Sobolev spaces on metric-measure spaces

Piotr Hajłasz

Contents

1. Introduction

This paper addresses various possible extensions of the classical theory of Sobolev spaces to the setting of metric spaces equipped with a Borel measure and contains new results and new proofs. It was my intention to make this paper accessible to graduate students and researchers looking for an introduction to this material. I tried to make the paper as self-contained as possible by writing detailed proofs. However, to keep the paper at a manageable length, certain important results were deleted. In addition, the reader will discover that what is missing the most are the examples, due to the fact that their generally complicated nature requires some specific knowledge of geometric analysis. The examples can be found in such areas of geometric analysis as quasi-conformal mappings, non-linear sub-elliptic equations, differential geometry, and analysis on graphs and provide the main motivation for future development of the subject.

All the bibliographical comments are moved to the last Section 12. Again, in the interest of length I decided to make only very short historical comments and provide a short list of references. Further references and comments can easily be found in the papers cited here.

The reader should be familiar with basic results from measure theory, maximal functions and classical theory of Sobolev spaces. Without familiarity with the subject, the reader will not have appropriate motivation for what is discussed here and will not see the similarities between the metric theory and the classical theory.

This paper is a modified version of my lectures delivered during the trimester *Heat Kernels, Random Walks & Analysis on Manifolds & Graphs* at the Centre Emile Borel, June 2002, Paris. I would like to express my deepest gratitude to the organizers, Pascal Auscher, Gérard Besson, Thierry Coulhon and Alexander Grigor'yan, for giving me opportunity to participate in this excellent event. The paper was completed during my stay in the Department of Mathematics at the University of Michigan. I wish to thank both the Centre Emilie Borel and the University of Michigan for their support and hospitality.

2. Classical Sobolev spaces

The theory of Sobolev spaces is a central analytic tool in the study of various aspects of partial differential equations and calculus of variations. However, the scope of its applications is much wider, including questions in differential geometry, algebraic topology, complex analysis, and in probability theory, just to name a few.

The classical Sobolev spaces are defined on open sets in the Euclidean spaces or Riemannian manifolds. If $\Omega \subset \mathbb{R}^n$ is open and $1 \leq p < \infty$ then the Sobolev space $W^{1,p}(\Omega)$ is defined as the collection of all functions $u \in L^p(\Omega)$ with the distributional gradient in L^p.[1] $W^{1,p}(\Omega)$ is a Banach space with respect to the norm $\|u\|_{W^{1,p}} = \|u\|_{L^p} + \|\nabla u\|_{L^p}$. Since locally Lipschitz (or even smooth) functions are dense in $W^{1,p}(\Omega)$, one can equivalently define the Sobolev space as the closure of the subspace of locally Lipschitz functions in the Sobolev norm.

The definition of the Sobolev space strongly involves the Euclidean structure of the underlying domain Ω suggesting that there is no reasonable way to extend this definition to the case in which Ω is replaced by a metric space equipped with a Borel measure.[2] The purpose of this paper is to show that this expectation is not correct. We will provide several different approaches to Sobolev spaces on metric-measure spaces and we will show that, in a surprisingly general setting, all the definitions are equivalent.

The theory of Sobolev spaces on metric-measure spaces turns out to have many applications to Carnot–Carathéodory spaces, sub-elliptic equations, quasi-conformal mappings on Carnot groups and more general Loewner spaces, analysis on topological manifolds, potential theory on infinite graphs, analysis on fractals, theory of Dirichlet forms, variational problems in the setting of metric spaces, hyperbolic buildings, Gromov hyperbolic spaces, rigidity results and other problems of metric differential geometry. Unfortunately there is no space here to discuss those applications. We will concentrate on the abstract approach only. For more comments and references, see Section 12.

---

[1] We do not consider the case $p = \infty$ just for simplicity's sake. In this case Sobolev functions are Lipschitz continuous and the theory has a slightly different (more classical) character.

[2] Called *metric-measure space* for short.

In order to define Sobolev spaces on metric-measure spaces we need to find another approach to Sobolev spaces, the one that will not be so strongly Euclidean and, in particular, will not involve derivatives. One such approach is suggested by the following

THEOREM 2.1. *For $u \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$, the following conditions are equivalent*

    (1) $u \in W^{1,p}(\mathbb{R}^n)$,

    (2) *There exists $0 \leq g \in L^p(\mathbb{R}^n)$ such that*

$$(2.1) \qquad \fint_B |u - u_B| \, dx \leq r \fint_B g \, dx$$

        on every ball B of any radius r.

    (3) *There exists $0 \leq g \in L^p(\mathbb{R}^n)$, and $\sigma \geq 1$ such that*

$$(2.2) \qquad \fint_B |u - u_B| \, dx \leq r \left(\fint_{\sigma B} g^p \, dx \right)^{1/p}$$

        on every ball B of any radius r.

    (4) *There exists $0 \leq g \in L^p(\mathbb{R}^n)$, and $\sigma \geq 1$ such that*

$$(2.3) \quad |u(x) - u(y)| \leq |x - y| \left(\left(\mathcal{M}_{\sigma|x-y|} g^p(x) \right)^{1/p} + \left(\mathcal{M}_{\sigma|x-y|} g^p(y) \right)^{1/p} \right) \ \ a.e.$$

Moreover each of the inequalities at 2.-4. implies that

$$|\nabla u| \leq Cg \ \ a.e.$$

Let us explain the notation. Here and in what follows the integral average is denoted by $u_B = \fint_B u \, d\mu = \mu(B)^{-1} \int_B u \, d\mu$, where μ is the measure with respect to which we integrate. In Theorem 2.1 we integrate with respect to the Lebesgue measure. By σB we denote a ball concentric with B and with radius σ times that of B. The symbols

$$(2.4) \qquad \mathcal{M}_R h(x) = \sup_{r < R} \fint_{B(x,r)} |h| \, d\mu, \qquad \mathcal{M}h(x) = \sup_{r > 0} \fint_{B(x,r)} |h| \, d\mu$$

denote the Hardy–Littlewood maximal functions. C is used to denote a general constant—its value can change even in a single string of estimates.

There are also natural versions of the theorem in which $\mathbb{R}^n$ is replaced by a bounded domain Ω with sufficiently regular boundary, but this will not be discussed here.

If $u \in W^{1,p}(\mathbb{R}^n)$, then (2.1) follows from the classical Poincaré inequality

$$\fint_B |u - u_B| \, dx \leq Cr \fint_B |\nabla u| \, dx$$

which holds true on every ball B of radius r. Inequality (2.2) is then a consequence of (2.1) and the Hölder inequality. To prove (2.3) for $u \in W^{1,p}(\mathbb{R}^n)$ we just need a well known pointwise inequality

$$(2.5) \qquad |u(x) - u(y)| \leq C|x - y| \left(\mathcal{M}_{2|x-y|} |\nabla u|(x) + \mathcal{M}_{2|x-y|} |\nabla u|(y) \right) \ \ a.e.$$

and the Hölder inequality.

We do not prove the remaining implications now. We will come back to this later and prove the theorem in a great generality of metric-measure spaces, see Theorems 9.4, 9.5 and 10.2.

Note that the conditions at 2.-4. do not involve derivatives, just the metric and measure properties of the underlying Euclidean space $\mathbb{R}^n$. Hence each of the conditions can be used to define a version of the Sobolev space in the setting of a general metric-measure spaces. In particular if (X, d, μ) is a metric space (X, d) equipped with a Borel measure μ, and we define the space $P^{1,p}(X, d, \mu)$ as the collection of all $u \in L^1_{\text{loc}}$,[3] such that there is a nonnegative $g \in L^p$ and $\sigma \geq 1$ so that

$$\fint_B |u - u_B|\, d\mu \leq r \left(\fint_{\sigma B} g^p\, d\mu \right)^{1/p}$$

on every ball $B \subset X$ of radius r, then $L^p \cap P^{1,p}$ is a natural generalization of $W^{1,p}$. The other characterizations in Theorem 2.1 could also be used to define Sobolev type spaces on metric-measure spaces; we will study this approach later in Sections 9 and 10.

Inequality (2.5) implies that for $u \in W^{1,p}(\mathbb{R}^n)$ we have

$$|u(x) - u(y)| \leq C|x - y|(\mathcal{M}|\nabla u|(x) + \mathcal{M}|\nabla u|(y)) \text{ a.e.}$$

This and the boundedness of the maximal function in L^p, $p > 1$,[4] give one direction of the proof of the following result. The other implication follows from (2.1) upon integrating (2.6) with respect to $x, y \in B$, (cf. (9.2)).

THEOREM 2.2. *Let $\Omega = \mathbb{R}^n$ or $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary and $1 < p < \infty$. Then $u \in W^{1,p}(\Omega)$, where $1 < p < \infty$, if and only if $u \in L^p(\Omega)$ and there is $0 \leq g \in L^p(\Omega)$ so that*

(2.6) $$|u(x) - u(y)| \leq |x - y|(g(x) + g(y)) \quad \text{a.e.}$$

Moreover[5] $\|\nabla u\|_{L^p} \approx \inf_g \|g\|_{L^p}$, where the infimum is taken over the class of all functions g satisfying (2.6).

This result leads to yet another approach to Sobolev spaces on metric-measure spaces, see Section 8. Observe that Theorem 2.1 allows for $p = 1$, while Theorem 2.2 requires $p > 1$.

Let us close this section by discussing one more characterization of the Sobolev space in terms of absolute continuity on lines.

Absolutely continuous functions are integrals of L^1 function. More precisely, u is *absolutely continuous* on an interval $[a, b]$ if $u(x) = c + \int_a^x h(t)\, dt$ for some $c \in \mathbb{R}$, $h \in L^1([a, b])$ and all $x \in [a, b]$.[6] We say that u is *locally absolutely continuous* on an open set $U \subset \mathbb{R}$ if it is absolutely continuous on each $[a, b] \subset U$.

Let $\Omega \subset \mathbb{R}^n$ be an open set. We say that $u \in ACL(\Omega)$ (*absolutely continuous on lines*) if u is Borel measurable and locally absolutely continuous on almost all lines parallel to coordinate axes.[7] Since absolutely continuous functions are differentiable a.e., $u \in ACL(\Omega)$ has partial derivatives a.e. and hence the vector of partial derivatives ∇u is defined a.e. Now we say that $u \in ACL^p(\Omega)$ if $u \in L^p(\Omega) \cap ACL(\Omega)$ and $|\nabla u| \in L^p(\Omega)$. The following result characterizes Sobolev spaces in terms of absolute continuity.

---

[3]By writing $u \in L^p_{\text{loc}}$ we designate that $u \in L^p(B)$ on every ball B.

[4]$\|\mathcal{M}g\|_{L^p} \leq C\|g\|_{L^p}$ for $1 < p < \infty$, see also Theorem 4.8.

[5]We write $A \approx B$ if $C^{-1}A \leq B \leq CA$ for some constant $C \geq 1$.

[6]Then it follows that u is differentiable a.e. and $u' = h$ a.e.

[7]More precisely, for almost every line ℓ parallel to one of the coordinate axes, the restriction of u to $\ell \cap \Omega$ is locally absolutely continuous on $\ell \cap \Omega$.

THEOREM 2.3. *If $\Omega \subset \mathbb{R}^n$ is open and $1 \leq p < \infty$, then $W^{1,p}(\Omega) = ACL^p(\Omega)$.*

The theorem asserts that each $ACL^p(\Omega)$ function belongs to $W^{1,p}(\Omega)$ and that the classical partial derivatives (which exist a.e. for elements of $ACL^p(\Omega)$) are equal to weak partial derivatives. On the other hand every element $u \in W^{1,p}(\Omega)$ can be alternated on a set of measure zero in a way that the resulting function belongs to $ACL^p(\Omega)$.

The proof of the inclusion $ACL^p(\Omega) \subset W^{1,p}(\Omega)$ is easy. It follows from the fact that integration by parts holds for absolutely continuous functions, the definition of the weak derivative and from the Fubini theorem. The opposite inclusion is more involved and will not be proved now. We will return to this result later, see Theorem 7.13.

In contrast with Theorems 2.1 and 2.2, there is no obvious way to use the characterization from Theorem 2.3 in order to define Sobolev type spaces on metric-measure spaces. This is because both, the notion of almost all lines parallel to coordinate axes and the notion of the gradient do not make sense for general metric-measure spaces. It is possible, however, to overcome these problems by introducing the *modulus of the path family* and the notion of *upper gradient*. The modulus of path family is an outer measure on the class of all nonconstant rectifiable curves and hence it allows us to talk about almost all curves and the notion of upper gradient is a natural replacement for the length of the gradient. This approach will be presented in Section 7.

The paper is organized as follows. In Section 3 we will discuss rectifiable curves in metric spaces. A short introduction to Borel and doubling measures is provided in Section 4. The modulus of the path family and the notion of upper gradient will be introduced in Sections 5 and 6 respectively. With these tools in hands we will introduce a Sobolev type space $N^{1,p}$ on a metric-measure space in Section 7. This approach is related to Theorem 2.3. The approaches related to Theorems 2.1 and 2.2 will be discussed in Sections 9 and 8 respectively. These two approaches are particularly effective if the measure satisfies the doubling condition. Yet another approach to Sobolev spaces based on the existence of an abstract linear operator of differentiation will be presented in Section 10. Finally Section 11 is devoted to theory of spaces supporting Poincaré inequalities. This is a special class of metric-measure spaces. It turns out that although, the definition of this class of metric-measure spaces is expressed in metric-measure terms, it has a surprisingly rich structure as it allows for the existence of a kind of a differentiable structure. Most of the results discussed in the paper are proved here. The last Section 12 is devoted to bibliographical comments and further reading.

The paper is at the same time a survey and research paper—many results are stated and proved here for the first time and some other results are furnished with new proofs. For description what is new, see Section 12.

3. Curves in metric spaces

Let (X, d) be a metric space. By a *curve* in X we will mean any continuous mapping $\gamma : [a, b] \to X$. The *image of the curve* will be denoted by $|\gamma| = \gamma([a, b])$. The *length* of γ is defined as

$$\ell(\gamma) = \sup \left\{ \sum_{i=0}^{n-1} d(\gamma(t_i), \gamma(t_{i+1})) \right\},$$

where the supremum is taken over all partitions $a = t_0 < t_1 < \ldots < t_n = b$. We say that the curve is *rectifiable* if $\ell(\gamma) < \infty$.

Observe that a curve $\gamma : [a, b] \to \mathbb{R}^n$ is rectifiable if and only if the coordinate functions are continuous and of bounded variation.

The *length function* associated with a rectifiable curve $\gamma : [a, b] \to X$ is $s_\gamma : [a, b] \to [0, \ell(\gamma)]$, given by $s_\gamma(t) = \ell(\gamma|_{[a,t]})$.

LEMMA 3.1. *Assume given a rectifiable curve* $\gamma : [a, b] \to X$. *Then the length function* $s_\gamma : [a, b] \to [0, \ell(\gamma)]$ *is nondecreasing and continuous.*

PROOF. Clearly, s_γ is nondecreasing. Continuity is less obvious. We will prove continuity of γ in the interior of $[a, b]$. The case of end points is similar and left to the reader. By contradiction suppose that there is $\tau \in (a, b)$ with

$$(3.1) \qquad \eta = \lim_{t \to \tau+} s_\gamma(t) - \lim_{t \to \tau-} s_\gamma(t) > 0.$$

Take a partition $a = t_0 < t_1 < \ldots < t_n = b$, such that

$$(3.2) \qquad \sum_{i=0}^{n-1} d(\gamma(t_i), \gamma(t_{i+1})) > \ell(\gamma) - \eta/3$$

and $d(\gamma(t_i), \gamma(t_{i+1})) < \eta/3$ for $i = 0, 1, 2, \ldots, n - 1$. We can always choose a partition such that τ is not among t_i's i.e. $\tau \in (t_i, t_{i+1})$ for some i (why?). Hence it follows from (3.1) that $\ell(\gamma|_{[t_i, t_{i+1}]}) \geq \eta$. Taking a subdivision of $[t_i, t_{i+1}]$, we can replace the summand $d(\gamma(t_i), \gamma(t_{i+1})) < \eta/3$ in (3.2) by a sum larger than $2\eta/3$. This implies, however, that the new sum in (3.2) is larger than $(\ell(\gamma) - \eta/3) - \eta/3 + 2\eta/3 = \ell(\gamma)$, an obvious contradiction. $\qquad \square$

If $\gamma : [a, b] \to X$ is a curve and $\alpha : [c, d] \to [a, b]$ is continuous, nondecreasing and onto, then we say that the curve $\gamma \circ \alpha$ is obtained from γ by a *nondecreasing change of variables.* Note that

$$(3.3) \qquad \ell(\gamma) = \ell(\gamma \circ \alpha).$$

We could also define the length of an arbitrary, not necessarily continuous, mapping $\gamma : [a, b] \to X$, in the same way as the length of a continuous curve. Of course, for rectifiability of such γ it is necessary that there are at most countably many points of discontinuity. Note that (3.3) holds true for an arbitrary mapping $\gamma : [a, b] \to X$ as well. We will need this observation just once, in the proof of Theorem 3.2, and it will be pointed out explicitly. In all other cases throughout the paper we will consider continuous curves only.

Every rectifiable curve admits a very nice parametrization by the *arc-length*. This follows from the next result.

THEOREM 3.2. *If* $\gamma : [a, b] \to X$ *is a rectifiable curve, then there is a unique curve* $\widetilde{\gamma} : [0, \ell(\gamma)] \to X$ *such that*

$$(3.4) \qquad \gamma = \widetilde{\gamma} \circ s_\gamma.$$

Moreover $\ell(\widetilde{\gamma}|_{[0,t]}) = t$ *for every* $t \in [0, \ell(\gamma)]$. *In particular* $\widetilde{\gamma} : [0, \ell(\gamma)] \to X$ *is a* 1-*Lipschitz*[8] *mapping.*

---

[8]A mapping $f : (X, d) \to (Y, \rho)$ between metric spaces is called *L-Lipschitz* if $\rho(f(x), f(y)) \leq Ld(x, y)$ for all $x, y \in X$. We call L *Lipschitz constant* of f.

REMARK 3.3. We call $\widetilde{\gamma}$ parametrized by the *arc-length* because $\ell(\widetilde{\gamma}|_{[0,t]}) = t$ for $t \in [0, \ell(\gamma)]$.

PROOF OF THEOREM 3.2. We can assume that $a = 0$. Let

$$h(t) = \inf\{s_\gamma^{-1}(t)\} \quad \text{for } t \in [0, \ell(\gamma)].$$

Since the sets $s_\gamma^{-1}(t)$ are compact, the infimum is attained i.e., $h(t) \in s_\gamma^{-1}(t)$. Hence

$$s_\gamma(h(t)) = t, \qquad h(s_\gamma(t)) \le t.$$

The last inequality follows from the observation that $t \in s_\gamma^{-1}(s_\gamma(t))$ and hence infimum of the set $s_\gamma^{-1}(s_\gamma(t))$ which, by definition, equals $h(s_\gamma(t))$ is less than or equal to t. Note that h need not be continuous. Actually intervals of constancy of s_γ correspond to jumps of h.

If the curve $\widetilde{\gamma} : [0, \ell(\gamma)] \to X$ satisfies (3.4), then $\gamma(h(t)) = \widetilde{\gamma}(s_\gamma(h(t)) = \widetilde{\gamma}(t)$ for all $t \in [0, \ell(\gamma)]$, and hence

(3.5)
$$\widetilde{\gamma}(t) = \gamma(h(t)) \quad \text{for } t \in [0, \ell(\gamma)].$$

This proves the uniqueness of $\widetilde{\gamma}$. Now it remains to show that if $\widetilde{\gamma}$ is defined by formula (3.5), then (3.4) and $\ell(\widetilde{\gamma}|_{[0,t]}) = t$ for all $t \in [0, \ell(\gamma)]$ hold true. Note that the last condition will imply that $\widetilde{\gamma}$ is 1-Lipschitz and hence continuous.

Since $h(s_\gamma(t)) \le t$, we have

$$d(\gamma(t), \gamma(h(s_\gamma(t)))) \le \ell(\gamma|_{[h(s_\gamma(t)),t]}) = s_\gamma(t) - s_\gamma(h(s_\gamma(t))) = s_\gamma(t) - s_\gamma(t) = 0.$$

Hence $(\widetilde{\gamma} \circ s_\gamma)(t) = \gamma(h(s_\gamma(t))) = \gamma(t)$ which is (3.4). The proof of the arc-length parametrization of $\widetilde{\gamma}$ is also easy

$$\ell(\widetilde{\gamma}|_{[0,t]}) = \ell(\widetilde{\gamma}|_{[0,s_\gamma(h(t))]}) = \ell(\widetilde{\gamma} \circ s_\gamma|_{[0,h(t)]}) = \ell(\gamma|_{[0,h(t)]}) = s_\gamma(h(t)) = t.$$

In the last equality we employed the fact that formula (3.3) holds also for discontinuous curves, since at this moment we did not know whether $\widetilde{\gamma}$ was continuous or not. However, now we can readily show that $\widetilde{\gamma}$ is 1-Lipschitz

$$d(\widetilde{\gamma}(t_1), \widetilde{\gamma}(t_2)) \le \ell(\widetilde{\gamma}|_{[t_1,t_2]}) = t_2 - t_1.$$

The proof is complete. $\square$

In particular, Theorem 3.2 shows that every rectifiable curve admits a 1-Lipschitz parametrization.

EXERCISE 3.4. Let $\gamma : [0,1] \to [0,1]$ be a standard Cantor staircase function. What is its arc-length parametrization $\widetilde{\gamma}$?

DEFINITION 3.5. For a curve $\gamma : [a,b] \to X$ we define *speed* at a point $t \in (a,b)$ as the limit

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

provided the limit exists.

THEOREM 3.6. *For every Lipschitz curve $\gamma : [a,b] \to X$ speed exists a.e. and*

(3.6)
$$\ell(\gamma) = \int_a^b |\dot{\gamma}|(t)\,dt.$$

PROOF. Let $\{x_n\}_{n=1}^\infty$ be a dense subset of $|\gamma| = \gamma([a,b])$. Let $\varphi_n(t) = d(\gamma(t), x_n)$. Functions $\varphi_n : [a,b] \to \mathbb{R}$ are Lipschitz continuous and hence differentiable a.e. Let $m(t) = \sup_n |\dot{\varphi}_n(t)|$. We will prove that

$$|\dot{\gamma}|(t) = m(t) \text{ a.e.} \tag{3.7}$$

Since each of the functions $x \mapsto d(x, x_n)$ is 1-Lipschitz we conclude that

$$\liminf_{h \to 0} \frac{d(\gamma(t+h), \gamma(t))}{|h|} \geq \liminf_{h \to 0} \frac{|\varphi_n(t+h) - \varphi_n(t)|}{|h|} = |\dot{\varphi}_n(t)| \text{ a.e.}$$

Taking the supremum over n yields

$$\liminf_{h \to 0} \frac{d(\gamma(t+h), \gamma(t))}{|h|} \geq m(t) \text{ a.e.} \tag{3.8}$$

In particular m is bounded by a Lipschitz constant of γ and hence integrable on $[a,b]$. On the other hand for $s \leq t$ we have
(3.9)

$$d(\gamma(t), \gamma(s)) = \sup_n |d(\gamma(t), x_n) - d(\gamma(s), x_n)| \leq \sup_n \int_s^t |\dot{\varphi}_n(\tau)| \, d\tau \leq \int_s^t m(\tau) \, d\tau.$$

Now at a Lebesgue point $t \in (a,b)$ of m we have

$$\limsup_{h \to 0} \frac{d(\gamma(t+h), \gamma(t))}{|h|} \leq \limsup_{h \to 0} \frac{1}{h} \int_t^{t+h} m(\tau) \, d\tau = m(t).$$

This together with (3.8) proves (3.7). We are left with the proof of (3.6). According to (3.9) and (3.7), for an arbitrary partition $a = t_0 < t_1 < \ldots < t_n = b$ we have

$$\sum_{i=0}^{n-1} d(\gamma(t_{i+1}), \gamma(t_i)) \leq \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} m(\tau) \, d\tau = \int_a^b |\dot{\gamma}|(\tau) \, d\tau.$$

Taking the supremum over partitions yields $\ell(\gamma) \leq \int_a^b |\dot{\gamma}|(\tau) \, d\tau$. To prove opposite inequality, fix $\varepsilon > 0$ and split $[a,b]$ into n segments of equal length i.e. $t_i = a + ih_n$, $h_n = (b-a)/n$, $i = 0, 1, 2, \ldots, n$. Take n so that $h_n < \varepsilon$. We have

$$\frac{1}{h_n} \int_a^{b-\varepsilon} d(\gamma(t+h_n), \gamma(t)) \, dt \quad \leq \quad \frac{1}{h_n} \int_0^{h_n} \sum_{i=0}^{n-2} d(\gamma(t+t_{i+1}), \gamma(t+t_i)) \, dt$$

$$\leq \quad \frac{1}{h_n} \int_0^{h_n} \ell(\gamma) = \ell(\gamma).$$

Now the definition of speed and Fatou's theorem imply

$$\int_a^{b-\varepsilon} |\dot{\gamma}|(t) \, dt \quad = \quad \int_a^{b-\varepsilon} \lim_{n \to \infty} \frac{d(\gamma(t+h_n), \gamma(t))}{h_n} \, dt$$

$$\leq \quad \liminf_{n \to \infty} \frac{1}{h_n} \int_a^{b-\varepsilon} d(\gamma(t+h_n), \gamma(t)) \, dt \leq \ell(\gamma).$$

Passing to the limit as $\varepsilon \to 0$ yields the desired inequality. $\qquad\square$

COROLLARY 3.7. $|\dot{\tilde{\gamma}}|(t) = 1$ for a.e. $t \in [0, \ell(\gamma)]$.

PROOF. $\ell(\gamma) = \ell(\tilde{\gamma}) = \int_0^{\ell(\gamma)} |\dot{\tilde{\gamma}}|(t) \, dt$. This and $|\dot{\tilde{\gamma}}|(t) \leq 1$ ($\tilde{\gamma}$ is 1-Lipschitz) implies $|\dot{\tilde{\gamma}}|(t) = 1$ a.e. $\qquad\square$

COROLLARY 3.8. *If $\gamma : [a,b] \to X$ is a Lipschitz curve then s_γ is Lipschitz and $\dot{s}_\gamma(t) = |\dot{\gamma}|(t)$ for a.e. $t \in (a,b)$.*

PROOF. For $a \le t_1 \le t_2 \le b$ we have

$$|s_\gamma(t_1) - s_\gamma(t_2)| = \ell(\gamma|_{[t_1,t_2]}) = \int_{t_1}^{t_2} |\dot{\gamma}|(\tau)\, d\tau \le L|t_1 - t_2|,$$

where L is a Lipschitz constant of γ, so the function s_γ is Lipschitz. Hence

$$\int_a^b |\dot{\gamma}|(\tau)\, d\tau = \ell(\gamma) = s_\gamma(b) - s_\gamma(a) = \int_a^b \dot{s}_\gamma(\tau)\, d\tau.$$

This and the obvious inequality $\dot{s}_\gamma \ge |\dot{\gamma}|$ yields the result. $\qquad\square$

We will use the existence of arc-length parametrizations of rectifiable curves to establish the existence of geodesics in metric spaces.

THEOREM 3.9. *If the space X is proper[9] and if there exists a rectifiable curve that joins given two points $x, y \in X$, then there exists a shortest curve that joins x and y in X.*

PROOF. Let L be the infimum of the lengths of all curves $\gamma : [a,b] \to X$, such that $\gamma(a) = x$, $\gamma(b) = y$. Clearly, $L < \infty$. Of course, we can assume that $L > 0$. We want to prove that there is a curve of length L. Let $\gamma_n : [a,b] \to X$, $\gamma_n(a) = x$, $\gamma_n(b) = y$ be a minimizing sequence i.e. $\ell(\gamma_n) \to L$ as $n \to \infty$. We can assume that $\ell(\gamma_n) = L_n < L + 1$ for every n. The family of curves $\eta_n : [0, L] \to X$, $\eta_n(t) = \widetilde{\gamma}_n(tL_n/L)$ is equicontinuous since all the curves are $(L+1)/L$-Lipschitz.[10] The images of the curves are contained in a compact set in X (as they are contained in a sufficiently large ball) and hence according to the Arzela–Ascoli theorem, we can extract a uniformly convergent subsequence $\eta_{n_k} \to \eta : [0, L] \to X$. It easily follows from the definition of the length of curve that $\ell(\eta) \le \liminf_{k \to \infty} \ell(\eta_{n_k}) = L$ (why?). Since η is a curve that joints x and y it cannot be shorter than L, hence $\ell(\eta) = L$. $\qquad\square$

A shortest curve joining two points need not be unique—think of poles on the sphere.

LEMMA 3.10. *A shortest curve $\gamma : [a,b] \to X$ that joints given two points is one-to-one.*

PROOF. Otherwise there would be "loops". Cutting loops would make the curve shorter. $\qquad\square$

Now we are ready to define the integral of a Borel function along a rectifiable curve.

DEFINITION 3.11. Let $\gamma : [a,b] \to X$ be a rectifiable curve and $\varrho : |\gamma| \to [0, \infty]$ be a Borel measurable function, where $|\gamma| = \gamma([a,b])$. Then we define

$$\int_\gamma \varrho := \int_0^{\ell(\gamma)} \varrho(\widetilde{\gamma}(t))\, dt,$$

---

[9]A metric space is proper if bounded and closed sets are compact. It is a stronger condition than being locally compact as $\mathbb{R}^n \setminus \{0\}$ is locally compact but not proper. Actually the claim of the theorem does not hold in $\mathbb{R}^n \setminus \{0\}$.

[10]Here $\widetilde{\gamma}_n$ denotes the arc-length parametrization of γ_n.

where $\widetilde{\gamma} : [0, \ell(\gamma)] \to X$ is the arc-length parametrization of γ.

THEOREM 3.12. *Let* $\gamma : [a, b] \to X$ *be a Lipschitz curve and let* $\varrho : |\gamma| \to [0, \infty]$ *be Borel measurable. Then*

$$\int_\gamma \varrho = \int_a^b \varrho(\gamma(t))|\dot{\gamma}|(t)\,dt.$$

PROOF. Since $\gamma = \widetilde{\gamma} \circ s_\gamma$ and s_γ is Lipschitz with $\dot{s}_\gamma(t) = |\dot{\gamma}|(t)$ a.e., we conclude, upon applying the classical change of variables formula, that

$$\int_\gamma \varrho = \int_0^{\ell(\gamma)} \varrho(\widetilde{\gamma}(t))\,dt = \int_a^b \varrho(\widetilde{\gamma}(s_\gamma(t)))\dot{s}_\gamma(t)\,dt = \int_a^b \varrho(\gamma(t))|\dot{\gamma}|(t)\,dt.$$

The proof is complete. $\square$

EXERCISE 3.13. There is an obvious way to define length of an "open" curve $\gamma : (a, b) \to X$. Show that if $\ell(\gamma) < \infty$, and the space is complete, then there is a unique extension of γ to a curve $\gamma : [a, b] \to X$ of the same length.

4. Borel and doubling measures

In the remaining part of the paper we will be concerned mainly with the metric-measure spaces. In this section we briefly review basic facts about Borel measures including the discussion on doubling measures. As we shall see, assuming that the metric space is equipped with a doubling measure leads to a very rich theory of Sobolev spaces.

Throughout the paper by a Borel measure we mean an outer, Borel-regular measure i.e., such an outer measure μ on a metric space (X, d) that all Borel sets are μ-measurable and for every set A there exists a Borel set B such that $A \subset B$ and $\mu(A) = \mu(B)$. With every measure μ (outer or countably additive) for which all Borel sets are measurable we can associate a Borel-regular measure by the formula

$$\widetilde{\mu}(A) = \inf\{\mu(E) : \ A \subset E \text{ and } E \text{ is Borel}\} \qquad \text{for every set } A \subset X.$$

Working with outer measures is convenient because sometimes we might want measure a set without being too concerned about its measurability. We will need the following important

THEOREM 4.1. *Suppose that* μ *is a Borel measure on* (X, d) *and that* X *can be represented as a countable union of open sets of finite measure. Then for all measurable sets* $A \subset X$ *we have*[11]

$$\mu(A) = \inf_{\substack{U \supset A \\ U - \text{open}}} \mu(U) = \sup_{\substack{C \subset A \\ C - \text{closed}}} \mu(C)\,.$$

Note that if the space X is locally compact, separable and $\mu(K) < \infty$ for every compact set K, then X can be written as a union of a countable family of open sets with finite measure so the above theorem applies.

As a corollary we obtain the following well known result.

THEOREM 4.2. *If* μ *is a Borel measure on a metric space* (X, d), *then for every* $1 \le p < \infty$, *continuous functions are dense in* $L^p(X)$.

---

[11]Actually, the first equality with the infimum over open sets holds for all (not necessarily measurable) sets $A \subset X$.

PROOF. Simple functions are dense and characteristic functions can be approximated via Theorem 4.1 and Urysohn's lemma. □

In what follows we will *always* assume that a Borel measure μ has the property that $0 < \mu(B) < \infty$ for every ball $B \subset X$.[12] By a *metric-measure space* we will *always* mean a metric space equipped with such a measure.

We say that the measure μ is *doubling* if there is a constant $C_d \geq 1$ (called *doubling constant*) such that $0 < \mu(2B) \leq C_d\mu(B) < \infty$ for every ball $B \subset X$.

Metric spaces equipped with a doubling measure are called *spaces of homogeneous type* and $s = \log_2 C_d = \log C_d / \log 2$ is called *homogeneous dimension*. Note that the homogeneous dimension is not uniquely associated with a given doubling measure as we can always take C_d larger.

Not every metric space can support a doubling measure. Indeed, we have

LEMMA 4.3. *If μ is doubling, then the metric space is doubling in the sense that there is a constant C so that every ball B of radius r can be covered by at most C balls of radius $r/2$.*

PROOF. Take a subset $\{x_i\}_{i \in I} \subset B$ maximal with respect to the property that $d(x_i, x_j) \geq r/2$ for all $i, j \in I$, $i \neq j$. Then $B \subset \bigcup_{i \in I} B(x_i, r/2)$ (by maximality) and the balls $B(x_i, r/4)$ are pairwise disjoint. Now it suffices to observe that the last condition together with the doubling condition imply that the cardinality of I is bounded by a finite number depending on C_d only. □

The doubling condition is quite restrictive as we have

LEMMA 4.4. *A doubling and complete metric space is proper.*

PROOF. Mimic standard proof of the Bolzano–Weierstrass theorem using Lemma 4.3. □

A deep result is the converse to Lemma 4.3.

THEOREM 4.5. *There is a doubling measure on a complete metric space X if and only if X is doubling.*[13]

In particular every closed subset of the Euclidean space carries a doubling measure. One can also construct doubling measures in the Euclidean space which are singular with respect to the Lebesgue measure.

An important class of doubling measures is formed by so called *n-regular* measures[14], which are measures for which there are constants $C \geq 1$ and $s > 0$ such that $C^{-1}r^s \leq \mu(B(x, r)) \leq Cr^s$ for all $x \in X$ and $0 < r < \operatorname{diam} X$. The *$s$-regular* measures are closely related to the Hausdorff measure $\mathcal{H}^s$ since we have

THEOREM 4.6. *If μ is a an s-regular measure, then there is a constant $C \geq 1$ so that $C^{-1}\mu(E) \leq \mathcal{H}^s(E) \leq C\mu(E)$ for every $E \subset X$. In particular $\mathcal{H}^s$ is s-regular too.*

---

[12]Clearly, this implies the assumptions of Theorem 4.1.

[13]The theorem is not true without the assumption that the spaces is complete: the set of rational numbers is doubling but there is no doubling measure on it (why?).

[14]Called also *Ahlfors–David* regular measures.

The proof is based on standard covering arguments and a version of Lemma 4.3 in which we estimate the number of balls of radius $r > 0$ needed to cover a ball of radius $R > 0$.

Although general doubling measure need not be regular, we always have a one-sided estimate.

LEMMA 4.7. *If the measure μ is doubling with the doubling constant C_d, and $s = \log_2 C_d$, then*

$$(4.1) \qquad \frac{\mu(B(x,r))}{\mu(B_0)} \geq 4^{-s} \left(\frac{r}{r_0} \right)^s$$

whenever B_0 is a ball of radius r_0, $x \in B_0$ and $r \leq r_0$.

PROOF. Take a positive integer k such that $2r_0 \leq 2^k r < 4r_0$. Then $B_0 \subset B(x, 2^k r)$ and hence $\mu(B_0) \leq \mu(B(x, 2^k r)) \leq C_d^k \mu(B(x,r))$. The rest is easy.   □

The exponent $s = \log_2 C_d$ cannot be lowered in general. Indeed, in the case in which μ is the Lebesgue measure in $\mathbb{R}^n$ we have $C_d = 2^n$ and hence $s = n$. Doubling measures have many properties similar to those of the Lebesgue measure in $\mathbb{R}^n$. For example the Hardy–Littlewood maximal function is bounded in L^p for $p > 1$ and it satisfies weak type estimates for $p = 1$. Moreover the Lebesgue differentiation theorem holds true.

THEOREM 4.8. *If μ is doubling, then*
(1) $\mu(\{x : \mathcal{M}g(x) > t\}) \leq Ct^{-1} \int_X |g| d\mu$ *for every $t > 0$.*
(2) $\|\mathcal{M}g\|_{L^p} \leq C\|g\|_{L^p}$, *for $1 < p < \infty$.*

THEOREM 4.9. *Assume that μ is doubling and $u \in L^1_{\mathrm{loc}}$. Fix $C > 0$. Then for μ-a.e. x the following is true. If $B(x_i, r_i)$ is a sequence of balls such that $x_i \to x$, $r_i \to 0$, and $r_i > Cd(x, x_i)$, then*

$$\lim_{i \to \infty} \fint_{B(x_i, r_i)} u \, d\mu = u(x).$$

There is an obvious way to define Lebesgue points and it easily follows from Theorem 4.9 that μ-a.e. point is Lebesgue.

5. Modulus of the path family

In this section we assume that (X, d, μ) is a metric-measure space.[15]

Let $\mathfrak{M}$ denote the family of all nonconstant rectifiable curves in X. It may well be that $\mathfrak{M} = \emptyset$, but we will be mainly interested in metric spaces for which the space $\mathfrak{M}$ is sufficiently large.

In the Euclidean space the notion of almost every straight line is self-explanatory. We want to generalize this to the setting of metric spaces. To this end we need define an outer measure on $\mathfrak{M}$. This measure will allow us talk about properties that hold for almost all curves $\gamma \in \mathfrak{M}$.

DEFINITION 5.1. For $\Gamma \subset \mathfrak{M}$, let $F(\Gamma)$ be the family of all Borel measurable functions $\varrho : X \to [0, \infty]$ such that

$$\int_\gamma \varrho \geq 1 \quad \text{for every } \gamma \in \Gamma.$$

---

[15]For detailed assumptions about the space, see Section 4.

Now for each $1 \leq p < \infty$ we define

$$\mathrm{Mod}_p(\Gamma) = \inf_{\varrho \in F(\Gamma)} \int_X \varrho^p \, d\mu.$$

The number $\mathrm{Mod}_p(\Gamma)$ is called *p-modulus* of the family Γ.

THEOREM 5.2. Mod_p *is an outer measure on* $\mathfrak{M}$ *i.e.,*

(1) $\mathrm{Mod}_p(\emptyset) = 0$,
(2) $\Gamma_1 \subset \Gamma_2 \Rightarrow \mathrm{Mod}_p(\Gamma_1) \leq \mathrm{Mod}_p(\Gamma_2)$,
(3) $\mathrm{Mod}_p(\bigcup_{i=1}^{\infty} \Gamma_i) \leq \sum_{i=1}^{\infty} \mathrm{Mod}_p(\Gamma_i)$.

PROOF. (1) $\mathrm{Mod}_p(\emptyset) = 0$ because $\varrho \equiv 0 \in F(\emptyset)$. (2) If $\Gamma_1 \subset \Gamma_2$, then $F(\Gamma_2) \subset F(\Gamma_1)$ and hence $\mathrm{Mod}_p(\Gamma_1) \leq \mathrm{Mod}_p(\Gamma_2)$. (3) We can assume that $\mathrm{Mod}_p(\Gamma_i) < \infty$ for all i. Take $\varrho_i \in F(\Gamma_i)$ such that, $\int_X \varrho_i^p \, d\mu < \mathrm{Mod}_p(\Gamma_i) + \varepsilon 2^{-i}$. Then $\varrho := (\sum_{i=1}^{\infty} \varrho_i^p)^{1/p} \in F(\bigcup_{i=1}^{\infty} \Gamma_i)$ and the claim easily follows. $\square$

The larger family of curves, the larger modulus, but also the shorter curves, the larger modulus. More precisely, we have.

LEMMA 5.3. *Let* $\Gamma_1, \Gamma_2 \subset \mathfrak{M}$. *If each curve* $\gamma \in \Gamma_1$ *contains a subcurve[16] that belongs to* Γ_2, *then* $\mathrm{Mod}_p(\Gamma_1) \leq \mathrm{Mod}_p(\Gamma_2)$

PROOF. $F(\Gamma_2) \subset F(\Gamma_1)$. $\square$

If some property holds for all curves $\gamma \in \mathfrak{M} \setminus \Gamma$, where $\mathrm{Mod}_p(\Gamma) = 0$, then we say that the property holds for *p-a.e. curve*.

In the Euclidean space the notion of *p*-a.e. curve is consistent with the notion of almost every line parallel to a given coordinate direction in the following sense.

THEOREM 5.4. *Points in* $Q^n = [0,1]^n = [0,1] \times Q^{n-1}$ *will be denoted by* $x = (x_1, x')$. *Let* $1 \leq p < \infty$. *For a Borel subset* $E \subset Q^{n-1} = [0,1]^{n-1}$ *consider the family of curves (straight segments) passing through* E *and parallel to* x_1 *i.e.,*

$$\Gamma_E = \{\gamma_{x'} : [0,1] \to Q^n : \gamma_{x'}(t) = (t, x'), \ x' \in E\}.$$

Then $\mathrm{Mod}_p(\Gamma_E) = 0$ *if and only if* $\mathcal{L}^{n-1}(E) = 0$.[17]

PROOF. $\Rightarrow$. For every $\varepsilon > 0$ there is $\varrho \in F(\Gamma_E)$ such that

$$\varepsilon > \left(\int_{\mathbb{R}^n} \varrho^p \right)^{1/p} \geq \int_E \int_0^1 \varrho(t, x') \, dt \, dx' \geq \mathcal{L}^{n-1}(E),$$

and hence $\mathcal{L}^{n-1}(E) = 0$. $\Leftarrow$. Obviously $\varrho = \chi_{[0,1] \times E} \in F(\Gamma_E)$. Since $\mathcal{L}^{n-1}(E) = 0$ we conclude that $\|\varrho\|_{L^p} = 0$ and hence $\mathrm{Mod}_p(\Gamma_E) = 0$. $\square$

The following theorem provides a convenient characterization of path families of *p*-modulus zero.

THEOREM 5.5. *Let* $\Gamma \subset \mathfrak{M}$. *Then* $\mathrm{Mod}_p(\Gamma) = 0$ *if and only if there exists a Borel measurable function* $0 \leq \varrho \in L^p(X)$ *such that,*

$$\int_\gamma \varrho = +\infty \quad \text{for every } \gamma \in \Gamma.$$

---

[16]A subcurve of $\gamma : [a,b] \to X$ is a curve of the form $\gamma' : [c,d] \to X$, where $[c,d] \subset [a,b]$ and $\gamma' = \gamma|_{[c,d]}$.

[17]$\mathcal{L}^k$ denotes the *k*-dimensional Lebesgue measure.

PROOF. ⇐. $\varrho/n \in F(\Gamma)$ for every n and hence $\operatorname{Mod}_p(\Gamma) \leq \lim_{n \to \infty} \|\varrho/n\|_{L^p}^p = 0$.

⇒. There exists $\varrho_n \in F(\Gamma)$ such that $\|\varrho_n\|_{L^p} < 2^{-n}$ and $\int_\gamma \varrho_n \geq 1$ for every $\gamma \in \Gamma$. Then $\varrho = \sum_{n=1}^\infty \varrho_n$ has all desired properties. □

COROLLARY 5.6. *If* $0 \leq g \in L^p$, $1 \leq p < \infty$, *is Borel measurable, then* $\int_\gamma g < \infty$ *for p-a.e.* $\gamma \in \mathfrak{M}$.

The following theorem will be very important in what follows.

THEOREM 5.7. *Let* $u_k : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ *be a sequence of Borel functions which converge to a Borel function* $u : X \to \overline{\mathbb{R}}$ *in* $L^p(X)$. *Then there is a subsequence* $(u_{k_j})_j$ *such that*

$$\int_\gamma |u_{k_j} - u| \to 0 \quad as \ j \to \infty,$$

for p-a.e. curve $\gamma \in \mathfrak{M}$.

PROOF. Take a subsequence $(u_{k_j})_j$ such that

$$(5.1) \qquad \int_X |u_{k_j} - u|^p \, d\mu < 2^{-pj-j}.$$

Set $g_j = |u_{k_j} - u|$, and let $\Gamma \subset \mathfrak{M}$ be the family of curves such that $\int_\gamma g_j$ does not converge to 0 as $j \to \infty$. We will show that $\operatorname{Mod}_p(\Gamma) = 0$. Denote by Γ_j the family of curves in $\mathfrak{M}$ for which $\int_\gamma g_j > 2^{-j}$. Then $2^j g_j \in F(\Gamma_j)$ and hence $\operatorname{Mod}_p(\Gamma_j) < 2^{-j}$ as a consequence of (5.1). This and the observation that $\Gamma \subset \bigcup_{j=i}^\infty \Gamma_j$ for every i implies that $\operatorname{Mod}_p(\Gamma) \leq 2^{-i+1}$ for every i and hence $\operatorname{Mod}_p(\Gamma) = 0$. □

6. Upper gradient

DEFINITION 6.1. Let $u : X \to \mathbb{R}$ be a Borel function. We say that a Borel function $g : X \to [0, \infty]$ is an *upper gradient* of u if

$$(6.1) \qquad |u(\gamma(a)) - u(\gamma(b))| \leq \int_\gamma g$$

for *every* rectifiable curve $\gamma : [a, b] \to X$. We say that g is a *p-weak upper gradient* of u if (6.1) holds on *p*-a.e. curve $\gamma \in \mathfrak{M}$.

If g is an upper gradient of u and $\widetilde{g} = g$, μ-a.e., is another nonnegative Borel function, then it may be that $\widetilde{g}$ is no longer upper gradient of u. This is an unpleasant situation. Fortunately *p*-weak upper gradients are more flexible from this point of view.

LEMMA 6.2. *If* g *is a p-weak upper gradient of* u *and* $\widetilde{g}$ *is another nonnegative Borel function such that* $\widetilde{g} = g$ μ-a.e., *then* $\widetilde{g}$ *is a p-weak upper gradient of* u *too.*

PROOF. The constant sequence $g_n = |g - \widetilde{g}|$ converges to 0 in L^p, so according to Theorem 5.7, $\int_\gamma |g - \widetilde{g}| = 0$ for *p*-a.e. $\gamma \in \mathfrak{M}$. This, however, easily implies the claim. □

The next result says that *p*-weak upper gradients can be nicely approximated by upper gradients.

LEMMA 6.3. *If g is a p-weak upper gradient of u which is finite a.e., then for every $\varepsilon > 0$ there is an upper gradient g_ε of u such that*

$$g_\varepsilon \geq g \text{ everywhere}, \qquad \text{and} \qquad \|g_\varepsilon - g\|_{L^p} < \varepsilon.$$

PROOF. Let $\Gamma \subset \mathfrak{M}$ be the family of all nonconstant rectifiable curves $\gamma : [a, b] \to X$ for which the inequality

$$|u(\gamma(a)) - u(\gamma(b))| \leq \int_\gamma g$$

is *not* satisfied. Then $\mathrm{Mod}_p(\Gamma) = 0$, and hence, according to Theorem 5.5, there exists $0 \leq \varrho \in L^p$ such that $\int_\gamma \varrho = +\infty$ for every $\gamma \in \Gamma$. Now it suffices to take $g_\varepsilon = g + \varepsilon \varrho / \|\varrho\|_{L^p}$. $\qquad\square$

The result below shows that the upper gradient is a natural generalization of $|\nabla u|$.

PROPOSITION 6.4. *If $u \in C^\infty(\Omega)$, $\Omega \subset \mathbb{R}^n$, then $|\nabla u|$ is an upper gradient of u. This upper gradient is the least one in the sense that if $g \in L^1_{\mathrm{loc}}(\Omega)$ is another upper gradient of u, then $g \geq |\nabla u|$ a.e.*[18]

REMARK 6.5. The above result is not true without the assumption that $g \in L^1_{\mathrm{loc}}$. To see this, let $u(x) = x$ on $[0, 1]$ and let $E \subset [0, 1]$ be a Cantor type set of positive Lebesgue measure. Then the function $g = \infty \cdot \chi_{[0,1]\setminus E}$ is an upper gradient of u, however, it is not true that $g \geq |u'| = 1$ a.e. One can easily modify the example to have $g < \infty$ everywhere.

PROOF OF PROPOSITION 6.4. Connecting x and y by a curve $\gamma : [a, b] \to \Omega$ parametrized by the arc-length and employing the fact that $|\dot\gamma| = 1$ a.e. we have

$$|u(x) - u(y)| = \left| \int_a^b \frac{d}{dt} u(\gamma(t)) \, dt \right| \leq \int_a^b |\nabla u(\gamma(t))| \, dt = \int_\gamma |\nabla u|.$$

The proof for the second part of the proposition is slightly more difficult. Let $\nu \in S^{n-1}$ be a direction in $\mathbb{R}^n$. Consider all the curves γ defined on finite closed intervals with $\dot\gamma = \nu$ i.e., the curves are oriented segments parallel to ν parametrized by the arc-length. Since g restricted to almost all such segments is integrable, it easily follows from the Lebesgue differentiation theorem that $g(x) \geq D_\nu u(x) = \nabla u(x) \cdot \nu$ for a.e. x. Hence this inequality holds for almost all x and all ν from a countable dense subset of S^{n-1}. Taking supremum over ν yields $g(x) \geq |\nabla u(x)|$ a.e. $\qquad\square$

It is natural to ask if there is a canonical construction of an upper gradient of a given function in the general metric space. It turns out that for an arbitrary Lipschitz function u on a metric space the function

$$|\nabla^+ u|(x) := \limsup_{\substack{y \to x \\ x \neq y}} \frac{|u(y) - u(x)|}{d(y, x)}$$

is an upper gradient.[19] One can even prove a better result that we now describe.

---

[18]Compare to Corollary 7.15.
[19]We put $|\nabla^+ u|(x) = 0$ if x is an isolated point.

DEFINITION 6.6. For a function $u : X \to \mathbb{R}$ we define the *upper* and the *lower Lipschitz constant* at a point $x \in X$ by

$$\operatorname{Lip} u(x) = \limsup_{r \to 0} \frac{L(x, u, r)}{r}, \qquad \operatorname{lip} u(x) = \liminf_{r \to 0} \frac{L(x, u, r)}{r},$$

where

$$L(x, u, r) = \sup\{|u(y) - u(x)| : d(x, y) \le r\} \quad \text{for } r > 0.$$

Clearly $\operatorname{lip} u(x) \le \operatorname{Lip} u(x) = |\nabla^+ u|(x)$. It is not difficult to show that for u Lipschitz, both functions $\operatorname{lip} u$ and $\operatorname{Lip} u$ are Borel measurable.

LEMMA 6.7. *If u is Lipschitz continuous, then $\operatorname{lip} u$ is an upper gradient of u.*

REMARK 6.8. If u is merely continuous, then $\operatorname{lip} u$ need not be upper gradient. Indeed, the standard Cantor staircase function u satisfies $u' = 0$ a.e. and hence $\operatorname{lip} u = 0$ is not an upper gradient of u.

PROOF OF LEMMA 6.7. Let $\gamma : [a, b] \to X$ be a rectifiable curve parametrized by arc-length that connects x and y. The function $u \circ \gamma$ is Lipschitz continuous and hence differentiable a.e. It is easily seen that $|(u \circ \gamma)'(t)| \le \operatorname{lip} u(\gamma(t))$ at every point t of differentiability of $u \circ \gamma$. Now the inequality

$$|u(x) - u(y)| = \left| \int_a^b \frac{d}{dt} u(\gamma(t)) \, dt \right| \le \int_a^b \operatorname{lip} u(\gamma(t)) \, dt$$

completes the proof.   □

7. Sobolev spaces $N^{1,p}$

This section is devoted to the development of the theory of Sobolev spaces on metric-measure spaces based on the notion of the upper gradient. We assume that (X, d, μ) is a metric-measure space.

Let $\widetilde{N}^{1,p}(X, d, \mu)$, $1 \le p < \infty$ be the class of all L^p integrable Borel functions on X for which there exists a p-weak upper gradient in L^p. For $u \in \widetilde{N}^{1,p}(X, d, \mu)$ we define

$$\|u\|_{\widetilde{N}^{1,p}} = \|u\|_{L^p} + \inf_g \|g\|_{L^p},$$

where the infimum is taken over all p-weak upper gradients g of u.

Lemma 6.3 shows that in the definition of $\widetilde{N}^{1,p}$ and $\|\cdot\|_{\widetilde{N}^{1,p}}$, p-weak upper gradients can be replaced by upper gradients.

DEFINITION 7.1. We define an equivalence relation in $\widetilde{N}^{1,p}$ by $u \sim v \equiv \|u - v\|_{\widetilde{N}^{1,p}} = 0$. Then the space $N^{1,p}(X, d, \mu)$[20] is defined as the quotient $\widetilde{N}^{1,p}(X, d, \mu)/\sim$ and is equipped with the norm

$$\|u\|_{N^{1,p}} := \|u\|_{\widetilde{N}^{1,p}}.$$

Note that if $u \in \widetilde{N}^{1,p}$ and $v = u$ μ-a.e. then it is not necessarily true that $v \in \widetilde{N}^{1,p}$. Nevertheless, we will show later that if $u, v \in \widetilde{N}^{1,p}$ and $u = v$ μ-a.e., then $\|u - v\|_{\widetilde{N}^{1,p}} = 0$ and hence u and v define the same element in $N^{1,p}$. We will also show that $N^{1,p}$ is a Banach space and that in the case of a domain in the Euclidean space $N^{1,p}(\Omega) = W^{1,p}(\Omega)$. We need first some definitions and auxiliary results.

---

[20]$N^{1,p}$ is known as *Newtonian space*.

DEFINITION 7.2. Let $u : X \to \mathbb{R}$ be a given function. We say that

- u is absolutely continuous along a curve $\gamma : [a, b] \to X$ if $u \circ \widetilde{\gamma}$ is absolutely continuous on $[0, \ell(\gamma)]$.
- u is absolutely continuous on p-a.e. curve (has ACC_p property) if for p-a.e. $\gamma \in \mathfrak{M}$, $u \circ \widetilde{\gamma}$ is absolutely continuous.

LEMMA 7.3. *If u is absolutely continuous along γ, then $u \circ \gamma$ is continuous.*

PROOF. $u \circ \gamma = u \circ \widetilde{\gamma} \circ s_\gamma$ and s_γ is continuous. $\square$

COROLLARY 7.4. *If u has the ACC_p property, then $u \circ \gamma$ is continuous for p-a.e. $\gamma \in \mathfrak{M}$.*

LEMMA 7.5. *If $u : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^n$ is open, is a Borel measurable function with the ACC_p property, then $u \in ACL(\Omega)$.*

PROOF. It is a direct consequence of Theorem 5.4. $\square$

LEMMA 7.6. *Every function $u \in \widetilde{N}^{1,p}(X, d, \mu)$ has the ACC_p property.*

PROOF. Let $0 \le g \in L^p$ be an upper gradient of u (the existence of g follows from Lemma 6.3). Then for every $\gamma \in \mathfrak{M}$ we have

$$|u(\widetilde{\gamma}(\beta)) - u(\widetilde{\gamma}(\alpha))| \le \int_\alpha^\beta g(\widetilde{\gamma}(\tau)) \, d\tau, \quad \text{provided } [\alpha, \beta] \subset [0, \ell(\gamma)].$$

Moreover $\int_\gamma g < \infty$ for p-a.e. $\gamma \in \mathfrak{M}$ (Corollary 5.6). For such $\gamma \in \mathfrak{M}$ we have

$$|u(\widetilde{\gamma}(\beta)) - u(\widetilde{\gamma}(\alpha))| \le \int_\alpha^\beta g(\widetilde{\gamma}(\tau)) \, d\tau < \infty \quad \text{for every } [\alpha, \beta] \subset [0, \ell(\gamma)].$$

This, in turn, implies absolute continuity of $u \circ \widetilde{\gamma}$ as a consequence of the absolute continuity of the integral. $\square$

COROLLARY 7.7. *If $u_1, u_2 \in \widetilde{N}^{1,p}(X, d, \mu)$, $u_1 = u_2$ μ-a.e., then $u_1 \sim u_2$, i.e. the two functions define exactly the same element in $N^{1,p}(X, d, \mu)$.*

PROOF. For $u = u_1 - u_2$ we have $u \in \widetilde{N}^{1,p}$ and $\|u\|_{L^p} = 0$. We want to prove that $\|u\|_{\widetilde{N}^{1,p}} = 0$. To this end it suffices to show that $u \circ \gamma \equiv 0$ for p-a.e. $\gamma \in \mathfrak{M}$.[21] Let $E = \{x : u(x) \ne 0\}$. Then $\mu(E) = 0$ and hence $g = +\infty \cdot \chi_E \in L^p$, $\|g\|_{L^p} = 0$. By Corollary 5.6 and Lemma 7.6 for p-a.e. $\gamma \in \mathfrak{M}$ we have

$$\int_\gamma g < \infty \quad \text{and} \quad u \circ \widetilde{\gamma} \text{ is absolutely continuous.}$$

The first condition implies that $g \circ \widetilde{\gamma} = 0$ a.e.,[22] so $\mathcal{L}^1(\widetilde{\gamma}^{-1}(E)) = 0$ and hence $u \circ \widetilde{\gamma} = 0$ a.e. Now continuity of $u \circ \widetilde{\gamma}$ implies that $u \circ \widetilde{\gamma} = 0$ everywhere and therefore $u \circ \gamma = 0$ everywhere. $\square$

LEMMA 7.8. *Assume that $(u_j)_{j=1}^\infty$ and $(g_j)_{j=1}^\infty$ are sequences in $L^p(X)$, $1 \le p < \infty$, weakly convergent in $L^p(X)$ to $u \in L^p(X)$ and $g \in L^p(X)$ respectively. If g_j is a p-weak upper gradient of u_j for $j = 1, 2, 3, \ldots$, then there is a representative[23] of u such that g is a p-weak upper gradient of u.*

---

[21]Because this will imply that $g \equiv 0$ is a p-weak upper gradient of u.

[22]Because $g \circ \widetilde{\gamma} = +\infty$ at points where $g \circ \widetilde{\gamma} \ne 0$.

[23]Representative in the class of functions that are equal μ-a.e.

REMARKS 7.9. (a) If we even assumed that g_j were upper gradients of u_j, then the limiting function g would be, in general, only a p-weak upper gradient of u.
(b) The choice of a suitable representative of u is essential as changing u on a set of measure zero can result in the fact that g will no longer be a p-weak upper gradient of u. However, the choice of a suitable representative of g is not needed, see Lemma 6.2.

PROOF OF LEMMA 7.8. According to Lemma 6.3 we can assume that g_j is an upper gradient of u_j. Then, observe that a convex combination $\sum_{j=1}^{n} \alpha_j g_j$ is an upper gradient of $\sum_{j=1}^{n} \alpha_j u_j$, so by Mazur's lemma[24] we can assume that $u_j \to u$ and $g_j \to g$, both in the norm of L^p (why?). Now taking a suitable subsequence we can assume that $u_{j_k} \to u$ μ-a.e. and that (Theorem 5.7)

$$\int_\gamma |g_{j_k} - g| \to 0 \quad \text{as } k \to \infty$$

for all $\gamma \in \mathfrak{M} \setminus \Gamma_1$, where $\mathrm{Mod}_p(\Gamma_1) = 0$.

Let Γ_2 be the family of all curves $\gamma \in \mathfrak{M}$ such that either $\int_\gamma g = +\infty$ or $\int_\gamma g_j = +\infty$ for some j. Then $\mathrm{Mod}_p(\Gamma_2) = 0$ (Theorem 5.5).

Let E be the set of all points x for which the convergence $u_{j_k}(x) \to u(x) < \infty$ does *not* hold. Clearly $\mu(E) = 0$ and hence the family

$$\Gamma_3 = \left\{ \gamma \in \mathfrak{M} : \mathcal{L}^1(\widetilde{\gamma}^{-1}(E)) > 0 \right\}$$

satisfies $\mathrm{Mod}_p(\Gamma_3) = 0$. Indeed, $\infty \cdot \chi_E \in F(\Gamma_3)$, $\|\infty \cdot \chi_E\|_{L^p} = 0$.

Now let $\gamma \in \mathfrak{M} \setminus (\Gamma_1 \cup \Gamma_2 \cup \Gamma_3)$. Then

$$\int_0^{\ell(\gamma)} |g_{j_k}(\widetilde{\gamma}(\tau)) - g(\widetilde{\gamma}(\tau))| \, d\tau \to 0 \quad \text{as } k \to \infty.$$

Since $\gamma \notin \Gamma_3$, we have that $\widetilde{\gamma}(\tau) \notin E$ for a.e. $\tau \in [0, \ell(\gamma)]$ and hence $u_{j_k} \circ \widetilde{\gamma}(\tau)$ converges to $u \circ \widetilde{\gamma}(\tau) < \infty$ for a.e. τ. Next we show that the family $u_{j_k} \circ \widetilde{\gamma}$ is equicontinuous on $[0, \ell(\gamma)]$. Note that both facts will imply that the sequence $u_{j_k} \circ \widetilde{\gamma}$ converges uniformly on $[0, \ell(\gamma)]$. Obviously we have

$$(7.1) \qquad |u_{j_k}(\widetilde{\gamma}(t)) - u_{j_k}(\widetilde{\gamma}(s))| \le \int_s^t g_{j_k}(\widetilde{\gamma}(\tau)) \, d\tau$$

for every $0 \le s < t \le \ell(\gamma)$. Now equicontinuity will be proved as soon as we show that for every $\varepsilon > 0$ there is $\delta > 0$ such that if $A \subset [0, \ell(\gamma)]$ satisfies $\mathcal{L}^1(A) < \delta$, then[25]

$$\sup_k \int_A g_{j_k}(\widetilde{\gamma}(\tau)) \, d\tau < \varepsilon.$$

To this end choose k_0 so large that

$$\int_0^{\ell(\gamma)} |g_{j_k}(\widetilde{\gamma}(\tau)) - g(\widetilde{\gamma}(\tau))| \, d\tau < \varepsilon/2 \quad \text{for each } k \ge k_0.$$

---

[24]Mazur's lemma says that if a sequence (x_n) of elements of a Banach space X weakly converges to $x \in X$, then there is a sequence of convex combinations of (x_n) that converges to x in the norm of X.

[25]The proof of this fact is nothing else, but the proof of the well known fact that a sequence convergent in L^1 forms an equiintegrable family of functions.

Since

$$\int_A g_{j_k}(\widetilde{\gamma}(\tau))\,d\tau \leq \int_A g(\widetilde{\gamma}(\tau))\,d\tau + \int_0^{\ell(\gamma)} |g_{j_k}(\widetilde{\gamma}(\tau)) - g(\widetilde{\gamma}(\tau))|\,d\tau,$$

it suffice to choose $\delta > 0$ such that

$$\mathcal{L}^1(A) < \delta \Rightarrow \begin{cases} \int_A g(\widetilde{\gamma}(\tau))\,d\tau < \varepsilon/2, \\ \int_A g_{j_i}(\widetilde{\gamma}(\tau))\,d\tau < \varepsilon \text{ for } i = 1, 2, \ldots, k_0 - 1. \end{cases}$$

This is possible according to the absolute continuity of the integral.[26] Thus equicontinuity of the family $(u_{j_k} \circ \widetilde{\gamma})_k$ is proved. Since the sequence $u_{j_k} \circ \widetilde{\gamma}$ converges a.e. and forms an equicontinuous family, we conclude that the sequence converges uniformly on $[0, \ell(\gamma)]$.

Now we choose a representative of u as follows

$$u(x) := \begin{cases} \lim_{k \to \infty} u_{j_k}(x) & \text{if the limit exists} \\ 0 & \text{otherwise.} \end{cases}$$

Note that the sequence $(u_{j_k})_k$ converges uniformly to u on the image $|\gamma|$ of every curve $\gamma \in \mathfrak{M} \setminus (\Gamma_1 \cup \Gamma_2 \cup \Gamma_3)$, hence passing to the limit in (7.1) with $s = 0$ and $t = \ell(\gamma)$ yields

$$|u(\widetilde{\gamma}(0)) - u(\widetilde{\gamma}(\ell(\gamma)))| \leq \int_0^{\ell(\gamma)} g(\widetilde{\gamma}(\tau))\,d\tau,$$

which means $|u(\gamma(a)) - u(\gamma(b))| \leq \int_\gamma g$ for all $\gamma \in \mathfrak{M} \setminus (\Gamma_1 \cup \Gamma_2 \cup \Gamma_3), \gamma : [a, b] \to X$. $\qquad\square$

COROLLARY 7.10. *If $u_j \in N^{1,p}(X, d, \mu)$, $1 < p < \infty$, is a bounded sequence and $u_j \rightharpoonup u$ weakly in $L^p(X)$, then there is a representative of u such that $u \in N^{1,p}(X, d, \mu)$ and*

$$\|u\|_{N^{1,p}} \leq \liminf_{j \to \infty} \|u_j\|_{N^{1,p}}$$

REMARKS 7.11. (a) The claim is not true for $p = 1$.
(b) Reflexivity of $N^{1,p}$ would readily imply the corollary, however, in the general case the reflexivity of $N^{1,p}$ is an open problem. We will come back to this and other related questions later, see Theorem 10.2, Corollary 11.7 and Theorem 8.5.

PROOF OF COROLLARY 7.10. This is a direct consequence of the fact that from a bounded sequence of upper gradients $g_j \in L^p(X)$ of u_j a subsequence weakly convergent in $L^p(X)$ can be extracted,[27] Lemma 7.8, and the weak lower semicontinuity of the L^p norm,[28] i.e. $h_j \rightharpoonup h$ in L^p implies $\|h\|_{L^p} \leq \liminf_{j \to \infty} \|h_j\|_{L^p}$. $\qquad\square$

THEOREM 7.12. *$N^{1,p}(X, d, \mu)$, $1 \leq p < \infty$ is a Banach space.*

PROOF. Let $(u_j)_{j=1}^\infty$ be a Cauchy sequence for the norm $\|\cdot\|_{N^{1,p}}$. Then $u_j \to u$ in L^p for some $u \in L^p$. We want to show that $u \in N^{1,p}$ and that $u_j \to u$ in $N^{1,p}$. To this end it suffices to show that $u \in N^{1,p}$ and that each subsequence of (u_j) contains a subsequence that converges to u in the norm of $N^{1,p}$.

---

[26]Since $\gamma \notin \Gamma_2$, each of the functions $g \circ \widetilde{\gamma}$ and $g_j \circ \widetilde{\gamma}$ is integrable on $[0, \ell(\gamma)]$.

[27]Because L^p, $1 < p < \infty$ is reflexive.

[28]More generally, weak convergence $x_n \rightharpoonup x$ in a Banach space implies $\|x\| \leq \liminf_{n \to \infty} \|x_n\|$. Indeed, for $x^* \in X^*$ with $\|x^*\| = 1$ we have $x^*(x) = \lim_{n \to \infty} x^*(x_n) \leq \liminf_{n \to \infty} \|x_n\|$ and taking supremum over x^* yields the claim (Hahn–Banach).

Take an arbitrary subsequence of (u_j) and select from this another subsequence (u_{j_k}) such that
$$\|u_{j_k} - u_{j_{k+1}}\|_{N^{1,p}} < 2^{-k}.$$
Let $\widetilde{g}_k$ be an upper gradient of $u_{j_k} - u_{j_{k+1}}$ satisfying $\|\widetilde{g}_k\|_{L^p} < 2^{-k}$. Then the function $g_k = \sum_{i=k}^{\infty} \widetilde{g}_i$ is an upper gradient of $u_{j_k} - u_{j_{k+\ell}}$ for any $\ell \geq 1$. Observe that $\|g_k\|_{L^p} < 2 \cdot 2^{-k} \to 0$ as $k \to \infty$. Since $u_{j_k} - u_{j_{k+\ell}} \to u_{j_k} - u$ in L^p and $g_k \to g_k$ as $\ell \to \infty$ (g_k is a constant sequence with respect to ℓ), Lemma 7.8 implies that g_k is a p-weak upper gradient (of a suitable representative) of $u_{j_k} - u$. Hence $u_{j_k} - u \in N^{1,p}$, $u \in N^{1,p}$ and
$$\|u_{j_k} - u\|_{N^{1,p}} \leq \|u_{j_k} - u\|_{L^p} + \|g_k\|_{L^p} \to 0$$
as $k \to \infty$. $\square$

The next theorem shows that the space $N^{1,p}$ is a natural generalization of the classical Sobolev space $W^{1,p}$ to the setting of metric spaces.

THEOREM 7.13. *If $\Omega \subset \mathbb{R}^n$ is open and $1 \leq p < \infty$, then*
$$N^{1,p}(\Omega, |\cdot|, \mathcal{L}^n) = ACL^p(\Omega) = W^{1,p}(\Omega)$$
as sets and the norms are equal.[29]

REMARKS 7.14. (a) $N^{1,p}(\Omega, |\cdot|, \mathcal{L}^n)$ is the $N^{1,p}$ space on Ω regarded as a metric space with the Euclidean metric and the Lebesgue measure.
(b) Functions in $ACL^p(\Omega)$ have partial derivatives a.e. and the space is equipped with the norm $\|u\|_{ACL^p} = \|u\|_{L^p} + \|\nabla u\|_{L^p}$. A priori this is not obvious how to relate this norm to that of the space $W^{1,p}$ as the gradient ∇u of $u \in ACL^p$ is understood in the pointwise sense, while the gradient in $W^{1,p}$ is the distributional one.

PROOF OF THEOREM 7.13. Lemmas 7.6 and 7.5 imply that $N^{1,p} \subset ACL$. Since one can easily show that $|\partial u/\partial x_i| \leq g$ for every locally integrable upper gradient g of $u \in N^{1,p}$ and all $i = 1, 2, \ldots, n$, we conclude that $N^{1,p} \subset ACL^p$. Actually, since we can rotate the coordinate system, the same argument shows that directional derivatives $D_\nu u$ exist a.e. and $|D_\nu u| \leq g$ a.e. for all directions ν. We cannot, however, follow the argument from the proof of Proposition 6.4 and conclude that $|\nabla u| \leq g$ a.e. since we do not know whether $D_\nu u = \nabla u \cdot \nu$.

The Fubini theorem, the definition of the weak derivative, and the fact that the integration by parts holds for absolutely continuos functions imply $ACL^p \subset W^{1,p}$ with $\|u\|_{ACL^p} = \|u\|_{W^{1,p}}$. More precisely the pointwise partial derivatives of $u \in ACL^p$ are equal to the distributional ones. Since $D_\nu u = \nabla u \cdot \nu$ a.e. for $u \in W^{1,p}$,[30] we conclude that this equality holds also for $u \in N^{1,p} \subset W^{1,p}$ as well. Now the same argument as in the proof of Proposition 6.4 yields $|\nabla u| \leq g$ a.e. for any locally integrable upper gradient g of u. Hence $\|u\|_{ACL^p} = \|u\|_{W^{1,p}} \leq \|u\|_{N^{1,p}}$. Thus we are left with the proof that $W^{1,p} \subset N^{1,p}$ along with the inequality $\|u\|_{N^{1,p}} \leq \|u\|_{W^{1,p}}$.

---

[29]The equality between sets $N^{1,p}$, ACL^p and $W^{1,p}$ has to be properly understood since the equivalence relations identifying Borel functions in $N^{1,p}$, ACL^p and $W^{1,p}$ are defined in a slightly different way for each of the spaces, see also comments to Theorem 2.3.

[30]This is obvious for $u \in C^\infty$ and the case of $u \in W^{1,p}$ follows from a standard approximation argument.

Let $u \in W^{1,p}(\Omega)$. Choose $u_k \in C^\infty(\Omega)$, such that $u_k \to u$ in the norm of $W^{1,p}$. Since $|\nabla u_k|$ is an upper gradient of u_k, $u_k \to u$ in L^p, and $|\nabla u_k| \to |\nabla u|$ in L^p, Lemma 7.8 implies that u has a representative for which $|\nabla u|$ is a p-weak upper gradient. This yields the embedding $W^{1,p} \subset N^{1,p}$ and $\|u\|_{N^{1,p}} \le \|u\|_{W^{1,p}}$. $\square$

As a consequence of the proof we also have the following result which generalizes Proposition 6.4.

COROLLARY 7.15. *Any function* $u \in W^{1,p}(\Omega)$, $1 \le p < \infty$ *has a representative for which* $|\nabla u|$ *is a p-weak upper gradient. On the other hand if* $g \in L^1_{\mathrm{loc}}$ *is a p-weak upper gradient of u, then* $g \ge |\nabla u|$ *a.e.*

Thus $|\nabla u|$ is the least p-weak upper gradient of $u \in W^{1,p}(\Omega)$.

THEOREM 7.16. *For every* $u \in N^{1,p}(X,d,\mu)$, $1 \le p < \infty$, *there exists the least p-weak upper gradient* $g_u \in L^p$ *of u. It is smallest in the sense that if* $g \in L^p$ *is another p-weak upper gradient of u, then* $g \ge g_u$ *μ-a.e.*[31]

PROOF. We need the following important

LEMMA 7.17. *Assume that* $u \in N^{1,p}(X,d,\mu)$, $1 \le p < \infty$, *and* $g, h \in L^p$ *are p-weak upper gradients of u. If* $E \subset X$ *is a closed set, then*

$$\varrho = g\chi_E + h\chi_{X \setminus E}$$

is a p-weak upper gradient of u as well.

PROOF. Let Γ_1 be the family of curves $\gamma \in \mathfrak{M}$ for which either $\int_\gamma (g+h) = +\infty$ or $u \circ \gamma$ is *not* continuous. Clearly $\mathrm{Mod}_p(\Gamma_1) = 0$.[32]

Let Γ_2' be the family of curves $\gamma \in \mathfrak{M}$, $\gamma : [a,b] \to X$ for which the inequality

$$|u(\gamma(a)) - u(\gamma(b))| \le \min\left(\int_\gamma g, \int_\gamma h\right)$$

is *not* satisfied. Define Γ_2 be the family of all curves $\gamma \in \mathfrak{M}$ which contain subcurves belonging to Γ_2'. Since $F(\Gamma_2') \subset F(\Gamma_2)$, we have $\mathrm{Mod}_p(\Gamma_2) \le \mathrm{Mod}_p(\Gamma_2') = 0$.[33] Now it remains to show that

$$|u(\gamma(a)) - u(\gamma(b))| \le \int_\gamma \varrho$$

for all $\gamma \in \mathfrak{M} \setminus (\Gamma_1 \cup \Gamma_2)$, $\gamma : [a,b] \to X$.

If $|\gamma| \subset E$ or $|\gamma| \subset X \setminus E$, then the inequality is obvious. Thus assume that the image $|\gamma|$ has a nonempty intersection both with E and with $X \setminus E$.

The set $\gamma^{-1}(X \setminus E)$ is open and hence it consists of a countable (or finite) number of open and disjoint intervals. Denote the intervals by $((t_i, s_i))_{i=1}^\infty$. Let $\gamma_i = \gamma|_{[t_i, s_i]}$. We have

$$\begin{aligned}
|u(\gamma(a)) - u(\gamma(b))| &\le |u(\gamma(a)) - u(\gamma(t_1))| + |u(\gamma(t_1)) - u(\gamma(s_1))| \\
&+ |u(\gamma(s_1)) - u(\gamma(b))| \le \int_{\gamma \setminus \gamma_1} g + \int_{\gamma_1} h,
\end{aligned}$$

---

[31]I do not know if the claim is true for $g \in L^1_{\mathrm{loc}}$, but I believe it is not.
[32]Theorem 5.5, Lemma 7.6 and Corollary 7.4.
[33]Cf. Lemma 5.3.

where $\gamma \setminus \gamma_1$ denotes the two curves obtained from γ by 'erasing' the interior part γ_1, i.e. the curves $\gamma|_{[a,t_1]}$ and $\gamma|_{[s_1,b]}$. Similarly we can erase a larger number of subcurves of γ. This yields

$$|u(\gamma(a)) - u(\gamma(b))| \leq \int_{\gamma \setminus \bigcup_{i=1}^n \gamma_i} g + \int_{\bigcup_{i=1}^n \gamma_i} h\,,$$

and the claim follows upon passing to the limit as $n \to \infty$ (why?). $\qquad\square$

Now we can complete the proof of the theorem. Let $m = \inf_g \|g\|_{L^p}$, where the infimum is taken over the set of all p-weak upper gradients of u. It suffices to show that there is a p-weak upper gradient g_u of u such that $\|g_u\|_{L^p} = m$. Indeed, if we suppose that $g \in L^p$ is another p-weak upper gradient of u such that the set $\{g < g_u\}$ has positive measure, then there is a closed set $E \subset \{g < g_u\}$ of positive measure[34] $\mu(E) > 0$ and hence the function $\varrho = g\chi_E + g_u\chi_{X \setminus E}$ is a p-weak upper gradient of u with $\|\varrho\|_{L^p} < m$, which is a contradiction.

Thus it remains to prove the existence of a p-weak upper gradient g_u with $\|g_u\|_{L^p} = m$.[35] Let $(g_i)_{i=1}^\infty$ be a sequence of p-weak upper gradients of u such that $\|g_i\|_{L^p} < m + 2^{-i}$. We will show that it is possible to modify the sequence (g_i) in such a way that we will obtain a new sequence of p-weak upper gradients $(\varrho_i)_{i=1}^\infty$ of u satisfying

$$\|\varrho_i\|_{L^p} < m + 2 \cdot 2^{-i}, \qquad \varrho_1 \geq \varrho_2 \geq \varrho_3 \geq \dots \ \mu\text{-a.e.}$$

The sequence $(\varrho_i)_{i=1}^\infty$ will be defined by induction. We set $\varrho_1 = g_1$. Suppose the p-weak upper gradients $\varrho_1, \varrho_2, \dots, \varrho_i$ have already been chosen. We will now define ϱ_{i+1}.

The measure ν defined by $\nu(A) = \int_A \varrho_i^p \, d\mu$ is finite and hence there is a closed set $E \subset \{g_{i+1} < \varrho_i\}$ such that

$$\nu(\{g_{i+1} < \varrho_i\} \setminus E) < 2^{-(i+1)p}.$$

Now we set $\varrho_{i+1} = g_{i+1}\chi_E + \varrho_i\chi_{X \setminus E}$. Then $\varrho_{i+1} \leq \varrho_i$ is a p-weak upper gradient and

$$
\begin{aligned}
\int_X \varrho_{i+1}^p \, d\mu &= \int_E g_{i+1}^p \, d\mu + \int_{\{g_{i+1} < \varrho_i\} \setminus E} \varrho_i^p \, d\mu + \int_{\{g_{i+1} \geq \varrho_i\}} \varrho_i^p \, d\mu \\
&\leq \int_E g_{i+1}^p \, d\mu + \int_{\{g_{i+1} \geq \varrho_i\}} g_{i+1}^p \, d\mu + 2^{-(i+1)p} \\
&\leq (m + 2^{-(i+1)})^p + 2^{-(i+1)p}.
\end{aligned}
$$

Hence $m \leq \|\varrho_{i+1}\|_{L^p} \leq m + 2 \cdot 2^{-(i+1)}$. The sequence of p-weak upper gradients $(\varrho_i)_i$ converges pointwise to a function ϱ. The dominated convergence theorem yields $\varrho_i \to \varrho$ in L^p. Obviously $\|\varrho\|_{L^p} = m$, and according to Lemma 7.8, ϱ is a p-weak upper gradient of u. The proof of the theorem is complete. $\qquad\square$

There are other possible ways to define Sobolev spaces via upper gradients. One such definition is presented below. It turns out, however, that the space defined below is isometrically isomorphic to $N^{1,p}$.

---

[34]Theorem 4.1.

[35]The case $1 < p < \infty$ easily follows from the reflexivity of L^p, Lemma 7.8 and lower semicontinuity of the L^p norm. However, since we want to cover the case $p = 1$ as well, we need a different argument.

DEFINITION 7.18. $C^{1,p}(X, d, \mu)$ is a collection of all $u \in L^p$ with the finite norm

$$(7.2) \qquad \|u\|_{C^{1,p}} = \|u\|_{L^p} + \inf_{(g_i)} \liminf_{i \to \infty} \|g_i\|_{L^p},$$

where the infimum is taken over all sequences (g_i) of nonnegative Borel functions, for which there exists a sequence $u_i \xrightarrow{L^p} u$, such that g_i is an upper gradient of u_i for all i.

THEOREM 7.19. *The spaces $C^{1,p}(X, d, \mu)$ and $N^{1,p}(X, d, \mu)$ are isometrically isomorphic for $1 < p < \infty$.*

PROOF. It easily follows from Corollary 7.10 that $\|u\|_{N^{1,p}} = \|u\|_{C^{1,p}}$. $\qquad \square$

8. Sobolev spaces $M^{1,p}$

Although the definition of $N^{1,p}$ makes sense in every metric-measure space, the theory becomes trivial if the structure of the metric space is not rich enough. Indeed, if there are no rectifiable curves, except for the constant ones, then $N^{1,p} = L^p$. This follows from an obvious observation that on such a space $g \equiv 0$ is an upper gradient of every Borel function u. This is the case if, for example, the underlying metric space is a Cantor type set or the Van Koch snowflake. As the ternary Cantor set and the Van Koch snowflake are very interesting from the point of view of analysis on these spaces, one could expect to have another approach that would provide a rich theory of Sobolev spaces on these metric-measure spaces.[36] The purpose of this section is to discuss $M^{1,p}$ spaces whose definition follows Theorem 2.2. This theory will be rich even if the underlying space contain constant rectifiable curves only, like the Cantor set or the Van Koch snowflake. If, however, the metric space has sufficiently many rectifiable curves which are, in addition, well distributed, then the two approaches are equivalent i.e. $M^{1,p} = N^{1,p}$. This will be discussed in Section 11.

Throughout this section (X, d, μ) will be a metric-measure space. At the end we will assume that μ is doubling.

DEFINITION 8.1. For $0 < p < \infty$ we define $M^{1,p}(X, d, \mu)$ to be the set of all functions $u \in L^p(X)$ for which there exists $0 \le g \in L^p(X)$ such that

$$(8.1) \qquad |u(x) - u(y)| \le d(x, y)(g(x) + g(y)) \quad \mu\text{-a.e.}$$

Denote by $D(u)$ the class of all nonnegative Borel functions g that satisfy (8.1).[37] Thus $u \in M^{1,p}$ if and only if $u \in L^p$ and $D(u) \cap L^p \ne \emptyset$. The space $M^{1,p}$ is linear and equipped with the norm[38]

$$\|u\|_{M^{1,p}} = \|u\|_{L^p} + \inf_{g \in D(u)} \|g\|_{L^p}.$$

We can also define the local space $M_{loc}^{1,p}$ by assuming that $u \in L_{loc}^p$ and $D(u) \cap L_{loc}^p \ne \emptyset$.

_____

[36]The ternary Cantor set and the Van Koch snowflake are equipped with the Euclidean metric and the Hausdorff measure in a suitable dimension.

[37]Inequality (8.1) holds a.e. in the sense that there is a set $N \subset X$ of measure zero and such that the inequality holds true for all $x, y \in X \setminus N$.

[38]Obviously $\| \cdot \|_{M^{1,p}}$ is a norm only when $1 \le p < \infty$. However, for simplicity, we will call it norm in the whole range $0 < p < \infty$.

According to Theorem 2.2 and Theorem 7.13 we have

$$(8.2) \qquad N^{1,p}(\Omega, |\cdot|, \mathcal{L}^n) = W^{1,p}(\Omega) = M^{1,p}(\Omega, |\cdot|, \mathcal{L}^n)$$

provided $1 < p < \infty$ and $\Omega \subset \mathbb{R}^n$ is a smooth bounded domain or $\Omega = \mathbb{R}^n$. Hence in this case we also have that $M^{1,p}(\Omega, |\cdot|, \mathcal{L}^n) = N^{1,p}(\Omega, |\cdot|, \mathcal{L}^n)$.

EXERCISE 8.2. Show that, in general, $M^{1,1} \neq W^{1,1}$ and hence $N^{1,1} \neq M^{1,1}$ by showing that $u(x) = -x/(|x| \log |x|)$ belongs to $W^{1,1}(I)$, but it does not belong to $M^{1,1}(I)$, where $I = (-1/4, 1/4)$.

Let us list now two basic properties of the space $M^{1,p}$.

THEOREM 8.3. $M^{1,p}(X, d, \mu)$ is a Banach space for $1 \leq p < \infty$.

Classical result on density of smooth functions in the Sobolev space has the following analogue.

THEOREM 8.4. Assume given $u \in M^{1,p}(X, d, \mu)$, $0 < p < \infty$. Then for every $\varepsilon > 0$ there is a Lipschitz function φ on X such that

(1) $\mu(\{u \neq \varphi\}) < \varepsilon$,
(2) $\|u - \varphi\|_{M^{1,p}} < \varepsilon$.

Note that in addition to the approximation in norm, the approximating function coincides with u off a set of arbitrarily small measure.

It is natural to inquire whether the space $M^{1,p}(X, d, \mu)$ is reflexive for $1 < p < \infty$. Surprisingly, it is not always the case.

THEOREM 8.5. If $C \subset [0, 1]$ is the classical ternary Cantor set, then the space $M^{1,p}(C, |\cdot|, \mathcal{H}^{\log 2/\log 3})$ contains a subspace isomorphic to ℓ^∞ and hence it is not reflexive nor is separable.

The theorem can be generalized to more general Cantor-type self-similar sets. Its proof is based on a wavelet characterization of the space $M^{1,p}$.

On the other hand if the metric-measure space is equipped with a doubling measure and admits sufficiently many well distributed rectifiable curves then the space $M^{1,p}$ is reflexive. This is true in the Euclidean case because of (8.2) and, more generally, when the underlying metric-measure space supports a Poincaré inequality, see Corollary 11.7.

As we shall see, in the general case of a metric-measure space, $M^{1,p} \subset N^{1,p}$. However, in most of the cases $M^{1,p}$ is a proper subset of $N^{1,p}$, i.e. $M^{1,p} \subsetneq N^{1,p}$. This is the case if, for example, X is the classical ternary Cantor set or the Van Koch snowflake. Indeed, in these cases $N^{1,p} = L^p$, but $M^{1,p} \subsetneq L^p$.

THEOREM 8.6. For $u \in M^{1,p}(X, d, \mu)$, where $1 \leq p < \infty$, there is a representative which belongs to $N^{1,p}(X, d, \mu)$. Moreover if $g \in D(u) \cap L^p$, then $2g$ is a p-weak upper gradient of u and hence we have a continuous embedding

$$M^{1,p}(X, d, \mu) \subset N^{1,p}(X, d, \mu), \qquad \|u\|_{N^{1,p}} \leq 2\|u\|_{M^{1,p}}.$$

PROOF. Let $u \in M^{1,p}(X, d, \mu)$. According to Theorem 8.4 there is a sequence u_k of Lipschitz functions that converge to u in the norm of $M^{1,p}$. It easily follows that we can choose $g_k \in D(u_k) \cap L^p$ such that $g_k \to g$ in L^p (why?). Modifying g_k on a set of measure zero we can assume that g_k is Borel measurable and that inequality $|u_k(x) - u_k(y)| \leq d(x, y)(g_k(x) + g_k(y))$ holds for all $x, y \in X$. We will prove that $2g_k$ is an upper gradient of u_k. Let $\gamma : [0, L] \to X$ be a rectifiable curve

parametrized by arc-length. We can assume that $\int_\gamma g_k < +\infty$, as otherwise the inequality from the definition of the upper gradient is obvious. It is a rather easy consequence of Luzin's theorem that there is a set $D \subset [0, L]$ of full measure such that for each $t_0 \in D$ there is a sequence $h_n \to 0$ for which $g_k(\gamma(t_0+h_n)) \to g_k(\gamma(t_0))$. The function $u_k \circ \gamma$ is differentiable a.e. as Lipschitz continuous and hence for a.e. $t_0 \in D$ we have

$$\left| \frac{d}{dt}\big|_{t=t_0} u_k(\gamma(t)) \right| = \left| \lim_{n\to\infty} \frac{u_k(\gamma(t_0 + h_n)) - u_k(\gamma(t_0))}{h_n} \right|$$

$$\leq \limsup_{n\to\infty} \left| \frac{d(\gamma(t_0 + h_n), \gamma(t_0))}{h_n} \right| (g_k(\gamma(t_0 + h_n)) + g_k(\gamma(t_0))) \leq 2g_k(\gamma(t_0)),$$

since $d(\gamma(t_0 + h_n), \gamma(t_0)) \leq |h_n|$. This in turn yields

$$|u_k(\gamma(L)) - u_k(\gamma(0))| = \left| \int_0^L \frac{d}{dt} u_k(\gamma(t))\, dt \right| \leq \int_0^L 2g_k(\gamma(t))\, dt,$$

which means $2g_k$ is an upper gradient of u_k. Since $u_k \to u$ in L^p and $2g_k \to 2g$ in L^p it follows from Lemma 7.8 that u has a representative for which $2g$ is a p-weak upper gradient. Hence $u \in N^{1,p}$ and $\|u\|_{N^{1,p}} \leq 2\|u\|_{M^{1,p}}$. $\qquad\square$

The most important result of the classical theory of Sobolev spaces is the Sobolev embedding theorem. Motivation for the development of the theory of $M^{1,p}$ spaces comes partially from the fact that the Sobolev embedding theorem holds in the general setting of $M^{1,p}$ spaces too. The character of the classical Sobolev embedding depends on the relation between p and the dimension of the Euclidean space. In order to extend the embedding theorem to the metric-measure case we replace the dimension by a lower bound for the growth of the measure. We say that the measure μ satisfies the $V(\sigma B_0, s, b)$ condition if

$$\mu(B(x, r)) \geq br^s, \quad \text{whenever } B(x, r) \subset \sigma B_0.$$

Here $s, b > 0$, $\sigma \geq 1$ are fixed constants and $B_0 \subset X$ is a fixed ball.

Observe that a similar estimate holds for doubling measures, Lemma 4.7, although the $V(\sigma B_0, s, b)$ condition is weaker than the inequality (4.1) and it does not imply that the measure is doubling.

THEOREM 8.7. *Assume that $u \in M^{1,p}(\sigma B_0, d, \mu)$, and $g \in D(u)$, where $0 < p < \infty$, $\sigma > 1$ and B_0 is a fixed ball of radius r_0. Assume also that the measure μ satisfies the $V(\sigma B_0, s, b)$ condition. Then there exist constants C, C_1 and C_2 depending on s, p and σ only such that*

(1) *If $0 < p < s$, then $u \in L^{p^*}(B_0)$, $p^* = sp/(s - p)$ and*

$$(8.3) \qquad \inf_{c\in\mathbb{R}} \left(\fint_{B_0} |u - c|^{p^*}\, d\mu \right)^{1/p^*} \leq C \left(\frac{\mu(\sigma B_0)}{br_0^s} \right)^{1/p} r_0 \left(\fint_{\sigma B_0} g^p\, d\mu \right)^{1/p}.$$

(2) *If $p = s$, then*

$$(8.4) \qquad \fint_{B_0} \exp \left(C_1 b^{1/s} \frac{|u - u_{B_0}|}{\|g\|_{L^s(\sigma B_0)}} \right) d\mu \leq C_2.$$

(3) *If $p > s$, then*

$$(8.5) \qquad \|u - u_{B_0}\|_{L^\infty(B_0)} \leq C \left(\frac{\mu(\sigma B_0)}{br_0^s} \right)^{1/p} r_0 \left(\fint_{\sigma B_0} g^p\, d\mu \right)^{1/p}.$$

In particular u is Hölder continuous[39] on B and

$$(8.6) \qquad |u(x) - u(y)| \leq Cb^{-1/p} d(x,y)^{1-s/p} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/p} \quad for \ x, y \in B.$$

REMARK 8.8. If $p^* \geq 1$, then (8.3) can be replaced by

$$(8.7) \qquad \left(\fint_{B_0} |u - u_{B_0}|^{p^*} \, d\mu \right)^{1/p^*} \leq C \left(\frac{\mu(\sigma B_0)}{br_0^s} \right)^{1/p} r_0 \left(\fint_{\sigma B_0} g^p \, d\mu \right)^{1/p}.$$

as the left hand sides of (8.3) and (8.7) are comparable—it easily follows from Hölder's inequality. If $p^* < 1$ then the function u need not be integrable on B_0, making impossible to evaluate the average value u_{B_0}. This is no longer the problem when $p \geq s$—it follows from the theorem that in this case the function u is integrable on B_0.

PROOF OF THEOREM 8.7. Throughout the proof, C will be a general constant depending on s, p, and σ only.[40] $A \approx B$ will mean that $C^{-1}A \leq B \leq CA$ for some constant $C = C(s, p, \sigma) \geq 1$.

We can assume that $\int_{\sigma B_0} g^p \, d\mu < \infty$ as otherwise the theorem is obvious. Subtracting a constant from u will not affect inequalities (8.3), (8.4), (8.5) and (8.6), so by selecting an appropriate constant we may assume that $\operatorname{ess\,inf}_E |u| = 0$, where $E \subset \sigma B_0$ is any subset of positive measure. The set E will be chosen later. With a correct choice of E we will manage to prove (8.3) with $(\fint_{B_0} |u|^{p^*} \, d\mu)^{1/p^*}$ on the left hand side, when $0 < p < s$, and similar modifications of inequalities (8.4) and (8.5) in the other two cases.

If $g = 0$ a.e., then u is constant and hence the theorem follows. Thus we may assume that $\int_{\sigma B_0} g^p \, d\mu > 0$. We may also assume that[41]

$$(8.8) \qquad g(x) \geq 2^{-(1+1/p)} \left(\fint_{\sigma B_0} g^p \, d\mu \right)^{1/p} > 0 \quad \text{for all } x \in \sigma B_0,$$

as otherwise we replace g by $\widetilde{g}(x) = g(x) + \left(\fint_{\sigma B_0} g^p \, d\mu \right)^{1/p}$.

Let us define auxiliary sets

$$E_k = \{ x \in \sigma B_0 : g(x) \leq 2^k \}, \quad k \in \mathbb{Z}.$$

Clearly $E_k \subset E_{k+1}$. Observe that

$$(8.9) \qquad \int_{\sigma B_0} g^p \, d\mu \approx \sum_{k=-\infty}^{\infty} 2^{kp} \mu(E_k \setminus E_{k-1}).$$

We set $a_k = \sup_{B_0 \cap E_k} |u|$. Obviously $a_k \leq a_{k+1}$. If $0 < p < s$, then

$$(8.10) \qquad \int_{B_0} |u|^{p^*} \, d\mu \leq \sum_{k=-\infty}^{\infty} a_k^{p^*} \mu(B_0 \cap (E_k \setminus E_{k-1})).$$

The idea of the proof of (8.3) is to estimate the right hand side of (8.10) in terms of the right hand side of (8.9). Similar ideas are used in the other two cases.

---

[39]After redefinition on a set of measure zero.

[40]Dependence on b will always be written explicitly.

[41]For $p \geq 1$ inequality (8.8) holds with $2^{-(1+1/p)}$ replaced by 2^{-1}.

Note that measure of the complement of E_k has a nice upper bound

$$(8.11) \qquad \mu(\sigma B_0 \setminus E_k) = \mu(\{x \in \sigma B_0 : g(x) > 2^k\}) \leq 2^{-kp} \int_{\sigma B_0} g^p \, d\mu,$$

by Chebyschev's inequality.[42]

From (8.1) it follows that

$$(8.12) \qquad |u(x) - u(y)| \leq 2^{k+1} d(x, y) \quad \text{for } x, y \in E_k,$$

so $u|_{E_k}$ is 2^{k+1}-Lipschitz. For $k \in \mathbb{Z}$ we define

$$(8.13) \qquad r_k = 2b^{-1/s} \mu(\sigma B_0 \setminus E_{k-1})^{1/s}.$$

For $x \in \sigma B_0$ let k be the least integer such that $x \in E_k$. Then a.e. point $x \in \sigma B_0$ has the property that $\mu(\sigma B_0 \setminus E_{k-1}) > 0$ (why?). Thus it suffices to consider only those k for which $\mu(\sigma B_0 \setminus E_{k-1}) > 0$ or, equivalently, $r_k > 0$.

Assume that $k \in \mathbb{Z}$ and $x_k \in E_k$ are such that

$$(8.14) \qquad \mu(\sigma B_0 \setminus E_{k-1}) > 0 \quad \text{and} \quad B(x_k, r_k) \subset \sigma B_0.$$

Then

$$\mu(B(x_k, r_k)) \geq b r_k^s > \mu(\sigma B_0 \setminus E_{k-1})$$

and hence $B(x_k, r_k) \cap E_{k-1} \neq \emptyset$. Thus there is $x_{k-1} \in E_{k-1}$ such that

$$d(x_k, x_{k-1}) < r_k \leq 2b^{-1/s} 2^{-(k-1)p/s} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/s},$$

by (8.13) and (8.11). Repeating this construction with $x_k \in E_k$ replaced by $x_{k-1} \in E_{k-1}$, and then by $x_{k-2} \in E_{k-2}$ ect. we finally obtain for $k > k_0$ a sequence

$$x_k \in E_k, \ x_{k-1} \in E_{k-1}, \ldots, x_{k_0} \in E_{k_0},$$

such that

$$(8.15) \qquad d(x_{k-i}, x_{k-(i+1)}) < r_{k-i} \leq 2b^{-1/s} 2^{-(k-(i+1))p/s} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/s}$$

for $i = 0, 1, 2, \ldots, k - k_0 - 1$. Hence

$$\begin{aligned} d(x_k, x_{k_0}) \quad &< \quad r_k + r_{k-1} + \ldots + r_{k_0+1} \leq 2b^{-1/s} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/s} \sum_{j=k_0}^{k-1} 2^{-jp/s} \\ &< \quad 2^{-k_0 p/s} \frac{2b^{-1/s}}{1 - 2^{-p/s}} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/s}. \end{aligned}$$

This is all true, provided $B(x_{k-i}, r_{k-i}) \subset \sigma B_0$ for $i = 0, 1, 2, \ldots, k - k_0 - 1$ (cf. (8.14)).[43] This condition may or may not be satisfied. If we choose, however, $x_k \in B_0$, $k > k_0$ and require that

$$(8.16) \qquad 2^{-k_0 p/s} \frac{2b^{-1/s}}{1 - 2^{-p/s}} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/s} \leq (\sigma - 1) r_0,$$

---

[42]Chebyschev's inequality says that $t^p \mu(\{|f| > t\}) \leq \int_X |f|^p \, d\mu$ for $f \in L^p$, $p > 0$ and all $t > 0$. This is true for any measure μ.

[43]We choose the first point in the sequence x_k, such that $\mu(\sigma B_0 \setminus E_{k-1}) > 0$. Then also $\mu(\sigma B_0 \setminus E_{k-i}) > 0$ for all i as above.

then clearly, all the balls $B(x_{k-i}, r_{k-i})$, $i = 0, 1, 2, \ldots, k - k_0 - 1$ are contained in σB_0.[44] Condition (8.16) is equivalent to

$$(8.17) \qquad 2^{k_0} \geq \left(\frac{2}{(1 - 2^{-p/s})(\sigma - 1)} \right)^{s/p} (br_0^s)^{-1/p} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/p}.$$

Lower bound (8.8) implies that $E_k = \emptyset$ for sufficiently small k. On the other hand $\mu(E_k) \uparrow \mu(\sigma B_0)$ as $k \to \infty$. Hence there is $\widetilde{k}_0$ such that

$$(8.18) \qquad \mu(E_{\widetilde{k}_0 - 1}) < \frac{\mu(\sigma B_0)}{2} \leq \mu(E_{\widetilde{k}_0}).$$

The inequality on the right yields $E_{\widetilde{k}_0} \neq \emptyset$ and hence according to (8.8)

$$2^{-(1+1/p)} \left(\fint_{\sigma B_0} g^p \, d\mu \right)^{1/p} \leq g(x) \leq 2^{\widetilde{k}_0}$$

for $x \in E_{\widetilde{k}_0}$. At the same time the left inequality at (8.18) and (8.11) imply

$$\frac{\mu(\sigma B_0)}{2} < \mu(\sigma B_0 \setminus E_{\widetilde{k}_0 - 1}) \leq 2^{-(\widetilde{k}_0 - 1)p} \int_{\sigma B_0} g^p \, d\mu.$$

Thus both inequalities yield

$$2^{-(1+1/p)} \left(\fint_{\sigma B_0} g^p \, d\mu \right)^{1/p} \leq 2^{\widetilde{k}_0} \leq 2^{1+1/p} \left(\fint_{\sigma B_0} g^p \, d\mu \right)^{1/p}.$$

Choose the least integer $\ell \in \mathbb{Z}$ such that

$$2^{\ell} > \max \left\{ 2^{1+1/p} \left(\frac{2}{(1 - 2^{-p/s})(\sigma - 1)} \right)^{s/p}, 1 \right\} \left(\frac{\mu(\sigma B_0)}{br_0^s} \right)^{1/p}$$

and set $k_0 = \widetilde{k}_0 + \ell$. The $V(\sigma B_0, s, b)$ condition implies that $\ell > 0$. Hence (8.18) yields $\mu(E_{k_0}) > 0$. Then inequality (8.17) holds true and

$$(8.19) \qquad 2^{k_0} \approx (br_0^s)^{-1/p} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/p}.$$

Recall that $a_k = \sup_{E_k \cap B_0} |u|$. For $k > k_0$ and $x_k \in E_k \cap B_0$ we choose a sequence $x_{k-1}, \ldots, x_{k_0}$ as above. From (8.12) we have

$$
\begin{aligned}
|u(x_k)| &\leq \left(\sum_{i=0}^{k-k_0-1} |u(x_{k-i}) - u(x_{k-(i+1)})| \right) + |u(x_{k_0})| \\
&\leq \left(\sum_{i=0}^{k-k_0-1} 2^{k-i+1} d(x_{k-i}, x_{k-(i+1)}) \right) + |u(x_{k_0})|.
\end{aligned}
$$

Hence (8.15), upon taking supremum over $x_k \in E_k \cap B_0$, yields

$$a_k \leq 8b^{-1/s} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/s} \sum_{j=k_0}^{k-1} 2^{j(1-p/s)} + \sup_{E_{k_0} \cap \sigma B_0} |u|.$$

---

[44]Inequality (8.16) comes from the estimate $\mathrm{dist}\,(B_0, X \setminus \sigma B_0) \geq (\sigma - 1)r_0$. Hence it implies $r_k + r_{k-1} + \ldots + r_{k_0+1} < \mathrm{dist}\,(B_0, X \setminus \sigma B_0)$. We need assume that $\sigma > 1$ in order to have the right hand side of (8.16) positive.

Next, we want to estimate the last term $\sup_{E_{k_0} \cap \sigma B_0} |u|$. Since $\mu(E_{k_0}) > 0$, we can assume that $\operatorname{ess\,inf}_{E_{k_0} \cap \sigma B_0} |u| = 0,$[45] i.e. there is a sequence $y_i \in E_{k_0}$ such that $u(y_i) \to 0$ as $i \to \infty$. Invoking (8.12), for $x \in E_{k_0} \cap \sigma B_0$ we have

$$|u(x)| = \lim_{i \to \infty} |u(x) - u(y_i)| \le 4\sigma r_0 2^{k_0}.$$

Concluding, for $k > k_0$ we have proved that

$$(8.20) \qquad a_k \le 8b^{-1/s} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/s} \sum_{j=k_0}^{k-1} 2^{j(1-p/s)} + 4\sigma r_0 2^{k_0}.$$

For $k \le k_0$ we will use the estimate $a_k \le a_{k_0} \le 4\sigma r_0 2^{k_0}$.
Case 1: $0 < p < s$. For every $k \in \mathbb{Z}$ we have

$$
\begin{aligned}
a_k &\le 8b^{-1/s} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/s} \sum_{j=-\infty}^{k} 2^{j(1-p/s)} + 4\sigma r_0 2^{k_0} \\
&= C 2^{k(1-p/s)} b^{-1/s} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/s} + 4\sigma r_0 2^{k_0}.
\end{aligned}
$$

We employed here the fact that $1 - p/s > 0$. Hence applying (8.10), (8.9) and (8.19) we get

$$
\begin{aligned}
\int_{B_0} |u|^{p^*} \, d\mu &\le \sum_{k=-\infty}^{\infty} a_k^{p^*} \mu(B_0 \cap (E_k \setminus E_{k-1})) \\
&\le C \left(b^{-p^*/s} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{p^*/s} \sum_{k=-\infty}^{\infty} 2^{kp} \mu(E_k \setminus E_{k-1}) + r_0^{p^*} 2^{k_0 p^*} \mu(B_0) \right) \\
&\le C \left(1 + \frac{\mu(B_0)}{br_0^s} \right) b^{-p^*/s} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{p^*/p}.
\end{aligned}
$$

Now observe that the $V(\sigma B_0, s, b)$ condition implies that $1 + \mu(B_0)/(br_0^s) \le 2\mu(B_0)/(br_0^s)$ and hence inequality (8.3) follows.
Case 2: $p = s$. It follows from Jensen's inequality[46] that

$$(8.21) \quad \left(\fint_{B_0} \exp \left(C_1 b^{1/s} \frac{|u - u_{B_0}|}{\|g\|_{L^s(\sigma B_0)}} \right) d\mu \right)^{1/2} \le \fint_{B_0} \exp \left(C_1 b^{1/s} \frac{|u|}{\|g\|_{L^s(\sigma B_0)}} \right) d\mu$$

and hence it is enough to estimate the integral on the right hand side of (8.21). It follows from (8.19) that

$$(8.22) \qquad a_{k_0} \le 4\sigma r_0 2^{k_0} \le C b^{-1/s} \left(\int_{\sigma B_0} g^s \, d\mu \right)^{1/s}.$$

Since $2^{j(1-s/s)} = 1$, (8.20) and (8.22) yield

$$a_k \le \widetilde{C} b^{-1/s} \left(\int_{\sigma B_0} g^s \, d\mu \right)^{1/s} (k - k_0)$$

---

[45]See a remark at the beginning of the proof.
[46]Jensen's inequality says that if μ is a finite measure on a set X and $\varphi : [0, \infty) \to [0, \infty)$ is a convex function, then $\varphi(\fint_X |f| \, d\mu) \le \fint_X \varphi(|f|) \, d\mu$ for any measurable function f.

for $k > k_0$. We split the integral on the right hand side of (8.21) into two parts: we estimate the integrals over $B_0 \cap E_{k_0}$ and $B_0 \setminus E_{k_0}$ separately. For the first part we have

$$\frac{1}{\mu(B_0)} \int_{B_0 \cap E_{k_0}} \ldots \leq \frac{1}{\mu(B_0)} \mu(B_0 \cap E_{k_0}) \exp\left(C_1 b^{1/s} \frac{a_{k_0}}{\|g\|_{L^s(\sigma B_0)}} \right) \leq \exp(C_1 C).$$

The second part is estimated as follows

$$\frac{1}{\mu(B_0)} \int_{B_0 \setminus E_{k_0}} \ldots \quad \leq \quad \frac{1}{\mu(B_0)} \sum_{k=k_0+1}^{\infty} \exp\left(C_1 b^{1/s} \frac{a_k}{\|g\|_{L^s(\sigma B_0)}} \right) \mu(B_0 \cap (E_k \setminus E_{k-1}))$$

$$\leq \quad \frac{1}{\mu(B_0)} \sum_{k=k_0+1}^{\infty} \exp(C_1 \widetilde{C}(k - k_0)) \mu(E_k \setminus E_{k-1}) = \heartsuit.$$

We choose the constant C_1 so that $\exp(C_1 \widetilde{C}) = 2^s$. We have

$$\heartsuit \leq \frac{1}{\mu(B_0)} 2^{-sk_0} \sum_{k=-\infty}^{\infty} 2^{ks} \mu(E_k \setminus E_{k-1}) \leq C,$$

where the last inequality is a direct consequence of (8.19), (8.9) and the $V(\sigma B_0, s, b)$ condition.

Case 3: $p > s$. It follows from (8.20) and (8.19) that

$$a_k \quad \leq \quad 8b^{-1/s} \left(\int_{\sigma B_0} g^p \, d\mu \right)^{1/s} \sum_{j=k_0}^{\infty} 2^{j(1-p/s)} + 4\sigma r_0 2^{k_0}$$

$$(8.23) \qquad\qquad \leq \quad C \left(\frac{\mu(\sigma B_0)}{br_0^s} \right)^{1/p} r_0 \left(\fint_{\sigma B_0} g^p \, d\mu \right)^{1/p}$$

for all $k > k_0$. We employed here the fact that $1 - p/s < 0$. This, however, implies that u is bounded on B_0[47] with the right hand side of (8.23) being the bound for the supremum norm of u. Now estimate (8.5) follows from the elementary inequality $\|u - u_{B_0}\|_{L^\infty(B_0)} \leq 2\|u\|_{L^\infty(B_0)}$. Thus we are left with the proof of (8.6).

If $x, y \in B_0$, $d(x,y) \leq (\sigma-1)r_0/(2\sigma)$, and $B_1 = B(x, 2d(x,y))$, then $\sigma B_1 \subset \sigma B_0$ and hence (8.5) applied on B_1 yields

$$|u(x) - u(y)| \leq 2\|u - u_{B_1}\|_{L^\infty(B_1)} \leq Cb^{-1/p} d(x,y)^{1-s/p} \left(\int_{\sigma B_1} g^p \, d\mu \right)^{1/p}.$$

If, however, $d(x,y) > (\sigma-1)r_0/(2\sigma)$, then the upper bound for $|u(x) - u(y)|$ follows directly from (8.5) applied on B_0. The proof is complete. $\square$

Let us assume now that the measure μ is doubling with the doubling constant C_d. Then Lemma 4.7 implies that the measure μ satisfies the $V(\sigma B_0, s, b)$ condition for every ball B_0, with $s = \log_2 C_d$ and $b = 4^{-s}\sigma^{-s} r_0^{-s} \mu(\sigma B_0)$. This gives the following corollary.

COROLLARY 8.9. *Assume that the measure μ is doubling with the doubling constant C_d and let $s = \log_2 C_d$ be the associated homogeneous dimension. Fix a ball B of radius r and $\sigma > 1$. Assume that $u \in M^{1,p}(\sigma B, d, \mu)$ and $g \in D(u)$, where $0 < p < \infty$. Then there exist constants C, C_1 and C_2 depending on C_d, p and σ only such that*

---

[47]More precisely, u is equal a.e. to a bounded function.

(1) *If $0 < p < s$, then $u \in L^{p^*}(B)$, $p^* = sp/(s-p)$, and*

$$(8.24) \qquad \inf_{c \in \mathbb{R}} \left(\fint_B |u-c|^{p^*} \, d\mu \right)^{1/p^*} \le Cr \left(\fint_{\sigma B} g^p \, d\mu \right)^{1/p}.$$

(2) *If $p = s$, then*

$$\fint_B \exp \left(C_1 \frac{\mu(\sigma B)^{1/s}}{r} \frac{|u - u_B|}{\|g\|_{L^s(\sigma B)}} \right) d\mu \le C_2.$$

(3) *If $p > s$, then u is Hölder continuous on B and*

$$|u(x) - u(y)| \le Cr^{s/p} d(x,y)^{1-s/p} \left(\fint_{\sigma B} g^p \, d\mu \right)^{1/p} \qquad \text{for } x, y \in B.$$

9. Sobolev spaces $P^{1,p}$

This section is devoted to an approach to Sobolev spaces on metric-measure spaces suggested by Theorem 2.1. Throughout this section we assume that (X, d, μ) is a metric-measure space, μ is doubling with the doubling constant C_d and $s = \log_2 C_d$ is the associated homogeneous dimension.

DEFINITION 9.1. Fix $\sigma \ge 1$ and $0 < p < \infty$. We say that the pair (u, g), $u \in L^1_{\text{loc}}$, $0 \le g \in L^p_{\text{loc}}$, satisfies the *p-Poincaré inequality* if

$$(9.1) \qquad \fint_B |u - u_B| \, d\mu \le r \left(\fint_{\sigma B} g^p \, d\mu \right)^{1/p} \qquad \text{on every ball } B \text{ of radius } r.$$

The class of $u \in L^1_{\text{loc}}$ for which there exists $0 \le g \in L^p_{\text{loc}}$ so that the pair (u, g) satisfies the *p*-Poincaré inequality (9.1) will be denoted by $P^{1,p}_{\sigma,\text{loc}}(X, d, \mu)$ and $P^{1,p}_{\text{loc}}(X, d, \mu) = \bigcup_{\sigma \ge 1} P^{1,p}_{\sigma,\text{loc}}(X, d, \mu)$. If in addition $g \in L^p$, then we do not write 'loc'.[48]

If $u \in M^{1,p}(X, d, \mu)$, $p \ge 1$, and $g \in D(u) \cap L^p$, then integrating inequality (8.1) and applying Hölder's inequality yield

$$(9.2) \qquad \fint_B |u - u_B| \, d\mu \le 4r \fint_B g \, d\mu \le 4r \left(\fint_B g^p \, d\mu \right)^{1/p}.$$

Thus the pair $(u, 4g)$ satisfies the *p*-Poincaré inequality with $\sigma = 1$ and hence $M^{1,p}(X, d, \mu) \subset P^{1,p}_1(X, d, \mu) \cap L^p$. The above argument does not apply for $0 < p < 1$. We can, however, go slightly below 1. To this end we need apply Corollary 8.9.

THEOREM 9.2. *Fix $\sigma > 1$. If $u \in M^{1,p}(X, d, \mu)$ and $g \in D(u)$, where $p \ge s/(s+1)$, then*

$$(9.3) \qquad \fint_B |u - u_B| \, d\mu \le Cr \left(\fint_{\sigma B} g^p \, d\mu \right)^{1/p} \qquad \text{on every ball } B \text{ of radius } r,$$

with C depending on C_d, p and σ only. In particular $M^{1,p}(X, d, \mu) \subset P^{1,p}(X, d, \mu) \cap L^p$.

PROOF. Fix a ball B. Then $u \in M^{1,p}(\sigma B, d, \mu) \subset M^{1,s/(s+1)}(\sigma B, d, \mu)$. Since the Sobolev exponent associated with $s/(s+1)$ equals $(s/(s+1))^* = 1$, inequalities (8.24), (8.7) and the Hölder inequality imply (9.3). $\qquad \square$

---

[48]Thus $u \in P^{1,p}$ if $u \in L^1_{\text{loc}}$ and there exist $\sigma \ge 1$ and $0 \le g \in L^p$ such that (9.1) holds true.

Inequality (9.3) seems much weaker than (8.1), but as we shall see, the two inequalities are almost equivalent. Indeed, we have the following

THEOREM 9.3. *The following conditions are equivalent for $s/(s+1) < p < \infty$:*
  (1) $u \in M^{1,p}(X, d, \mu)$.
  (2) $u \in L^p(X)$ *and there exist $0 \le g \in L^p$ and constants $\sigma \ge 1$ and $0 < q < p$ such that*

$$(9.4) \qquad \fint_B |u - u_B|\, d\mu \le r \left(\fint_{\sigma B} g^q\, d\mu \right)^{1/q}$$

  on every ball B of radius r.

PROOF. The implication from 1. to 2. follows from Theorem 9.2 with $q = s/(s+1)$ and $\sigma > 1$. The implication from 2. to 1. follows from the result below.

THEOREM 9.4. [49] *If the pair (u, g) satisfies the p-Poincaré inequality (9.1), $p > 0$, then*

$$|u(x) - u(y)| \le Cd(x,y) \left(\left(\mathcal{M}_{2\sigma d(x,y)} g^p(x) \right)^{1/p} + \left(\mathcal{M}_{2\sigma d(x,y)} g^p(y) \right)^{1/p} \right), \quad \mu\text{-a.e.}$$

where $\mathcal{M}_R g^p$ is the Hardy–Littlewood maximal function.[50]

Indeed, by the assumption $0 \le g \in L^p$ and the pair (u, g) satisfies the q-Poincaré inequality (9.4). Since $g^q \in L^{p/q}$ and $p/q > 1$, Theorem 4.8 implies that $\mathcal{M}g^q \in L^{p/q}$ and hence $(\mathcal{M}g^q)^{1/q} \in L^p$. Now the implication follows from the inequality

$$|u(x) - u(y)| \le Cd(x,y) \left((\mathcal{M}g^q)^{1/q}(x) + (\mathcal{M}g^q)^{1/q}(y) \right)$$

which is a direct consequence of (9.4) and Theorem 9.4. Thus we are left with the proof of Theorem 9.4.

PROOF OF THEOREM 9.4. Let $x, y \in X$ be Lebsegue points of u. Let $B_i(x) = B(x, r_i) = B(x, 2^{-i}d(x,y))$, $i = 0, 1, 2, \ldots$ Then $u_{B_i(x)} \to u(x)$ as $i \to \infty$ (cf. Theorem 4.9). We have

$$
\begin{aligned}
|u(x) - u_{B_0(x)}| &\le \sum_{i=0}^{\infty} |u_{B_i(x)} - u_{B_{i+1}(x)}| \le \sum_{i=0}^{\infty} \frac{\mu(B_i(x))}{\mu(B_{i+1}(x))} \fint_{B_i(x)} |u - u_{B_i(x)}|\, d\mu \\
&\le C \sum_{i=0}^{\infty} r_i \left(\fint_{\sigma B_i(x)} g^p\, d\mu \right)^{1/p} \le C \sum_{i=0}^{\infty} r_i \left(\mathcal{M}_{\sigma d(x,y)} g^p(x) \right)^{1/p} \\
&= Cd(x,y) \left(\mathcal{M}_{\sigma d(x,y)} g^p(x) \right)^{1/p},
\end{aligned}
$$

and by symmetry a similar estimate holds at y too

$$|u(y) - u_{B_0(y)}| \le Cd(x,y) \left(\mathcal{M}_{\sigma d(x,y)} g^p(y) \right)^{1/p}.$$

We need one more inequality to go

$$
\begin{aligned}
|u_{B_0(x)} - u_{B_0(y)}| &\le |u_{B_0(x)} - u_{2B_0(x)}| + |u_{B_0(y)} - u_{2B_0(x)}| \\
&\le C \fint_{2B_0(x)} |u - u_{2B_0(x)}|\, d\mu \le Cd(x,y) \left(\mathcal{M}_{2\sigma d(x,y)} g^p(x) \right)^{1/p}.
\end{aligned}
$$

---

[49] In the Euclidean case Theorem 9.4 is the implication 3. to 4. of Theorem 2.1.
[50] See (2.4).

Now the theorem follows from the triangle inequality

$$|u(x) - u(y)| \leq |u(x) - u_{B_0(x)}| + |u_{B_0(x)} - u_{B_0(y)}| + |u(y) - u_{B_0(y)}|$$

and the above estimates. This completes the proof of Theorem 9.4 and hence that for Theorem 9.3. $\qquad\square$

It turns out that Theorem 9.4 can be inverted.

THEOREM 9.5. *If the functions $u \in L^1_{\text{loc}}$ and $g \in L^p_{\text{loc}}$, $s/(s+1) < p < \infty$, satisfy the pointwise inequality*

$$|u(x) - u(y)| \leq d(x,y) \left(\left(\mathcal{M}_{\sigma d(x,y)} g^p(x) \right)^{1/p} + \left(\mathcal{M}_{\sigma d(x,y)} g^p(y) \right)^{1/p} \right) \quad \mu\text{-a.e.}$$

with some constant $\sigma \geq 1$, then

$$\fint_B |u - u_B| \, d\mu \leq Cr \left(\fint_{6\sigma B} g^p \, d\mu \right)^{1/p}$$

for every ball B of radius r.

PROOF. Fix a ball B of radius r. For $x, y \in 2B$ we have

$$(9.5) \qquad |u(x) - u(y)| \leq d(x,y) \left(\left(\mathcal{M}(g^p \chi_{6\sigma B})(x) \right)^{1/p} + \left(\mathcal{M}(g^p \chi_{6\sigma B})(y) \right)^{1/p} \right).$$

The weak type estimate for the maximal function, Theorem 4.8, implies that

$$(9.6) \qquad \mu \left(\left\{ (\mathcal{M}(g^p \chi_{6\sigma B}))^{1/p} > t \right\} \right) \leq \frac{C}{t^p} \int_{6\sigma B} g^p \, d\mu.$$

Hence $(\mathcal{M}(g^p \chi_{6\sigma B}))^{1/p}$ belongs to the Marcinkiewicz space,[51] $(\mathcal{M}(g^p \chi_{6\sigma B}))^{1/p} \in L^p_w(X)$.

Clearly $L^p \subset L^p_w$, by Chebyschev's inequality, but, in general, L^p_w functions need not be L^p integrable.[52] We only have local integrability with exponents less than p as the following result shows.

LEMMA 9.6. [53] *If $0 < \mu(E) < \infty$, then $L^p_w(E, \mu) \subset L^q(E, \mu)$ for all $0 < q < p$. Moreover if u satisfies*

$$(9.7) \qquad \mu(\{x \in E : |u(x)| > t\}) \leq mt^{-p} \quad \text{for all } t > 0,$$

then

$$(9.8) \qquad \left(\fint_E |u|^q \, d\mu \right)^{1/q} \leq 2^{1/q} \left(\frac{q}{p-q} \right)^{1/p} \left(\frac{m}{\mu(E)} \right)^{1/p}.$$

PROOF. We represent the integral $\int_E |u|^q \, d\mu$ using the Cavalieri principle.[54] Then we split it into two integrals $\int_0^{t_0} \ldots$ and $\int_{t_0}^{\infty} \ldots$ Next, for $t < t_0$ we estimate measure of the level sets by $\mu(X)$ and for $t > t_0$ using (9.7). Choosing appropriate constant t_0 yields the result. $\qquad\square$

---

[51]We say that the function u belongs to the Marcinkiewicz space $L^p_w(X)$, $p > 0$ (called also *weak L^p*), if there is a constant $m > 0$ such that $\mu(\{|u| > t\}) \leq mt^{-p}$ for all $t > 0$.

[52]$x^{-1} \in L^1_w(0,1)$, but $x^{-1} \notin L^1(0,1)$.

[53]The lemma is true for any finite measure.

[54]Cavalieri's principle says that for any σ-finite measure μ and $f \in L^q$, $q > 0$, we have $\int_E |f|^q \, d\mu = q \int_0^{\infty} t^{q-1} \mu(\{|f| > t\}) \, dt$. It easily follows from Fubini's theorem.

Thus $\left(\mathcal{M}(g^p\chi_{6\sigma B})\right)^{1/p} \in L^q(2B)$, for every $0 < q < p$. Therefore from (9.5) we conclude that $u \in M^{1,q}(2B, d, \mu)$, and $\left(\mathcal{M}(g^p\chi_{6\sigma B})\right)^{1/p} \in D(u) \cap L^q$. Taking $s/(s+1) = q < p$, and applying Corollary 8.9 we obtain

$$\fint_B |u - u_B|\, d\mu \leq Cr \left(\fint_{2B} \left(\mathcal{M}(g^p\chi_{6\sigma B})\right)^{q/p} d\mu\right)^{1/q} \leq Cr \left(\fint_{6\sigma B} g^p\, d\mu\right)^{1/p}.$$

The last inequality follows from (9.8) and (9.6). □

Let us define the space $M_w^{1,p}(X, d, \mu)$ as the class of all $u \in L^p$ with $D(u) \cap L_w^p \neq \emptyset$. If the pair (u, g) satisfies the p-Poincaré inequality (9.1), $p > 0$, then Theorem 9.4 implies

$$|u(x) - u(y)| \leq Cd(x, y)\left((\mathcal{M}g^p)^{1/p}(x) + (\mathcal{M}g^p)^{1/p}(y)\right).$$

This together with the weak type estimate for the maximal function yields $P^{1,p}(X, d, \mu) \cap L^p \subset M_w^{1,p}(X, d, \mu)$.

The above results, all together, give the following inclusions

$$M^{1,p} \subset P^{1,p} \cap L^p \subset M_w^{1,p} \subset M_{\text{loc}}^{1,q}$$

The first inclusion holds for $p \geq s/(s+1)$, while the remaining ones for all $0 < q < p$.

Let us close this section with a general Sobolev embedding theorem for $P^{1,p}$ spaces. The case $0 < p < s$ follows from the embedding $P^{1,p} \subset M_{\text{loc}}^{1,\alpha}$ for all $0 < \alpha < p$ and Corollary 8.9.

THEOREM 9.7. *Assume that the pair $u \in L_{\text{loc}}^1$, and $0 \leq g \in L_{\text{loc}}^p$ satisfies the p-Poincaré inequality (9.1) with $0 < p < \infty$ and $\sigma \geq 1$.*

(1) If $0 < p < s$, then for every $0 < h < p^* = sp/(s - p)$

$$(9.9) \qquad \inf_{c \in \mathbb{R}} \left(\fint_B |u - c|^h\, d\mu\right)^{1/h} \leq Cr \left(\fint_{6\sigma B} g^p\, d\mu\right)^{1/p}.$$

If in addition $g \in L_{\text{loc}}^q$, $p < q < s$, then

$$(9.10) \qquad \inf_{c \in \mathbb{R}} \left(\fint_B |u - c|^{q^*}\, d\mu\right)^{1/q^*} \leq Cr \left(\fint_{6\sigma B} g^q\, d\mu\right)^{1/q},$$

where $q^* = sq/(s - q)$ and B is any ball of radius r.
(2) If $p = s$, then

$$\fint_B \exp\left(\frac{C_1 \mu(6\sigma B)^{1/s}|u - u_B|}{r\|g\|_{L^s(6\sigma B)}}\right) d\mu \leq C_2.$$

(3) If $p > s$, then u is locally Hölder continuous and

$$|u(x) - u(y)| \leq Cr^{s/p}d(x, y)^{1 - s/p} \left(\fint_{6\sigma B_0} g^p\, d\mu\right)^{1/p}$$

for all $x, y \in B$, where B is an arbitrary ball of radius r_0.
The constants in the theorem depend on p, q, h, C_d, and σ.

REMARKS 9.8. The theorem is not optimal:
(a) $6\sigma B$ can be reduced to $(1 + \varepsilon)\sigma B$. Under some geometric assumptions about the space one can further reduce $6\sigma B$ to B. It suffices to assume that the space is proper and that the distance between any two points equals infimum of lengths of curves connecting the points.

(b) In the case $0 < p < s$ one can prove that $u \in L_w^{p^*}(B)$, along with a corresponding estimate. This is more than just proving integrability for all $h < p^*$. If one makes an additional assumption that the pair (u, g) satisfies the so called truncation property, then one can even prove integrability with the exponent p^*.

(c) Assuming geometric property discussed in 1. and the truncation property discussed in 2. one obtains the optimal version of the Sobolev inequality, which, when applied to the Euclidean case gives the sharp version of the Sobolev embedding. Sharp, except for the constant because the "metric" approach is to general to give the sharp constant.

(d) Assuming connectivity of the space, one can improve inequality in the case $p = s > 1$ by proving the integrability of $\exp((...)^{s/(s-1)})$, known as the Trudinger inequality.

PROOF OF THEOREM 9.7 IN THE CASE $0 < p < s$. Fix a ball B of radius r. By Poincaré inequality (9.1) and Theorem 9.4 for $x, y \in \frac{6}{5}B$ we have

$$|u(x) - u(y)| \leq Cd(x, y) \left(\left(\mathcal{M}(g^p \chi_{6\sigma B})(x)\right)^{1/p} + \left(\mathcal{M}(g^p \chi_{6\sigma B})(y)\right)^{1/p} \right).$$

When $0 < \alpha < p$, weak type estimates and Lemma 9.6 yield $\left(\mathcal{M}(g^p \chi_{6\sigma B})\right)^{1/p} \in L^\alpha(\frac{6}{5}B)$ with

$$\left(\fint_{\frac{6}{5}B} \left(\mathcal{M}(g^p \chi_{6\sigma B})\right)^{\alpha/p} \right)^{1/\alpha} \leq C \left(\fint_{6\sigma B} g^p \, d\mu \right)^{1/p}.$$

Hence by Corollary 8.9,

$$\inf_{c \in \mathbb{R}} \left(\fint_B |u - c|^{\alpha^*} \, d\mu \right)^{1/\alpha^*} \leq Cr \left(\fint_{6\sigma B} g^p \, d\mu \right)^{1/p}.$$

Since $h = \alpha^* \uparrow p^*$ as $\alpha \uparrow p$, inequality (9.9) follows. Assume now that $q \in L_{loc}^q$, $p < q < s$, then $\left(\mathcal{M}(g^p \chi_{6\sigma B})\right)^{1/p} \in L^q(6\sigma B)$ with

$$\left(\fint_{\sigma B} \left(\mathcal{M}(g^p \chi_{6\sigma B})\right)^{q/p} \right)^{1/q} \leq C \left(\fint_{6\sigma B} g^q \, d\mu \right)^{1/q}$$

and again, inequality (9.10) follows from Corollary 8.9.   □

I believe that in the remaining two cases $p = s$ and $p > s$ a similar argument can be used. It would, however, require a closer look at the proof of Theorem 8.7 to see whether the weak L^p estimates that we obtain[55] are enough to conclude desired inequalities.

10. Abstract derivative and Sobolev spaces $H^{1,p}$.

In order to get a richer theory than the one developed in the previous sections, we need assume more about the space. Here we will assume the existence of an abstract operator that shares many properties with the derivative. This will lead to a nice characterization of the space $P^{1,p}$. In the next section we will see that the existence of such abstract derivative is a surprising consequence of another very natural assumption about the space, the condition of supporting Poincaré inequality.

---

[55]Instead of L^p estimates assumed in Theorem 8.7.

THEOREM 10.1. *Let (X, d, μ) be a metric-measure space equipped with a doubling measure and let N be a positive integer. Suppose that there is a constant $C > 0$ and a linear operator D which associates with each locally Lipschitz function u a measurable function $Du : X \to \mathbb{R}^N$ in such a way that*

(1) *If u is L-Lipschitz with $L \geq 1$, then $|Du| \leq CL$ μ-a.e.*
(2) *If u is locally Lipschitz and constant in an open set $\Omega \subset X$, then $Du = 0$ μ-a.e. in Ω.*

Let $H^{1,p}(X, d, \mu)$ be the Banach space defined as the closure of the set of locally Lipschitz functions with respect to the norm $\|u\| = \|u\|_{L^p} + \|Du\|_{L^p}$. Then $P^{1,p}(X, d, \mu) \cap L^p \subset H^{1,p}(X, d, \mu)$ for $1 \leq p < \infty$.

In general, there may be a problem with the definition of Du for a given $u \in H^{1,p}$. Namely, if u_k and v_k are two sequences of locally Lipschitz functions such that both converge to u in L^p, and $Du_k \to g$ in L^p, $Dv_k \to h$ in L^p, $g \neq h$, then (u, g) and (u, h) represent two different elements in $H^{1,p}$.[56] In such a case we say that the gradient is not uniquely determined. Fortunately, for a reasonable class of spaces we have uniqueness of the gradient.

We say that the *uniqueness of the gradient* holds if for every sequence of locally Lipschitz functions, u_n, such that $u_n \to 0$ in L^p and $Du_n \to g$ in L^p, we have $g = 0$ a.e. In such a case we can associate with each $u \in H^{1,p}$ a unique Du obtained by taking the limit of 'gradients' for any approximating sequence of locally Lipschitz functions.

THEOREM 10.2. *Let (X, d, μ) be a metric-measure space equipped with a doubling measure and let N be a positive integer. Suppose that there is a constant $C > 0$ and a linear operator D which associates with each locally Lipschitz function u a measurable function $Du : X \to \mathbb{R}^N$ in such a way that*

(1) *If u is L-Lipschitz with $L \geq 1$, then $|Du| \leq CL$ μ-a.e.*
(2) *If u is locally Lipschitz and constant in a measurable set $E \subset X$, then $Du = 0$ μ-a.e. in E.*

Let $1 \leq p < \infty$ and $\sigma \geq 1$. Assume that for every locally Lipschitz function u, the pair $(u, |Du|)$ satisfies the p-Poincaré inequality (9.1). Define $H^{1,p}(X, d, \mu)$ as in Theorem 10.1. Then $H^{1,p}(X, d, \mu) = P^{1,p}(X, d, \mu) \cap L^p$, the uniqueness of the gradient holds and $|Du| \leq C'g$ a.e., whenever (u, g) satisfies the p-Poincaré inequality.[57] Moreover for $1 < p < \infty$ the space $H^{1,p}(X, d, \mu)$ is reflexive.

PROOF OF THEOREM 10.1 FOR $p > 1$. Assume that $u \in P^{1,p} \cap L^p$ i.e., there exists $0 \leq g \in L^p$ and $\sigma \geq 1$ such that the p-Poincaré inequality (9.1) holds true.

We need construct a sequence of locally Lipschitz functions approximating u. To this end fix $\varepsilon > 0$ and choose $\{\widetilde{B}_i\}$ to be a maximal disjoint family of balls in X of radius $\varepsilon/4$. Then the family $\{2\widetilde{B}_i\}$ forms a covering of X. Set $B_i = 4\widetilde{B}_i$. The doubling property implies that there is a constant C such that no point of X belongs to more than C balls B_i.[58] Now we construct a Lipschitz partition of unity

---

[56]Hence the embedding $P^{1,p} \cap L^p \subset H^{1,p}$ has to be properly understood: for every $u \in P^{1,p} \cap L^p$ there is a Cauchy sequence with respect to the $H^{1,p}$ norm of locally Lipschitz functions that converge to u in L^p.

[57]Here the pair (u, g) may satisfy the p-Poincaré inequality (9.1) with σ replaced by another constant $\tau \geq 1$.

[58]Cf. the proof of Lemma 4.3.

$\{\varphi_i\}$ subordinated to the given family of balls $\{B_i\}$, i.e., $\sum_i \varphi_i \equiv 1$, $0 \le \varphi_i \le 1$, supp $\varphi_i \subset B_i$ and all the functions φ_i are $C\varepsilon^{-1}$-Lipschitz. To this end it is enough to choose $\varphi_i(x) = \psi(d(x_i, x)/\varepsilon)(\sum_k \psi(d(x_k, x)/\varepsilon))^{-1}$, where ψ is a smooth function, $\psi \equiv 1$ on $[0, 1/2]$, $\psi \equiv 0$ on $[3/4, \infty)$, $0 \le \psi \le 1$, and x_i is the center of B_i for $i = 1, 2, \ldots$ We put $u_\varepsilon = \sum_i \varphi_i u_{B_i}$.

Now we show that $u_\varepsilon \to u$ in $L^p(X)$ as $\varepsilon \to 0$. According to Theorem 9.7
(10.1)
$$\left(\fint_B |u - u_B|^p \, d\mu \right)^{1/p} \le Cr \left(\fint_{6\sigma B} g^p \, d\mu \right)^{1/p} \quad \text{on every ball } B \text{ of radius } r.$$

Using the fact that the sum in the definition of u_ε has uniformly bounded number of nonzero terms at a given point we obtain

$$|u - u_\varepsilon|^p \le C \sum_i |\varphi_i|^p |u - u_{B_i}|^p \le C \sum_i |u - u_{B_i}|^p \chi_{B_i}.$$

Hence the p-Poincaré inequality (10.1) and the uniform bound for overlapping of the balls $6\sigma B_i$ yield

$$\int_X |u - u_\varepsilon|^p \, d\mu \le C \sum_i \int_{B_i} |u - u_{B_i}|^p \, d\mu \le C\varepsilon^p \sum_i \int_{6\sigma B_i} g^p \, d\mu \le C\varepsilon^p \int_X g^p \, d\mu \,.$$

This proves that $u_\varepsilon \in L^p(X)$ and $u_\varepsilon \to u$ in $L^p(X)$ as $\varepsilon \to 0$. We will also need the following

LEMMA 10.3. *Assume that the pair (u, g) is as above. Then*

$$|Du_\varepsilon| \le C \sum_i \left(\fint_{3\sigma B_i} g^p \, d\mu \right)^{1/p} \chi_{B_i} \,.$$

PROOF. Fix a ball in the covering $\{B_i\}$ and denote it by B_0. Then for a.e. $x \in B_0$ we have

$$|Du_\varepsilon(x)| = |D(u_\varepsilon - u_{B_0})(x)| \le \sum_{i:\, B_0 \cap B_i \ne \emptyset} |D\varphi_i(x)||u_{B_i} - u_{B_0}|$$

$$\le \quad C\varepsilon^{-1} \sum_{i:\, B_0 \cap B_i \ne \emptyset} \fint_{B_i} \fint_{B_0} |u(y) - u(z)| \, d\mu(y) \, d\mu(z)$$

$$\le \quad C\varepsilon^{-1} \fint_{3B_0} |u(y) - u_{3B_0}| \, d\mu(y) \le C \left(\fint_{3\sigma B_0} g^p \, d\mu \right)^{1/p} \,.$$

In the proof of the first inequality we used finite additivity of D and property 2. of D. The proof of the lemma is complete. $\square$

It follows from the lemma that $\sup_\varepsilon \int_X |Du_\varepsilon|^p \le C \int_X g^p$. Take a sequence $\varepsilon_n \to 0$. Since $p > 1$, reflexivity of L^p implies[59] that there is a subsequence of $(Du_{\varepsilon_n})_n$ weakly convergent in L^p.[60] Now, by Mazur's lemma a sequence of convex combinations of u_{ε_n} converges in the norm of $H^{1,p}$. Since it converges to u in L^p we conclude that $u \in H^{1,p}$. $\square$

---

[59]Every bounded sequence of a reflexive space contains a weakly convergent subsequence.

[60]This is the only moment when we use the assumption $p > 1$. The existence of a weakly convergent subsequence is true for $p = 1$ as well, but the proof is more difficult as the space L^1 is not reflexive.

PROOF OF THEOREM 10.2. First we prove the uniqueness of the gradient. Let u_n be a sequence of locally Lipschitz functions such that $u_n \to 0$ in L^p and $Du_n \to g$ in L^p. We have to prove that $g = 0$ a.e. By selecting a subsequence we may assume that

$$(10.2) \qquad \int_X \left(|u_{n+1} - u_n|^p + |Du_{n+1} - Du_n|^p \right) d\mu \leq 10^{-np}.$$

This implies that the sequences (u_n) and (Du_n) converge a.e. Let $u_{n+1} - u_n = v_n$. Since by the assumption the pair $(v_n, |Dv_n|)$ satisfies the p-Poincaré inequality, Theorem 9.4 gives

$$|(u_{n+1} - u_n)(x) - (u_{n+1} - u_n)(y)| \leq Cd(x,y) \left((\mathcal{M}|Dv_n|^p(x))^{1/p} + (\mathcal{M}|Dv_n|^p(y))^{1/p} \right).$$

Hence for $\ell \geq k \geq k_0$

$$|(u_\ell - u_k)(x) - (u_\ell - u_k)(y)| \leq Cd(x,y) \left(g_{k_0}(x) + g_{k_0}(y) \right),$$

where

$$g_{k_0}(x) = \sum_{n=k_0}^{\infty} \left(\mathcal{M}|Dv_n|^p(x) \right)^{1/p}.$$

Taking the limit as $\ell \to \infty$ we obtain

$$(10.3) \qquad |u_k(x) - u_k(y)| \leq Cd(x,y)(g_{k_0}(x) + g_{k_0}(y)),$$

for all $k \geq k_0$ and almost every x and y. Now we estimate size of the level sets of the function g_{k_0}:

$$\mu(\{g_{k_0} > t\}) \leq \sum_{n=k_0}^{\infty} \mu \left(\left\{ (\mathcal{M}|Dv_n|^p)^{1/p} > \frac{t}{2^{n-k_0+1}} \right\} \right)$$

$$\leq \sum_{n=k_0}^{\infty} C \frac{2^{(n-k_0+1)p}}{t^p} \int_X |Dv_n|^p \, d\mu \leq C' t^{-p} 10^{-k_0 p}.$$

In the middle inequality we used the weak type estimate for the maximal function, Theorem 4.8, while the last inequality follows from (10.2).

Let $E_{k_0,t} = \{g_{k_0} > t\}$. Note that (10.3) implies that $u_k|_{X \setminus E_{k_0,t}}$ is Ct-Lipschitz.

Observe now that if u is locally Lipschitz and $u|_F$ is L-Lipschitz, then $|Du| \leq CL$ almost everywhere in F. Indeed, $u|_F$ can be extended to a globally L-Lipschitz function $\bar{u}$ on X.[61] Hence $|D\bar{u}| \leq CL$ a.e. Since $u - \bar{u} = 0$ in F, then $|D(u - \bar{u})| = 0$ a.e. in F and thus $|Du| \leq CL$ a.e. in F.

Returning to the theorem we get $|Du_k| \leq Ct$ a.e. in $X \setminus E_{k_0,t}$, and hence $|g| \leq Ct$ a.e. in $X \setminus E_{k_0,t}$. Thus $\mu(\{|g| > Ct\}) \leq \mu(E_{k_0,t}) \to 0$ as $k_0 \to \infty$. Since $t > 0$ can be arbitrarily small we conclude that $g = 0$ a.e. and the uniqueness of the gradient follows.

By Theorem 10.1 we know that $P^{1,p}(X,d,\mu) \cap L^p \subset H^{1,p}(X,d,\mu)$. The converse inclusion follows from the definition of $H^{1,p}(X,d,\mu)$ and the fact that the p-Poincaré inequality holds for locally Lipschitz functions.

Now we will prove that if the pair (u,g) satisfies p-Poincaré inequality, then $|Du| \leq Cg$ a.e.

As in the proof of Theorem 10.1 we find a sequence u_{ε_n} of locally Lipschitz functions such that $u_{\varepsilon_n} \to u$ in L^p and Du_{ε_n} is weakly convergent in L^p. Then

---

[61]McShane's lemma states that every L-Lipschitz function defined on a subset of a metric space can be extended to the entire space as a L-Lipschitz function.

by Mazur's lemma a sequence of convex combinations of u_{ε_n} converges to u in the norm of $H^{1,p}$.

It follows from Lemma 10.3 that $\limsup_{n\to\infty} |Du_{\varepsilon_n}| \le Cg$ a.e. Indeed, this inequality is fairly easily seen to hold at the Lebesgue points of g^p. Since convex combinations of Du_{ε_n} converge to Du in L^p, we conclude that $|Du| \le Cg$ a.e.

Finally reflexivity of the space $H^{1,p}$ follows from the fact that this space is isomorphic to a closed subspace of $L^p(X, \mathbb{R}^{N+1})$. The isomorphism is given by $u \mapsto (u, Du)$. The proof is complete. $\qquad\square$

In a general setting of metric spaces there is a natural operator $u \mapsto |\nabla^+ u| = \operatorname{Lip} u$ which assigns an upper gradient to each locally Lipschitz function u, see Lemma 6.7. Unfortunately this operator is not linear so Theorems 10.1 and 10.2 do not apply. The operator is, however, easily seen to be sublinear[62] and hence we need a version of the theorems valid for sublinear operators.

The reason why the proof of Theorem 10.1 does not cover the case of the operator $|\nabla^+ u|$ is the lack of linearity by which one cannot use Mazur's lemma as in the last step of the proof of Theorem 10.1 to convert a sequence $u_k \xrightarrow{L^p} u$ with $|\nabla^+ u_k|$ weakly convergent in L^p into a Cauchy sequence. Therefore in order to generalize Theorem 10.1 to the case of sublinear operators we need assume that the Sobolev type space that we consider is a function space, closed under a kind of weak convergence.

Clearly the operator $|\nabla^+ u|$ satisfies assumptions of the following theorem.

THEOREM 10.4. *Let (X, d, μ) be a metric-measure space equipped with a doubling measure. Suppose that there is a constant $C > 0$ and an operator $|D|$ which associates with each locally Lipschitz function u a nonnegative function $|D|u : X \to [0, \infty)$ in such a way that*

(1) *$|D|(u+v) \le C(|D|u + |D|v)$ and $|D|(\lambda u) \le C|\lambda||D|u$ a.e. whenever u, v are locally Lipschitz and λ is a real.*
(2) *If u is L-Lipschitz, then $|D|u \le CL$ a.e.*
(3) *If u is locally Lipschitz and constant on an open set $\Omega \subset X$, then $|D|u = 0$ a.e. in Ω.*

Assume that $W^{1,p}(X, d, \mu)$, $1 \le p < \infty$ is a function space equipped with a norm $\|\cdot\|$ and with the following properties

(1) *If $u \in L^p$ is locally Lipschitz and such that $|D|u \in L^p$ then $u \in W^{1,p}(X, d, \mu)$ and $\|u\| \le C(\|u\|_p + \||D|u\|_p)$.*
(2) *If $u_k \in W^{1,p}(X, d, \mu) \cap \operatorname{Lip}_{\operatorname{loc}}$ converges in L^p to w and $|D|u_k$ is a sequence weakly convergent in L^p then $w \in W^{1,p}(X, d, \mu)$.*

Then $P^{1,p}(X, d, \mu) \cap L^p \subset W^{1,p}(X, d, \mu)$.

As the proof is almost the same as that for Theorem 10.1, we leave it to the reader.

The above result together with Lemma 7.8 implies that if the measure is doubling, then $P^{1,p} \cap L^p \subset N^{1,p}$ for $1 \le p < \infty$. This and (9.2) yield the following inclusions.[63]

---

[62]An operator T is sublinear if $|T(x+y)| \le |Tx| + |Ty|$. Sublinearity is the reason for which we use here $|\nabla^+ u|$ instead of $\operatorname{lip} u$.

[63]Cf. Theorem 8.6.

COROLLARY 10.5. *If the measure μ is doubling, then $M^{1,p}(X, d, \mu) \subset P^{1,p}(X, d, \mu) \cap L^p \subset N^{1,p}(X, d, \mu)$ for all $1 \le p < \infty$.*

11. Spaces supporting Poincaré inequality

In this section we will *always* assume that (X, d, μ) is a complete metric space equipped with a doubling measure.

It is natural to look for restrictions on the class of metric spaces equipped with a doubling measure so that: (1) the additional conditions are formulated in metric-measure terms; (2) the Euclidean space and Riemannian manifolds are among examples; (3) there are many examples far from being Euclidean-like; (4) the theory becomes much deeper and much more beautiful, than the one developed for general metric-measure spaces equipped with doubling measures.

The assumption about the existence of a linear operator D like in Theorems 10.1 and 10.2 does not really meet the standards as it is not formulated in metric-measure terms. However, the class of spaces supporting p-Poincaré inequalities, discussed below, meets all the standards. Surprisingly, for these spaces one can prove the existence of a linear operators of differentiation D as in Theorem 10.2.

DEFINITION 11.1. We say that (X, d, μ) *supports a p-Poincaré inequality*, $1 \le p < \infty$, if there exist constants C_P and $\sigma \ge 1$ such that for every Borel measurable function $u : X \to \mathbb{R}$ and every upper gradient $g : X \to [0, \infty]$ of u, the pair (u, g) satisfies the family of p-Poincaré inequalities[64]

$$(11.1) \qquad \fint_B |u - u_B| \, d\mu \le C_P r \left(\fint_{\sigma B} g^p \, d\mu \right)^{1/p} \qquad \text{on every ball } B \text{ of radius } r.$$

Clearly Euclidean space supports p-Poincaré inequalities for all $p \ge 1$. Also compact Riemannian manifolds and complete Riemannian manifolds of nonnegative Ricci curvature are amongst examples. There are, however, many important examples which have nothing to do with the Euclidean structure; for references, see Section 12.

Directly from the definition we have $N^{1,p} \subset P^{1,p} \cap L^p$. This and Corollary 10.5 gives

THEOREM 11.2. *If the space supports the p-Poincaré inequality, $1 \le p < \infty$, then $N^{1,p}(X, d, \mu) = P^{1,p}(X, d, \mu) \cap L^p$.*

Thus for spaces supporting p-Poincaré inequalities we have $M^{1,p} \subset P^{1,p} \cap L^p = N^{1,p}$. If $p = 1$ we have already noted that there is no equality between $N^{1,1}$ and $M^{1,1}$.[65] At the same time we have equality of the spaces $M^{1,p} = P^{1,p} \cap L^p = N^{1,p}$ in the Euclidean case when $p > 1$. This result generalizes to the setting of spaces supporting Poincaré inequalities as follows.

THEOREM 11.3. *If $1 < p < \infty$ and the space supports the q-Poincaré inequality for some $1 \le q < p$, then $M^{1,p} = P^{1,p} \cap L^p = N^{1,p}$.*

This covers the Euclidean case as $\mathbb{R}^n$ supports 1-Poincaré inequality.

---

[64]There is a tiny difference between (9.1) and (11.1), namely the existence of the constant C_P in (11.1).

[65]See an exercise proceeding Theorem 8.3. In particular $M^{1,1} \ne P^{1,1} \cap L^1$.

The definition of the space supporting the p-Poincaré inequality is very natural, but not so easy as one needs develop theory of upper gradients in order to formulate the definition. One could think about a more elementary approach involving Lipschitz functions and some particular upper gradients like $|\nabla^+ u|$. This is possible as the following result shows.

THEOREM 11.4. *A space (X, d, μ) supports the p-Poincaré inequality, $1 \leq p < \infty$, if and only if there exist constants $C > 0$, $\sigma \geq 1$ such that for every compactly supported Lipschitz function u,*

$$\fint_B |u - u_B| \, d\mu \leq Cr \left(\fint_{\sigma B} |\nabla^+ u|^p \, d\mu \right)^{1/p} \quad \text{on every ball } B \text{ of radius } r.$$

Note that the space supporting the p-Poincaré inequality must contain a lot of nonconstant rectifiable paths. Indeed, if the only rectifiable paths are constant, then clearly, $g \equiv 0$ is an upper gradient of any Borel function and hence the p-Poincaré inequality cannot be satisfied. This can be turned into a quantitative form as follows.

THEOREM 11.5. *A space (X, d, μ) supporting p-Poincaré inequality for some $1 \leq p < \infty$, is quasiconvex, i.e. there is a constant $C \geq 1$ such that any two points x and y in the space can be connected by a rectifiable path γ with the length bounded by $\ell(\gamma) \leq Cd(x, y)$.*

Perhaps the most surprising results concerning spaces supporting Poincaré inequalities are those about the existence of a "metric differentiable structure". A weak version of such a result can be formulated as follows.

THEOREM 11.6. *Suppose that the space (X, d, μ) supports the p-Poincaré inequality $1 < p < \infty$. Then there is a positive integer N, a constant $C \geq 1$ and a linear operator D which associates with each locally Lipschitz function u a measurable function $Du : X \to \mathbb{R}^N$ such that*

    (1) *If u is L-Lipschitz, then $|Du| \leq CL$, μ-a.e.*
    (2) *If u is locally Lipschitz and constant on a measurable set $E \subset X$, then $Du = 0$, μ-a.e. in E.*
    (3) *For locally Lipschitz functions u and v, $D(uv) = uDv + vDu$.*
    (4) *For each locally Lipschitz function u, $\text{Lip}\, u \leq |Du| \leq C \lip u$[66] and hence the pair $(u, |Du|)$ satisfies the p-Poincaré inequality.*

Note that, in particular, the operator D satisfies the assumptions of Theorem 10.2. Actually there are stronger versions of Theorem 11.6 but they are slightly more difficult to state.

As a consequence of Theorems 11.6, 10.2 and 11.3 we obtain

COROLLARY 11.7. *If the space (X, d, μ) supports the p-Poincaré inequality, $1 < p < \infty$, then $N^{1,p}(X, d, \mu)$ is reflexive. If, in addition, the space supports the q-Poincaré inequality for some $1 \leq q < p$, then $M^{1,p}(X, d, \mu)$ is reflexive too.*

---

[66]In particular $\lip u \approx \text{Lip}\, u$ which is also a consequence of supporting the p-Poincaré inequality. Recall that $\text{Lip}\, u = |\nabla^+ u|$.

12. Historical notes

The references below are by no means complete and many important contributions are omitted.

There are several excellent textbooks on Sobolev spaces. The one that seems most appropriate as an introduction to the paper is the book by Evans and Gariepy [23]. Theorems 2.1 and 2.2 are relatively new, while Theorem 2.3 is classical [23]. Theorem 2.2 was proved by Hajłasz [34]. Equivalence of conditions 1. and 2. in Theorem 2.1 is due to Calderón [12] for $p > 1$ and due to Franchi, Hajłasz and Koskela [26] when $p = 1$. The equivalence of 1. and 3. is due to Koskela and MacManus [50] when $p > 1$ and is proved in [26] in the general case $p \geq 1$. The equivalence of 3. and 4. for $p > 1$ was proved by Hajłasz and Koskela [37] and Heinonen and Koskela [42]. The case $p = 1$ was settled by Hajłasz [35]; it was used there to prove that 4. characterizes $W^{1,1}$ for $p = 1$.

The theory of rectifiable curves in metric spaces is a mathematical folklore, see also Väisälä [65]. Theorem 3.9 was proved by Busemann [10], but, perhaps, it is older. It seems that only notion of the speed and Theorem 3.6 are more recent. They both appear in books by Ambrosio and Tilli [4] and Burago, Burago and Iwanov [9]. Our proof follows [4].

The approach to Borel-regular measures presented here is standard and follows Federer [24]. Theorem 4.1 was proved in [24]. Coifman and Weiss [19] were the first to realize how important doubling measures are for the development of harmonic analysis on metric spaces. Many examples of doubling measures can be found in Stein's book [63]. Theorem 4.6 is standard, see Semmes [59, Lemma C.3] or Heinonen [40, Exercise 8.11]. Theorem 4.5 is due to Volberg and Konyagin [66] in the case of a compact metric space. The general case of a complete metric space has been established by Luukkainen and Saksman [54]. In the case of the Lebesgue measure in $\mathbb{R}^n$ Theorems 4.8 and 4.9 are proved e.g. in Stein [62]. The proofs easily generalize to general doubling measures.

The concept of the modulus of the path family was introduced by Ahlfors and Beurling [1] and developed by Fuglede [30]. This theory constitutes a basic tool for the theory of quasiregular mappings, see Rickman [55] and Väisälä [65]. Although all this was developed in Euclidean or Riemannian setting, the extension to the metric space case is straightforward.

The notion of upper gradient, as in Section 6, has beed introduced by Heinonen and Koskela [42]. All the results in Section 6 are standard by now.

Being motivated by the work of Heinonen and Koskela [42], Shanmugalingam [61] introduced the $N^{1,p}$ spaces and Cheeger [15] introduced, independently, spaces denoted here by $C^{1,p}$. Theorem 7.19 is due to Shanmugalingam [61]. Most of the results discussed in Section 7 are proved in [61], although the presentation and some of the proofs are somewhat different here. Theorem 7.16 for $p > 1$ was proved by Cheeger [15]; the case $p = 1$ is new.

The $M^{1,p}$ spaces have been introduced by Hajłasz [34] (and denoted by $W^{1,p}$). It seems it was the first attempt to define a Sobolev type space in the setting of general metric-measure spaces. For references to papers involving $M^{1,p}$ spaces, see Hajłasz [35]. It was assumed in [34] that $p > 1$ and that the space is of bounded diameter, but the extension to the case of $p > 0$ and unbounded spaces is, in many instances, straightforward. In particular proofs of Theorems 8.3 and 8.4 follow exactly the same argument as in [34]. Theorem 8.5 and its generalizations to other

Cantor-type self-similar sets are due to Rissanen [**56**]. Theorem 8.6 appears in Shanmugalingam [**61**]; the proof presented here is a new one. Theorem 8.7 is new. In the case in which $p > 1$ and X is of bounded diameter a global version of the theorem with integrals over entire X has been proved in [**34**]. The same method as in [**34**] easily extends to the general case $p > 0$, however, the local version with integrals over balls, as stated in Theorem 8.7, is much harder. It was assumed in [**34**] that $\mu(B(x,r)) \geq br^s$ for every x and all $0 < r < \operatorname{diam} X$. If we would like to apply the argument from [**34**] to the case considered in Theorem 8.7 we would have to assume that for every $x \in B_0$ and $r < r_0$, $\mu(B(x,r) \cap B_0) \geq br^s$, which is much more than the $V(\sigma B_0, s, b)$ condition. Even if $\operatorname{diam} X < \infty$, the Hölder continuity of $u \in M^{1,p}$ for $p > s$ is new. Indeed, to prove it we need local estimates over balls and global estimate proved in [**34**] is not enough (it implies boundedness of u only).

The $P^{1,p}$ spaces have been introduced by Hajłasz and Koskela in [**36**] (cf. Franchi, Lu and Wheeden [**28**]) and developed in [**37**]. The case of $s/(s+1) \leq p < 1$ of Theorem 9.2 is new. Theorem 9.4 and Theorem 9.3 in the case $p > 1$ are due to Hajłasz and Koskela [**37**]. Theorem 9.5 was proved by Hajłasz and Koskela [**37**] and Heinonen and Koskela [**42**] when $p > 1$. The case $p = 1$ was proved by Hajłasz [**35**] and the case $s/(s+1) < p < 1$ is new. Theorem 9.7 and its improvements mentioned in Remarks 9.8 are due to Hajłasz and Koskela [**37**], but the proof presented here is new.

Section 10 is based on [**26**]. Only Corollary 10.5 is new. Although this corollary is a straightforward consequence of results in [**26**] it seems that this observation has beed overlooked.

Spaces supporting Poincaré inequalities have been introduced in Heinonen and Koskela [**42**]. They developed theory of quasiconformal mappings between such spaces, see Heinonen, Koskela, Shanmugalingam and Tyson [**43**] for further development. The existing examples of spaces supporting the p-Poincaré inequality make the theory very important. These are: Euclidean space (obvious), compact Riemannian manifolds (easy), complete Riemannian manifolds of nonnegative Ricci curvature (a difficult result essentially due to Buser [**11**] with a recent direct proof by Saloff-Coste [**58**]), Carnot-Carathéodory spaces (Franchi and Lanconelli [**27**], Jerison [**45**], Franchi, Gutiérrez and Wheeden [**25**], Garofalo and Nhieu [**31**], Hajłasz and Koskela [**37**], Lanconelli and Morbidelli [**52**]). Some classes of topological manifolds (Semmes [**59**]). Boundaries of hyperbolic buildings (Bourdon and Pajot [**8**]) and some other exotic examples (Laakso [**51**] and Heinonen and Hanson [**41**]). Theorem 11.2 seems new. Theorem 11.3 is due to Koskela and MacManus [**50**]. Theorem 11.4 is due to Keith [**48**]. Theorem 11.5 was proved in Cheeger [**15**] and in Hajłasz and Koskela [**37**]. Amazing Theorem 11.6 and Corollary 11.7 are due to Cheeger [**15**]; for generalizations see Keith [**49**].

Other papers related to what was discussed here include Ambrosio [**2**] (BV spaces on metric spaces), Ambrosio and Kirchheim [**3**] (currents in metric spaces), Kinnunen and Martio [**47**] (capacities in metric spaces), Liu, Lu and Wheeden [**53**] (Sobolev spaces on metric spaces with higher order derivatives), Björn, Macmanus and Shanmugalingam [**7**] (potential theory on metric spaces), Goldshtein and Troyanov [**33**] (axiomatic theory of Sobolev spaces), Biroli and Mosco [**6**] (Dirichlet forms), Grigor'yan, Hu and Lau [**32**] and Sturm [**64**] (Heat kernel estimates), Delmotte [**22**], Holopainen and Soardi [**44**] and Woess [**67**] (analysis on graphs), Capogna and Garofalo [**13**], Capogna, Garofalo and Nhieu [**14**], Chernikov and

Vodop'yanov [**16**], Citti, Lanconelli and Montanari [**17**], Hajłasz and Strzelecki [**39**] (nonlinear sub-elliptic equations), Danielli, Garofalo and Nhieu, [**21**], Cohn and Lu [**18**], (Sobolev spaces on Carnot–Carathéodory spaces), Franchi, Serapioni and Serra-Casano [**29**] (Finite perimeter in the Heisenberg group), Coulhon, Holopainen and Saloff-Coste [**20**], Saloff-Coste, [**57**] (Harnack's inequality), Bakry, Coulhon, Ledoux and Saloff-Coste [**5**] (abstract approach to Sobolev inequalities), Hajłasz and Martio [**38**] and Kigami [**46**] (analysis on fractals).

Recent books on analysis on metric spaces include Ambrosio and Tilli [**4**], Heinonen [**40**] and Semmes [**60**].

References

[1] L. Ahlfors, A. Beurling, A., *Conformal invariants and function-theoretic null-sets,* Acta Math. **83** (1950), 101–129.

[2] L. Ambrosio, *Some fine properties of sets of finite perimeter in Ahlfors regular metric measure spaces,* Adv. Math. **159** (2001), 51–67.

[3] L. Ambrosio, B. Kirchheim, *Currents in metric spaces,* Acta Math. **185** (2000), 1–80.

[4] L. Ambrosio, P. Tilli, *Selected topics on "Analysis on metric spaces".* Scuola Normale Superiore, Pisa 2000.

[5] D. Bakry, T. Coulhon, M. Ledoux, L. Saloff-Coste, *Sobolev inequalities in disguise,* Indiana Univ. Math. J. **44** (1995), 1033–1074.

[6] M. Biroli, U. Mosco, *Sobolev inequalities on homogeneous spaces,* Pot. Anal. **4** (1995), 311–324.

[7] J. Björn, P. MacManus, N. Shanmugalingam, *Fat sets and pointwise boundary estimates for p-harmonic functions in metric spaces,* J. Anal. Math. **85** (2001), 339–369.

[8] M. Bourdon, H. Pajot, *Poincaré inequalities and quasiconformal structure on the boundary of some hyperbolic buildings,* Proc. Amer. Math. Soc. **127** (1999), 2315–2324.

[9] D. Burago, Y. Burago, S. Ivanov, *A course in metric geometry.* Graduate Studies in Mathematics, 33. American Mathematical Society, Providence, RI, 2001.

[10] H. Busemann, *The geometry of geodesics,* Academic Press, New York, 1955.

[11] P. Buser, *A note on the isoperimetric constant,* Ann. Sc. Ecole Norm. Sup. **15** (1982), 213–230.

[12] A. P. Calder"on, *Estimates for singular integral operators in terms of maximal functions,* Studia Math. **44** (1972), 563–582.

[13] L. Capogna, N. Garofalo, *Regularity of minimizers of the calculus of variations in Carnot groups via hypoellipticity of systems of Hörmander type.* J. Eur. Math. Soc. **5** (2003), 1–40.

[14] L. Capogna, N. Garofalo, D.-M. Nhieu, *Properties of harmonic measures in the Dirichlet problem for nilpotent Lie groups of Heisenberg type,* Amer. J. Math. **124** (2002), 273–306.

[15] J. Cheeger, *Differentiability of Lipschitz functions on metric measure spaces,* Geom. Funct. Anal. **9** (1999), 428–517.

[16] V. M. Chernikov, S. K. Vodop'yanov, *Sobolev Spaces and hypoelliptic equations I,II,* Siberian Advances in Mathematics. **6** (1996) no. 3, 27–67, no. 4, 64–96. Translation from: Trudy In-ta matematiki RAN. Sib. otd-nie. **29** (1995), 7–62

[17] G. Citti, E. Lanconelli, A. Montanari, *Smoothness of Lipchitz-continuous graphs with non-vanishing Levi curvature,* Acta Math. **188** (2002), 87–128.

[18] W. S.Cohn, G. Lu, *Best constants for Moser-Trudinger inequalities on the Heisenberg group.* Indiana Univ. Math. J. **50** (2001), 1567–1591.

[19] R. Coifman, G. Weiss, *Extensions of Hardy spaces and their use in analysis,* Bull. Amer. Math. Soc. **83** (1977), 569–645.

[20] T. Coulhon, I. Holopainen, L. Saloff-Coste, *Harnack inequality and hyperbolicity for subelliptic p-Laplacians with applications to Picard type theorems,* Geom. Funct. Anal. **11** (2001), 1139–1191.

[21] D. Danielli, N. Garofalo, D. M. Nhieu, *Non-doubling Ahlfors measures, perimeter measures, and the characterization of the trace spaces of Sobolev functions in Carnot-Carathéodory spaces,* Memoirs Amer. Math. Soc., to appear.

[22] T. Delmotte, *Inégalité de Harnack elliptique sur les graphes,* Colloq. Math. **72** (1997), 19–37.

[23] L. C. Evans, R. F. Gariepy, *Measure Theory and Fine Properties of Functions.* Studies in Advanced Mathematics, 1992.

[24] H. Federer, *Geometric measure theory,* Springer, 1969.

[25] B. Franchi, C. E. Gutiérrez , R. L. Wheeden, *Weighted Sobolev–Poincaré inequalities for Grushin type operators,* Comm. Partial Differential Equations **19** (1994), 523–604.

[26] B. Franchi, P. Hajłasz, P. Koskela, *Definitions of Sobolev classes on metric spaces,* Ann. Inst. Fourier (Grenoble) **49** (1999), 1903–1924.

[27] B. Franchi, E. Lanconelli, *Hölder regularity theorem for a class of non uniformly elliptic operators with measurable coefficients, Ann. Scuola Norm. Sup. Pisa* **10** (1983), 523–541.

[28] B. Franchi, G. Lu, R. L. Wheeden, *Representation formulas and weighted Poincaré inequalities for Hörmander vector fields,* Int. Mat. Res. Notices **1996,** no. 1, 1–14.

[29] B. Franchi, R. Serapioni, F. Serra Cassano, *Rectifiability and perimeter in the Heisenberg group,* Math. Ann. **321** (2001), 479–531.

[30] B. Fuglede, *Extremal length and functional completion,* Acta Math. **98** (1957), 171–219.

[31] N. Garofalo, D.-M. Nhieu, *Isoperimetric and Sobolev inequalities for Carnot-Carathéodory spaces and the existence of minimal surfaces,* Comm. Pure Appl. Math. **49** (1996), 1081–1144.

[32] A. Grigor'yan, J. Hu, H.-S. Lau, *Heat kernels on metric measure spaces and an application to semilinear elliptic equations,* Trans. Amer. Math. Soc. **355** (2003), 2065-2095.

[33] V. Gol'dshtein, M. Troyanov, *Axiomatic theory of Sobolev spaces,* Expo. Math. **19** (2001), 289–336.

[34] P. Hajłasz, *Sobolev spaces on an arbitrary metric space,* Potential Anal. **5** (1996), 403–415.

[35] P. Hajłasz, *A new characterization of the Sobolev space,* Studia Math. to appear.

[36] P. Hajłasz, P. Koskela, *Sobolev meets Poincaré,* C. R. Acad. Sci. Paris Ser. I Math. **320** (1995), 1211–1215.

[37] P. Hajłasz, P. Koskela, *Sobolev met Poincaré,* Memoirs Amer. Math. Soc. **145** (2000), no. 688, x+101 pp.

[38] P. Hajłasz, O. Martio, *Traces of Sobolev functions on fractal type sets and characterization of extension domains,* J. Funct. Anal. **143** (1997), 221–246.

[39] P. Hajłasz, P. Strzelecki, *Subelliptic p-harmonic maps into spheres and the ghost of Hardy spaces,* Math. Ann. **312** (1998), 341–362.

[40] J. Heinonen, *Lectures on analysis on metric spaces,* Universitext. Springer-Verlag, New York, 2001.

[41] B. Hanson, J. Heinonen, *An n-dimensional space that admits a Poincaré inequality but has no manifold points,* Proc. Amer. Math. Soc. **128** (2000), 3379–3390.

[42] J. Heinonen, P. Koskela, *Quasiconformal maps on metric spaces with controlled geometry,* Acta Math. **181** (1998), 1–61.

[43] J. Heinonen, P. Koskela, N. Shanmugalingam, J. T. Tyson, *Sobolev classes of Banach space-valued functions and quasiconformal mappings,* J. Anal. Math. **85** (2001), 87–139.

[44] I. Holopainen, P. Soardi, *p-harmonic functions on graphs and manifolds,* Manuscripta Math. **94** (1997), 95–110.

[45] D. Jerison, *The Poincaré inequality for vector fields satisfying Hörmander's condition,* Duke Math. J. **53** (1986), 503–523.

[46] J. Kigami, *Analysis on fractals.* Cambridge Tracts in Mathematics, 143. Cambridge University Press, Cambridge, 2001.

[47] J. Kinnunen, O. Martio, *The Sobolev capacity on metric spaces.* Ann. Acad. Sci. Fenn. Math. **21** (1996), 367–382.

[48] S. Keith, *Modulus and the Poincaré inequality on metric measure spaces,* Math. Z. to appear.

[49] S. Keith, *A differentiable structure for metric measure spaces,* Adv. Math. to appear.

[50] P. Koskela, P.MacManus, *Quasiconformal mappings and Sobolev spaces,* Studia Math. **131** (1998), 1–17.

[51] T. J. Laakso, *Ahlfors Q-regular spaces with arbitrary Q > 1 admitting weak Poincaré inequality,* Geom. Funct. Anal. **10** (2000), 111–123.

[52] E. Lanconelli, D. Morbidelli, *On the Poincaré inequality for vector fields,* Ark. Mat. **38** (2000), 327–342.

[53] Y. Liu, G. Lu, R. L. Wheeden, *Some equivalent definitions of high order Sobolev spaces on stratified groups and generalizations to metric spaces,* Math. Ann. **323** (2002), 157–174.

[54] J. Luukkainen, E. Saksman, *Every complete doubling metric space carries a doubling measure,* Proc. Amer. Math. Soc. **126** (1998), 531–534.

[55] S. Rickman, *Quasiregular mappings,* Ergebnisse der Mathematik und ihrer Grenzgebiete (3), 26. Springer-Verlag, Berlin, 1993.

[56] J. Rissanen, *Wavelets on self-similar sets and the structure of the spaces $M^{1,p}(E, u)$,* Ann. Acad. Sci. Fenn. Math. Diss. **125** (2002), 46 pp.

[57] L. Saloff-Coste, *A note on Poincaré, Sobolev, and Harnack inequalities,* Internat. Math. Res. Notices **1992**, no. 2, 27–38.

[58] L. Saloff-Coste, *Aspects of Sobolev-type inequalities,* London Mathematical Society Lecture Note Series, 289. Cambridge University Press, Cambridge, 2002.

[59] S. Semmes, *Finding curves on general spaces through quantitative topology, with applications to Sobolev and Poincaré inequalities,* Selecta Math. (N.S.) **2** (1996), 155–295.

[60] S. Semmes, *Some novel types of fractal geometry,* Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2001.

[61] N. Shanmugalingam, *Newtonian spaces: an extension of Sobolev spaces to metric measure spaces,* Rev. Mat. Iberoamericana **16** (2000), 243–279.

[62] E. M. Stein, *Singular integrals and differentiability properties of functions,* Princeton Mathematical Series, No. 30 Princeton University Press, Princeton, N.J. 1970

[63] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals,* Princeton Mathematical Series, 43, Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993.

[64] K. T. Sturm, *Analysis on local Dirichlet spaces III. The parabolic Harnack inequality,* J. Math. Pures Appl. **75** (1996), 273–297.

[65] J. Väisälä, *Lectures on n-dimensional quasiconformal mappings,* Lecture Notes in Math. 299, Springer 1971.

[66] A. L. Volberg, S. V. Konyagin, *On measures with the doubling condition,* Izv. Acad. Nauk SSSR **51** (1987), 666–675 (in Russian); English transl. Math. USSR Izv. **30** (1988), 629–638.

[67] W. Woess, *Random walks on infinite graphs and groups,* Cambridge Tracts in Mathematics, 138. Cambridge University Press, Cambridge, 2000.

INSTITUTE OF MATHEMATICS, WARSAW UNIVERSITY, UL. BANACHA 2, 02–097 WARSZAWA, POLAND

E-mail address: `hajlasz@mimuw.edu.pl`