_K \mathrm{d}Df(\boldsymbol{b}) : \boldsymbol{b} \in \mathrm{Der}_b,\ \ \exists\, \tilde h \in C_b(X) \text{ such that } |\boldsymbol{b}| \le \tilde h,\ \tilde h|_K \le h\right\} \tag{7.3.6}$$

where the supremum of the empty set is meant to be $-\infty$. Equation (7.3.4) guarantees that

$$\Psi_1(h) \le \Psi_2(h) \qquad \forall h \in Y. \tag{7.3.7}$$

Moreover, as before, Ψ_2 is convex and continuous while Ψ_1 is concave; by Hahn-Banach theorem we can find a continuous linear functional L on $C_b(K)$ such that

$$\Psi_1(h) \le L(h) \le \Psi_2(h).$$

In particular there exists a measure μ_K such that $L(h) = \int_K h\,\mathrm{d}\mu_K$ and, thanks to (7.3.5), we have $\mu_K(K) \le C$. Moreover, thanks to (7.3.6) we have that if $h \in C_b(X)$ is such that $|\boldsymbol{b}| \le h$ for some $\boldsymbol{b} \in \mathrm{Der}_b$ then

$$\int_K \mathrm{d}Df(\boldsymbol{b}) \le \int_K h\,\mathrm{d}\mu_K;$$

since for every $k \in C_b(X)$, we have $|k\boldsymbol{b}| \le |k|h$ we obtain also

$$\int_K k\,\mathrm{d}Df(\boldsymbol{b}) \le \int_K |k|h\,\mathrm{d}\mu_K.$$

In particular, optimizing in k we obtain also that $|Df(\boldsymbol{b})|$, the total variation of $Df(\boldsymbol{b})$, restricted to K, is less then or equal to $h\mu_K$. This implies that the following set is nonempty:

$$A_K = \{\nu \in \mathcal{M}_+(K)\ :\ |Df(\boldsymbol{b})||_K \le h\nu \text{ whenever } \boldsymbol{b} \in \mathrm{Der}_b, h \in C_b(X) \text{ s.t. } |\boldsymbol{b}| \le h\}\,.$$

Clearly this set is convex, weakly-$*$ closed and a lattice, in particular there exists the minimum, that we call ν_K. We can drop the dependence on K since it is easy to see that if $A \subset K_1 \cap K_2$ then $\nu_{K_1}(A) = \nu_{K_2}(A)$; suppose on the contrary that $\nu_{K_1}(A) > \nu_{K_2}(A)$. Then we can consider the measure $\tilde\nu(B) = \nu_{K_1}(B \setminus A) + \nu_{K_2}(B \cap A)$ that would be a strictly better competitor than μ_{K_1} in A_{K_1}.

Now we can extend ν to a measure on the whole space

$$\nu(B) = \sup_{K \subseteq B} \nu(K) \qquad \forall B \subseteq X \text{ Borel};$$

this is easily seen to be a measure, that is also finite since $\nu(K) \le \mu_K(K) \le C$ for all K compact and in particular we get $\nu(X) \le C$. Thanks to the finiteness of $|Df(\boldsymbol{b})|$ and ν, using that $\nu|_K \in A_K$, we find that

$$|Df(\boldsymbol{b})| \le h\nu \text{ whenever } \boldsymbol{b} \in \mathrm{Der}_b, h \in C_b(X) \text{ s.t. } |\boldsymbol{b}| \le h,$$

in particular, integrating in A we get

$$\int_A \mathrm{d}Df(\boldsymbol{b}) \le \int_A h\,\mathrm{d}\nu,$$

and taking the infimum in h we obtain (7.3.2), recalling that if $g \in L^\infty$ then

$$g^*(x) = \inf\{h(x)\ :\ h \in C_b(X),\ \ h \ge g\ \mathfrak{m}\text{-a.e.}\}.$$

For the last assertion in the theorem we already proved $C \ge \nu(X)$, while the other inequality is trivial taking $A = X$ in (7.3.2). $\square$

Theorem 7.3.4 (Representation formula for $|Df|$) *Let $f \in BV$. Then the classical representation formula holds true: for every open set A*

$$|Df|(A) = \sup\left\{\int_A f \cdot \mathrm{div}(\boldsymbol{b})\,\mathrm{d}\mathfrak{m} \;:\; \boldsymbol{b} \in \mathrm{Der}_b,\ |\boldsymbol{b}| \le 1,\ \mathrm{supp}(\boldsymbol{b}) \Subset A\right\}. \tag{7.3.8}$$

Proof. Let us consider two open sets A_1, A_2 and a closed set C such that $A_1 \Subset C \Subset A_2$. We will consider $(C, \mathsf{d}, \mathfrak{m})$ as a separable metric measure space, and relate the definitions of bounded variation in X and C. Let us consider a function $f \in BV(X, \mathsf{d}, \mathfrak{m})$; it is clear that $f \in BV(C, \mathsf{d}, \mathfrak{m})$ since $\mathrm{Der}_b(C) \subset \mathrm{Der}_b(X)$ (it is sufficient to set $\boldsymbol{b}_X(f) = \boldsymbol{b}_C(f|_C)$), and consequently $|Df|_X \ge |Df|_C$ by (7.3.2).

Moreover it is true that $|Df|_X(A_1) = |Df|_C(A_1)$. This is true because there exists a Lipschitz function $0 \le \chi \le 1$ such that $\chi = 0$ in $X \setminus C$ and $\chi = 1$ on a neighborhood of A_1; then we have that if $\boldsymbol{b} \in \mathrm{Der}_b(X)$ implies that $\chi\boldsymbol{b} \in \mathrm{Der}_b(C)$ and so in (7.3.2) we can imagine that $\boldsymbol{b} \in \mathrm{Der}_b(C)$ whenever $A \subset A_1$; but then we get that the measure ν defined as

$$\nu(B) = |Df|_X(B \setminus A_1) + |Df|_C(B \cap A_1)$$

is a good candidate in (7.3.2) and so, by the minimality of $|Df|_X$ we get $|Df|_C(A_1) = |Df|_X(A_1)$.

Now, denoting by $\mu(A)$ the set function defined in the left hand side of (7.3.8), it is obvious that $\mu(A_2) \le |Df|(A_2)$. But it is also obvious that $\mu(A_2) \ge |Df|_C(C) \ge |Df|_C(A_1) = |Df|_X(A_1)$. Letting $A_1 \uparrow A_2$ we get the desired inequality. □

7.3.1 Equivalence of BV spaces

We just sketch the equivalence with the other definitions given in literature: in particular we refer to [5] (or Chapter 4), where the authors consider the spaces BV_* and $w-BV$ and show their equivalence. As we did for $W^{1,p}$ we show $BV_* \subseteq BV \subseteq w-BV$.

Lemma 7.3.5 *Let $f \in BV_*(X, \mathsf{d}, \mathfrak{m})$. Then we have $f \in BV(X, \mathsf{d}, \mathfrak{m})$ and $|Df| \le |Df|_*$ as measures.*

Proof. By hypothesis, we know that there is a sequence $(f_n) \subset \mathrm{Lip}_0(X, \mathsf{d})$ such that $\mathrm{lip}_a(f_n) \rightharpoonup |Df|_*$ in duality with $C_b(X)$; in particular, for every $\boldsymbol{b} \in \mathrm{Der}_b$ we have

$$\left|\int_X f_n \cdot \mathrm{div}\,\boldsymbol{b}\,\mathrm{d}\mathfrak{m}\right| = \left|\int_X \boldsymbol{b}(f_n)\,\mathrm{d}\mathfrak{m}\right| \le \int_X |\boldsymbol{b}| \cdot \mathrm{lip}_a(f_n)\,\mathrm{d}\mathfrak{m}$$

taking limits and recalling that whenever $\nu_n \rightharpoonup \nu$ and $g \ge 0$, we have $\liminf_{n\to\infty} \int_X g\,\mathrm{d}\mu_n \le \int_X g^*\,\mathrm{d}\mu$, we have that

$$\left|\int_X f \cdot \mathrm{div}\,\boldsymbol{b}\right| \le \int_X |\boldsymbol{b}|^*\,\mathrm{d}|Df|_* \qquad \forall \boldsymbol{b} \in \mathrm{Der}_b.$$

Now this inequality would guarantee that $|Df| \le |Df|_*$ once we construct the linear functional $L_f : \mathrm{Der}_b \to \mathcal{M}(X)$ In order to find $L_f(\boldsymbol{b})$ we proceed exactly as in Section 7.2.1, and so we omit the construction. □

Lemma 7.3.6 *Let $f \in BV(X, \mathsf{d}, \mathfrak{m})$. Then we have $f \in w - BV(X, \mathsf{d}, \mathfrak{m})$ and $|Df|_w(X) \leq |Df|(X)$.*

Proof. As for the second inclusion it is sufficient to recall Proposition 7.2.3: we know that for every ∞-plan $\boldsymbol{\pi}$ we can associate a derivation $\boldsymbol{b_\pi} \in \mathrm{Der}_b$; we use this derivation in the definition of BV and, using also Theorem 7.3.3, we obtain

$$-\int_X f \cdot \operatorname{div} \boldsymbol{b_\pi} \, \mathrm{d}\mathfrak{m} \leq \int_X |\boldsymbol{b_\pi}|^* \, \mathrm{d}|Df|.$$

Now, using (7.2.6) and (7.2.7), we obtain

$$\int_{AC} (f(\gamma_0) - f(\gamma_1)) \mathrm{d}\boldsymbol{\pi} \leq C(\boldsymbol{\pi}) \cdot |Df|(X) \|\mathrm{Lip}(\gamma)\|_{L^\infty(\boldsymbol{\pi})} \qquad \forall \boldsymbol{\pi} \ \infty\text{-plan}. \tag{7.3.9}$$

Now we can use Remark 7.2 in [5] to conclude that $f \in w - BV$ and $|Df|_w(X) \leq |Df|(X)$ □

Using this two lemmas in conjunction with the equivalence result in [5] we can conlcude.

Theorem 7.3.7 *Let $(X, \mathsf{d}, \mathfrak{m})$ be a complete and separable metric space, such that $\mathfrak{m}$ is finite on bounded sets; then $BV(X, \mathsf{d}, \mathfrak{m}) = BV_*(X, \mathsf{d}, \mathfrak{m}) = w - BV(X, \mathsf{d}, \mathfrak{m})$. Moreover $|Df| = |Df|_* = |Df|_w$ for every function $f \in BV$.*

Proof. From Lemma 7.3.5 and 7.3.6 we know that $BV_* \subseteq BV \subseteq w - BV$ and moreover $|Df| \leq |Df|_*$ and $|Df|(X) \geq |Df|_w(X)$. Thanks to the equivalence theorem in [5] we get $BV = BV_* = w - BV$ and $|Df|_w = |Df|_*$, in particular $|Df|_w(X) = |Df|_*(X) \geq |Df|(X)$, and so $|Df|(X) = |Df|_*(X) = |Df|_w(X)$. This equality, along with $|Df| \leq |Df|_*$ let us conclude that the three definitions of total variation coincide. □

7.4 Sobolev Bundle

Here we suppose that X is a Banach space with separable dual, in which we can use our new definition to give a precise value to $|\nabla f|_p$ in the case of $f \in C^1_{\mathrm{loc}}(X)$ and $p > 1$. In particular we will define the Sobolev bundle S_p, i.e. a map that at each point $x \in X$ assigns a closed subspace of the tangent space at x, identified with X. We will prove that $|\nabla f|_p = |df|_{S_p}|$ $\mathfrak{m}$-a.e. A similar result has been already proved in [22] for finite dimensional spaces; the author use a bidual argument that can be adapted to reflexive Banach spaces. Our approach is different and more general since we can recover the result also for non-reflexive Banach spaces.

We first state a characterization of derivations in $D(\mathrm{div})$ as vector fields, in the case X is a Banach space:

Lemma 7.4.1 *Let $\boldsymbol{b} \in D(\mathrm{div})$. Then there exists and $\mathfrak{m}$-measurable map $v_{\boldsymbol{b}} : X \to X$ such that $\boldsymbol{b}(f) = df(v_{\boldsymbol{b}}) = \frac{\partial f}{\partial v_{\boldsymbol{b}}}$, for every $f \in C^1(X, \mathsf{d})$; moreover we have $|\boldsymbol{b}|(x) = |v_{\boldsymbol{b}}(x)|$ for $\mathfrak{m}$-a.e. $x \in X$.*

Proof. We will not enter in too many technical details; since this is a local statement we can assume that $\boldsymbol{b} \in \mathrm{Der}^{1,1}$. Then we can apply the results in [70] to $\boldsymbol{b}\mathfrak{m}$, that is easily seen to be a normal current, in order to find an integral representation of $\boldsymbol{b}$ through derivations along curves: this is in some sense dual to Proposition 7.2.3; in particular we have

$$\int g \cdot \boldsymbol{b}(f)\,\mathrm{d}\mathfrak{m} = \iint_0^1 g(\gamma_t) \cdot (f \circ \gamma)'(t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi} \qquad \forall g \in L^\infty(X, \mathfrak{m})$$

$$\int g|\boldsymbol{b}|\,\mathrm{d}\mathfrak{m} = \iint_0^1 g(\gamma_t)|\dot{\gamma}_t|\,\mathrm{d}t\,\mathrm{d}\mathfrak{m} \qquad \forall g \in L^\infty(X, \mathfrak{m}) \tag{7.4.1}$$

Then we know that for $\gamma \in \mathrm{AC}([0,1]; X)$, there exists the tangent vector for almost every time $t \in [0,1]$. This leads to the definition of the vector field $v_{\boldsymbol{b}}$ by duality:

$$\int \langle \rho, v_{\boldsymbol{b}} \rangle\,\mathrm{d}\mathfrak{m} = \iint_0^1 \langle \rho, \dot{\gamma}_t \rangle\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi} \qquad \forall \rho \in L^\infty(X; X^*). \tag{7.4.2}$$

Combining this with (7.4.1) we obtain $|v_{\boldsymbol{b}}| \leq |\boldsymbol{b}|$. For every $f \in C^1$ it is obvious that $df(v_{\boldsymbol{b}}) = \boldsymbol{b}(f)$ (using $\rho = g \cdot df$ in (7.4.2)) and so we have

$$|\boldsymbol{b}(f)| \leq |v_{\boldsymbol{b}}| \cdot \mathrm{lip}_a f. \tag{7.4.3}$$

If we suppose X finite dimensional then we can obtain that (7.4.3) holds also for $\mathfrak{m}$-a.e. for Lipschitz functions f, by approximating f with convolutions. Let us now assume X is infinite dimensional; we know that $\mathfrak{m}$ is supported on a σ-compact set S such that there exist linear projections $\pi_n : X \to \mathbb{R}^n \subset X$ such that $\pi_n x \to x$ for every $x \in S$. Let us consider $f_n(x) = f(\pi_n(x))$; using convolutions in $\pi_n(X)$, as in the finite dimensional case we find that (7.4.3) holds f_n. In particular we have

$$|\boldsymbol{b}(f_n)(x)| \leq |v_{\boldsymbol{b}}(x)| \cdot \mathrm{lip}_a(f, \pi_n x) \leq |v_{\boldsymbol{b}}(x)| \cdot \mathrm{Lip}(f, B(x, r_m(x))) \qquad \forall m \leq n, \tag{7.4.4}$$

where we may take $r_n(x) = 2\|x - \pi_n x\|$, that is decreasing in n. Since we have $f_n \to f$ pointwise in S, thanks to the integration by parts formula we have also $\boldsymbol{b}(f_n) \rightharpoonup \boldsymbol{b}(f)$ in L^1 and in particular in (7.4.4) we can let $n \to \infty$ and then $m \to \infty$, to obtain that (7.4.3) is true $\mathfrak{m}$-a.e (using that $r_m(x) \to 0$ for every $x \in S$). This proves that $|\boldsymbol{b}| \leq |v_{\boldsymbol{b}}|$ and so we have $|\boldsymbol{b}| = |v_{\boldsymbol{b}}|$. □

In the sequel, we will often identify $\boldsymbol{b} \in D(\mathrm{div})$ with the vector field $v_{\boldsymbol{b}}$ given by Proposition 7.4.1, through the equality $\boldsymbol{b} = v_{\boldsymbol{b}}$. Let us denote by $\mathcal{B}(X)$ the set of Borel maps from X to the set of closed subspaces of X, denoted by $Cl(X)$.

Definition 7.4.2 *S_p is the p-Sobolev bundle if*

(i) For every $\boldsymbol{b} \in \mathrm{Der}^{q,q}$ with $\boldsymbol{b} = v_{\boldsymbol{b}}$ have $v_{\boldsymbol{b}}(x) \in S_p(x)$ for $\mathfrak{m}$-a.e. $x \in X$

(ii) For every S' that satisfies (i) we have that $S_p(x) \subseteq S'(x)$ for $\mathfrak{m}$-a.e. $x \in X$.

In order to prove existency it is sufficient to find a map $F : Cl(X) \to [0,1]$, stictly increasing by inclusion, namely if $Y_1 \subsetneq Y_2$ then $F(Y_1) < F(Y_2)$. Then we minimize the quantity

$$\int_X F(S(x))\,\mathrm{d}\mathfrak{m} \tag{7.4.5}$$

among all $S \in \mathcal{B}(X)$ that satisfy (i); here we point out that this set is nonempty, considering the constant map $S(x) = X$. Let us say that a sequence S_n realizes the infimum, and now let us consider

$$S_p(x) = \bigcap_{n \in \mathbb{N}} S_n(x).$$

It is clear that S_p still satisfies (i), and of course S_p minimizes (7.4.5); now suppose that S_p doesn't satisfy (ii), and so there exists S' such that S' satisfy (i) and $\mathfrak{m}\{S_p(x) \nsubseteq S'(x)\} > 0$; define $S'' = S' \cap S_p$. We still have that S'' satisfies (i) and moreover $\mathfrak{m}\{S''(x) \subsetneq S_p(x)\} = \mathfrak{m}\{S_p(x) \nsubseteq S'(x)\} > 0$ and so, thanks to the strict monotonicity, we have that

$$\int_H F(S''(x))\, d\mathfrak{m} < \int_H F(S_p(x))\, d\mathfrak{m},$$

against the minimality of S_p.

The following lemma shows that there exists a map with these properties, when X^* is separable.

Lemma 7.4.3 *Let $\{e_i\}_{i \in \mathbb{N}}$ be a dense set in $B_{X^*}(0,1)$ (with the strong topology). Let $F : Cl(X) \to [0,1]$ be defined in this way:*

$$F(Y) = \sum_{i=1}^{\infty} 2^{-i} \sup\{\langle e_i, y\rangle \ : y \in Y, \|y\| \leq 1\}.$$

Then F is strictly increasing, namely if we consider two closed subsets $Y \subsetneq Y'$ then $F(Y) < F(Y')$.

Proof. The map F is increasing and its image is clearly contained in $[0,1]$. In order to prove the strict monotonicity we consider two closed subspaces $Y \subsetneq Y'$ and a point $y' \in Y' \setminus Y$. Since Y' is closed and convex and $\{y'\}$ is compact, applying Hahn-Banach theorem we know that there exists a linear functional l that separates y' and Y, in particular there exists $r \in \mathbb{R}$ such that

$$l(y) < r \leq l(y') \qquad \forall y \in Y;$$

In particular we can take $l(y') = r$ and, since Y is a vector space, we have $l|_Y = 0$. We know that there exists a sequence $e_{i_k} \to l$ strongly; it is clear that

$$\sup\{\langle e_{i_k}, y\rangle \ : y \in Y, \|y\| \leq 1\} \to \sup\{\langle l, y\rangle \ : y \in Y, \|y\| \leq 1\} = 0,$$

since the functions $e_{i_k} : B(0,1) \cap Y \to \mathbb{R}$ are converging uniformly to l. Moreover

$$e_{i_k}(y') \to e_{i_k}(y') = r,$$

and so we can find k_0 such that

$$\sup\{\langle e_{i_{k_0}}, y\rangle \ : y \in Y, \|y\| \leq 1\} \leq \frac{r}{3} < \frac{2r}{3} \leq \langle e_{i_{k_0}}, y'\rangle \leq \sup\{\langle e_{i_{k_0}}, y\rangle \ : y \in Y', \|y\| \leq 1\},$$

and so we get $F(Y') > F(Y)$ since all the other terms in the sum that defines F are increasing. □

Definition 7.4.4 *Let $\Upsilon = \{\boldsymbol{b}_i\}_{i\in\mathbb{N}}$ be a countable set of admissible derivations, where $\boldsymbol{b}_i = v_i$ with v_i a Borel vector field defined everywhere; we define the bundle generated by Υ*

$$S_\Upsilon(x) = \overline{\mathrm{span}\{v_i(x)\}_{i\in\mathbb{N}}}.$$

This definition is well posed in the sense that choosing different Borel representative $\tilde{v}_i$ we have that $\tilde{v}_i(x) = v_i(x)$ for $\mathfrak{m}$-a.e. x and so also $\tilde{S}_\Upsilon(x) = S_\Upsilon(x)$ for $\mathfrak{m}$-almost every $x \in H$, and this is sufficient for our pourposes. The next proposition assures that S_p is countably generated in the sense of 7.4.4.

Proposition 7.4.5 *There exists a countable set Υ such that $S_\Upsilon = S_p$.*

Proof. First we note that for every countable family Υ we have $S_\Upsilon \subset S_p$ $\mathfrak{m}$-almost everywhere, so it is sufficient to prove the converse inequality for a certain family. Arguing as we did before to prove existence of S_p we try to maximize

$$\int_H F(S_\Upsilon)\,\mathrm{d}\mathfrak{m} \tag{7.4.6}$$

among all countable admissible families Υ. Taking Υ_n a maximizing sequence we consider $\Upsilon = \cup_{n\in\mathbb{N}}\Upsilon_n$ that is still a countable family of admissible currents that clearly maximizes (7.4.6). Now we want to prove that $S_p \subseteq S_\Upsilon$; suppose on the contrary that this isn't true. This means that S_Υ does not satisfy (i) and so there exists $\boldsymbol{b} \in \mathrm{Der}^{q,q}$, with $\boldsymbol{b} = v_{\boldsymbol{b}}$, such that $v_{\boldsymbol{b}}(x) \notin S_\Upsilon(x)$ in a set of positive measure; but then the set $\Upsilon_1 = \Upsilon \cup \{\boldsymbol{b}\}$ has a greater value in (7.4.6) than Υ, but this goes against the maximality of Υ. □

Now we want to prove that the Sobolev bundle is exactely the set of direction in which we can't neglect the behaviour of the function when we relax. We first need some analytical tools.

Proposition 7.4.6 *Let $\boldsymbol{b} \in \mathrm{Der}^{q,q}$, with $\boldsymbol{b} = v_{\boldsymbol{b}}$. Then for every C^1 function f we have*

$$|\boldsymbol{b}(f)| \le |\nabla f|_w |v_{\boldsymbol{b}}| \qquad \mathfrak{m}\text{-almost everywhere.}$$

Proof. This follows by Theorem 7.1.6 and the equality $|\boldsymbol{b}| = |v_{\boldsymbol{b}}|$ in Lemma 7.4.1. □

Lemma 7.4.7 *Let Υ be a countable family of admissible currents. Let us consider a subset $A \subset \mathbb{R}^n$ such that $\mu(A) > 0$ and a Borel section v of the bundle S_Υ, defined on A, and a treshold $\varepsilon : A \to (0, +\infty)$. Then there exists a derivation $\boldsymbol{b} \in$ span$\{\Upsilon\}$ such that*

$$\mu\left\{x \in A \ : \ \|v_{\boldsymbol{b}}(x) - v(x)\| \le \varepsilon(x)\right\} > 0. \tag{7.4.7}$$

Proof. Possibly enlarging Υ with all finite $\mathbb{Q}$-linear combinations, we can say that by definition of S_Υ, we have that

$$v(x) \in S_\Upsilon(x) \quad \Longleftrightarrow \quad \inf_{i\in\mathbb{N}} \|v_{\boldsymbol{b}_i}(x) - v(x)\| = 0, \tag{7.4.8}$$

where $\{\boldsymbol{b}_i\}_{i\in\mathbb{N}} = \Upsilon$, because Υ is still countable. Now, for every $x \in A$ we can choose an index i that realizes $\|v_{\boldsymbol{b}_i}(x) - v(x)\| \le \varepsilon(x)$, for example:

$$i(x) := \inf\left\{i \in \mathbb{N} \ : \ \|v_{\boldsymbol{b}_i}(x) - v(x)\| \le \varepsilon(x)\right\}. \tag{7.4.9}$$

Now, letting $A_k = \{x \ : \ i(x) = k\}$, it is clear by (7.4.8), (7.4.9) that $\bigcup_k A_k = A$ and in particular we have that there exists a k_0 such that $\mu(A_{k_0}) > 0$ (all the measurability issues are trivial because the maps $v_{\boldsymbol{b}_i}$ and v are Borel). Now it is clear that $\boldsymbol{b} = \boldsymbol{b}_{k_0}$ satisfies (7.4.7).

□

Theorem 7.4.8 *Let $f \in W^{1,p}(X, \|\cdot\|_X, \mathfrak{m}) \cap C^1(X)$. Then $|\nabla f|_p(x) = |df|_{S_p}|(x)$ for $\mathfrak{m}$-almost every $x \in X$.*

Proof. We begin to show that $|df|_{S_p}|$ is a weak gradient. In fact it is easy to see that for every $\boldsymbol{b} \in \mathrm{Der}^{q,q}$ we have

$$|\boldsymbol{b}(f)| = |df(v_{\boldsymbol{b}})| \leq |df|_{S_p}| \cdot |v_{\boldsymbol{b}}| = |df|_{S_p}| \cdot |\boldsymbol{b}|.$$

In order to show the other inequality we suppose that there exists a set of positive measure A where $|df|_{S_p}| > |\nabla f|_p$; in this set we can find a unit vector in $v \in S_p$ such that $|df|_{S_p}| = df(v)$ (the map v is Borel thanks to the continuity of df). Then for every $x \in A$ we define $\varepsilon(x) = \frac{|df|_{S_p}| - |\nabla f|_p}{2|df|_{S_p}|} > 0$, and then we apply Lemma 7.4.7 to obtain an admissible derivation $\boldsymbol{b}$ such that

$$\mu \{x \in A \ : \ \|v_{\boldsymbol{b}}(x) - v(x)\| \leq \varepsilon(x)\} > 0. \tag{7.4.10}$$

Now, thanks to Proposition 7.4.6 we should have

$$\begin{aligned} |\nabla f|_p &\geq \frac{|df(v_{\boldsymbol{b}})|}{\|v_{\boldsymbol{b}}\|} \geq \frac{|df(v + (v_{\boldsymbol{b}} - v)|}{\|v + (v_{\boldsymbol{b}} - v)\|} \\ &\geq \frac{|df(v)| - |df_{S_p}| \cdot \|v_{\boldsymbol{b}} - v\|}{\|v\| + \|v_{\boldsymbol{b}} - v\|} \\ &= |df|_{S_p}| \frac{1 - \varepsilon}{1 + \varepsilon} \\ &> |df|_{S_p}|(1 - 2\varepsilon) = |\nabla f|_p, \end{aligned}$$

getting a contradiction. ☐

Bibliography

[1] L. Ahlfors and A. Beurling, "Conformal invariants and function-theoretic null-sets," *Acta Math.*, vol. 83, pp. 101–129, 1950.

[2] N. Aïssaoui, "Another extension of Orlicz-Sobolev spaces to metric spaces," *Abstr. Appl. Anal.*, no. 1, pp. 1–26, 2004.

[3] G. Alberti and A. Marchese, "On the differentiability of lipschitz functions with respect to measures in the euclidean space," 2014.

[4] L. Ambrosio, M. Colombo, and S. Di Marino, "Sobolev spaces in metric measure spaces: reflexivity and lower semicontinuity of slope," 2013.

[5] L. Ambrosio and S. Di Marino, "Equivalent definitions of BV space and of total variation on metric measure spaces," *J. Funct. Anal.*, vol. 266, no. 7, pp. 4150–4188, 2014.

[6] L. Ambrosio, S. Di Marino, and G. Savaré, "On the duality between p-modulus and probability measures," 2013.

[7] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, ser. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000, pp. xviii+434.

[8] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Second, ser. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008, pp. x+334.

[9] ——, "Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces," *Rev. Mat. Iberoam.*, vol. 29, no. 3, pp. 969–996, 2013.

[10] ——, "Heat flow and calculus on metric measure spaces with Ricci curvature bounded below—the compact case," in *Analysis and numerics of partial differential equations*, ser. Springer INdAM Ser. Vol. 4, Springer, Milan, 2013, pp. 63–115.

[11] ——, "Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below," *Invent. Math.*, vol. 195, no. 2, pp. 289–391, 2014.

[12] L. Ambrosio, A. Pinamonti, and G. Speight, "Tensorization of cheeger energies, the space $H^{1,1}$ and the area formula for graphs," 2014.

[13] L. Ambrosio and P. Tilli, *Topics on analysis in metric spaces*, ser. Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, Oxford, 2004, vol. 25, pp. viii+133.

[14] L. Ambrosio and D. Trevisan, "Well posedness of lagrangian flows and continuity equations in metric measure spaces," 2014.

[15] D. Bate, "Structure of measures in Lipschitz differentiability spaces," *J. Amer. Math. Soc.*, to appear.

[16] G. Bellettini, V. Caselles, and M. Novaga, "The total variation flow in $\mathbb{R}^N$," *J. Differential Equations*, vol. 184, no. 2, pp. 475–525, 2002.

[17] A. Björn and J. Björn, *Nonlinear potential theory on metric spaces*, ser. EMS Tracts in Mathematics. European Mathematical Society (EMS), Zürich, 2011, vol. 17, pp. xii+403.

[18] ——, "Obstacle and dirichlet problems on arbitrary nonopen sets, and fine topology," 2012.

[19] J. Björn, S. Buckley, and S. Keith, "Admissible measures in one dimension," *Proc. Amer. Math. Soc.*, vol. 134, no. 3, 703–705 (electronic), 2006.

[20] V. I. Bogachev, *Measure theory. Vol. I, II.* Springer-Verlag, Berlin, 2007, Vol. I: xviii+500 pp., Vol. II: xiv+575.

[21] ——, *Differentiable measures and the Malliavin calculus*, ser. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2010, vol. 164, pp. xvi+488.

[22] G. Bouchitte, G. Buttazzo, and P. Seppecher, "Energies with respect to a measure and applications to low-dimensional structures," *Calc. Var. Partial Differential Equations*, vol. 5, no. 1, pp. 37–54, 1997.

[23] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co. Inc., New York, 1973, pp. vi+183, North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).

[24] T. Champion, L. De Pascale, and P. Juutinen, "The ∞-Wasserstein distance: local solutions and existence of optimal transport maps," *SIAM J. Math. Anal.*, vol. 40, no. 1, pp. 1–20, 2008.

[25] J. Cheeger, "Differentiability of Lipschitz functions on metric measure spaces," *Geom. Funct. Anal.*, vol. 9, no. 3, pp. 428–517, 1999.

[26] V. Chiadò Piat and F. Serra Cassano, "Relaxation of degenerate variational integrals," *Nonlinear Anal.*, vol. 22, no. 4, pp. 409–424, 1994.

[27] M. Christ, "A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral," *Colloq. Math.*, vol. 60/61, no. 2, pp. 601–628, 1990.

[28] R. R. Coifman and G. Weiss, *Analyse harmonique non-commutative sur certains espaces homogènes*, ser. Lecture Notes in Mathematics, Vol. 242. Springer-Verlag, Berlin-New York, 1971, pp. v+160, Étude de certaines intégrales singulières.

[29] M. Colombo, L. De Pascale, and S. Di Marino, "Multimarginal optimal transpor maps for 1-dimensional repulsive costs," *Canad. J. Math.*, Mar. 2014.

[30] M. Colombo and S. Di Marino, "Equality between monge and kantorovich multimarginal problems with coulomb cost," *Ann. Mat. Pura Appl.*, Sep. 2013.

[31] G. Dal Maso, *An introduction to* Γ*-convergence*, ser. Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston Inc., Boston, MA, 1993, pp. xiv+340.

[32] G. De Philippis, S. Di Marino, and M. Focardi, "Lower semicontinuity for non-coercive polyconvex integrals in the limit case," 2014.

[33] C. Dellacherie and P.-A. Meyer, *Probabilités et potentiel.* Hermann, Paris, 1975, pp. x+291, Chapitres I à IV, Édition entièrement refondue, Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. XV, Actualités Scientifiques et Industrielles, No. 1372.

[34] S. Di Marino and G. Speight, "The p-weak gradient depends on p," 2014.

[35] E. B. Fabes, C. E. Kenig, and R. P. Serapioni, "The local regularity of solutions of degenerate elliptic equations," *Comm. Partial Differential Equations*, vol. 7, no. 1, pp. 77–116, 1982.

[36] K. J. Falconer, *The geometry of fractal sets*, ser. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1986, vol. 85, pp. xiv+162.

[37] B. Fuglede, "Extremal length and functional completion," *Acta Math.*, vol. 98, pp. 171–219, 1957.

[38] N. Fusco and G. Moscariello, "L^2-lower semicontinuity of functionals of quadratic type," *Ann. Mat. Pura Appl. (4)*, vol. 129, 305–326 (1982), 1981.

[39] N. Gigli, "Nonsmooth differential geometry - an approach tailored for spaces with ricci curvature bounded from below," 2014.

[40] N. Gigli and B. Han, "Independence on p of weak upper gradients in rcd spaces," 2014.

[41] N. Gigli, K. Kuwada, and S.-I. Ohta, "Heat flow on Alexandrov spaces," *Comm. Pure Appl. Math.*, vol. 66, no. 3, pp. 307–331, 2013.

[42] N. Gozlan, C. Roberto, and P.-M. Samson, "Hamilton Jacobi equations on metric spaces and transport entropy inequalities," *Rev. Mat. Iberoam.*, vol. 30, no. 1, pp. 133–163, 2014.

[43] P. Hajlasz and P. Koskela, "Sobolev met poincaré," *Mem. Amer. Math. Soc.*, vol. 154, no. 688, 2000.

[44] H. Hakkarainen, J. Kinnunen, P. Lathi, and P. Lehtelä, "Relaxation and integral representation for functionals of linear growth on metric measure spaces," 2014.

[45] J. Heinonen, *Lectures on analysis on metric spaces*, ser. Universitext. Springer-Verlag, New York, 2001, pp. x+140.

[46] ——, *Lectures on analysis on metric spaces*, ser. Universitext. Springer-Verlag, New York, 2001, pp. x+140.

[47] ——, "Nonsmooth calculus," *Bull. Amer. Math. Soc. (N.S.)*, vol. 44, no. 2, pp. 163–232, 2007.

[48] J. Heinonen, T. Kilpeläinen, and O. Martio, *Nonlinear potential theory of degenerate elliptic equations*, ser. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1993, pp. vi+363, Oxford Science Publications.

[49] J. Heinonen and P. Koskela, "Quasiconformal maps in metric spaces with controlled geometry," *Acta Math.*, vol. 181, no. 1, pp. 1–61, 1998.

[50] ——, "A note on Lipschitz functions, upper gradients, and the Poincaré inequality," *New Zealand J. Math.*, vol. 28, no. 1, pp. 37–42, 1999.

[51] J. Heinonen, P. Koskela, N. Shanmugalingam, and J. Tyson, *Sobolev spaces on metric measure spaces: an approach based on upper gradients*, ser. New Mathematical Monographs. Cambridge University Press, 2015.

[52] P. E. Herman, R. Peirone, and R. S. Strichartz, "p-energy and p-harmonic functions on Sierpinski gasket type fractals," *Potential Anal.*, vol. 20, no. 2, pp. 125–148, 2004.

[53] E. Hewitt and K. Stromberg, *Real and abstract analysis*. Springer-Verlag, New York-Heidelberg, 1975, pp. x+476, A modern treatment of the theory of functions of a real variable, Third printing, Graduate Texts in Mathematics, No. 25.

[54] S. Keith, "A differentiable structure for metric measure spaces," *Adv. Math.*, vol. 183, no. 2, pp. 271–315, 2004.

[55] M. Kell, "On interpolation and curvature via wasserstein geodesics," 2014.

[56] T. Kipeläinen, P. Koskela, and H. Masaoka, "Lattice property of p-admissible weights," 2014.

[57] P. Koskela and P. MacManus, "Quasiconformal mappings and Sobolev spaces," *Studia Math.*, vol. 131, no. 1, pp. 1–17, 1998.

[58] K. Kuwada, "Duality on gradient estimates and Wasserstein controls," *J. Funct. Anal.*, vol. 258, no. 11, pp. 3758–3774, 2010.

[59] ——, "Gradient estimate for markov kernels, wasserstein control and hopf-lax formula," in *Potential Theory and its Related Fields*, K. Hirata, Ed., ser. RIMS Kôkyûroku Bessatsu, 2013, pp. 61–80.

[60] B. Levi, "Sul principio di dirichlet," *Rend. Circ. Mat. Palermo*, vol. 22, pp. 293–359, 1906.

[61] S. Lisini, "Characterization of absolutely continuous curves in Wasserstein spaces," *Calc. Var. Partial Differential Equations*, vol. 28, no. 1, pp. 85–120, 2007.

[62] ——, "Absolutely continuous curves in extended Wasserstein-Orlicz spaces," 2014.

[63] J. Lott and C. Villani, "Weak curvature conditions and functional inequalities," *J. Funct. Anal.*, vol. 245, no. 1, pp. 311–333, 2007.

[64] J. Luukkainen and E. Saksman, "Every complete doubling metric space carries a doubling measure," *Proc. Amer. Math. Soc.*, vol. 126, no. 2, pp. 531–534, 1998.

[65] L. Malý, "Minimal weak upper gradients in Newtonian spaces based on quasi-Banach function lattices," *Ann. Acad. Sci. Fenn. Math.*, vol. 38, no. 2, pp. 727–745, 2013.

[66] P. Marcellini, "Some problems of semicontinuity and of Γ-convergence for integrals of the calculus of variations," in *Proceedings of the International Meeting on Recent Methods in Nonlinear Analysis (Rome, 1978)*, Pitagora, Bologna, 1979, pp. 205–221.

[67] M. Miranda Jr., "Functions of bounded variation on "good" metric spaces," *J. Math. Pures Appl. (9)*, vol. 82, no. 8, pp. 975–1004, 2003.

[68] M. Mocanu, "On the minimal weak upper gradient of a Banach-Sobolev function on a metric space," *Sci. Stud. Res. Ser. Math. Inform.*, vol. 19, no. 1, pp. 119–129, 2009.

[69] B. Muckenhoupt, "Weighted norm inequalities for the Hardy maximal function," *Trans. Amer. Math. Soc.*, vol. 165, pp. 207–226, 1972.

[70] E. Paolini and E. Stepanov, "Structure of metric cycles and normal one-dimensional currents," *J. Funct. Anal.*, vol. 264, no. 6, pp. 1269–1295, 2013.

[71] A. Pratelli, "On the equality between Monge's infimum and Kantorovich's minimum in optimal mass transportation," *Ann. Inst. H. Poincaré Probab. Statist.*, vol. 43, no. 1, pp. 1–13, 2007.

[72] M. M. Rao and Z. D. Ren, *Applications of Orlicz spaces*, ser. Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker, Inc., New York, 2002, vol. 250, pp. xii+464.

[73] A. Schioppa, "Derivations and alberti representations," 2013.

[74] ——, "Metric currents and alberti representations," 2014.

[75] N. Shanmugalingam, "Newtonian spaces: an extension of Sobolev spaces to metric measure spaces," *Rev. Mat. Iberoamericana*, vol. 16, no. 2, pp. 243–279, 2000.

[76] K.-T. Sturm, "How to construct diffusion processes on metric spaces," *Potential Anal.*, vol. 8, no. 2, pp. 149–161, 1998.

[77] ——, "Generalized Orlicz spaces and Wasserstein distances for convex-concave scale functions," *Bull. Sci. Math.*, vol. 135, no. 6-7, pp. 795–802, 2011.

[78] H. Tuominen, "Orlicz-Sobolev spaces on metric measure spaces," *Ann. Acad. Sci. Fenn. Math. Diss.*, no. 135, p. 86, 2004, Dissertation, University of Jyväskylä, Jyväskylä, 2004.

[79] C. Villani, *Optimal transport*, ser. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 2009, vol. 338, pp. xxii+973, Old and new.

[80] N. Weaver, "Lipschitz algebras and derivations. II. Exterior differentiation," *J. Funct. Anal.*, vol. 178, no. 1, pp. 64–112, 2000.