Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Definitions of Sobolev classes on metric spaces

by

Bruno Franchi, Piotr Hajłasz and Pekka Koskela

Preprint-Nr.: 11 1998

Definitions of Sobolev classes on metric spaces

Bruno Franchi, Piotr Hajłasz and Pekka Koskela*

Abstract

There have been recent attempts to develop the theory of Sobolev spaces $W^{1,p}$ on metric spaces that do not admit any differentiable structure. We prove that certain definitions are equivalent. We also define the spaces in the limiting case $p = 1$.

1. Introduction. Let $\Omega \subset \mathbb{R}^n$ be an open set. By the classical Sobolev space $W^{1,p}(\Omega)$ we mean the Banach space of those p-summable functions whose distributional gradients are p-summable as well. The space is equipped with the norm $\|u\|_{W^{1,p}} = \|u\|_p + \|\nabla u\|_p$. Here and in what follows by $\|\cdot\|_p$ we denote the L^p norm.

There are several ways to generalize the notion of the Sobolev space to the setting of metric spaces equipped with a Borel measure. Two natural definitions are the following. For $1 \leq p < \infty$, define the Sobolev space $M^{1,p}(S, d, \mu)$ as the set of all $u \in L^p(S)$ for which there exists $0 \leq g \in L^p(S)$ such that the inequality

$$|u(x) - u(y)| \leq d(x, y)(g(x) + g(y)) \tag{1}$$

holds μ-a.e. In the classical setting g corresponds to the maximal function of $|\nabla u|$. By $P^{1,p}(S, d, \mu)$ we denote the set of all functions $u \in L^p(S)$ such that there exist $0 \leq g \in L^p(S)$, $C > 0$ and $\lambda \geq 1$ so that the $(1, p)$-Poincaré inequality

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

holds on every ball B in S, where r is the radius of B, u_B is the average value of u on B, and $\fint$ denotes the average value of the integral. Notice that when $S = \mathbb{R}^n$, μ is the Lebesgue measure and d the euclidean distance, $u \in W^{1,p}(\mathbb{R}^n)$ and $g = |\nabla u|$ this

*B.F. was partially supported by Italian MURST and GNAFA of CNR, Italy, P.H. by KBN grant no. 2–PO3A–055–14, and P.K. by the Academy of Finland, SA-34082. B.F. thanks the Banach Center in Warsaw for its hospitality during the preparation of the manuscript. The research was completed during the stay of P.H. in the Max-Planck Institute in Mathematics in the Sciences in Leipzig, 1998. He thanks the Institute for the support and the hospitality.

inequality follows from the usual Poincaré inequality. These two different approaches both result in the usual Sobolev class when $p > 1$, $S = \mathbb{R}^n$, and μ is the Lebesgue measure and d the euclidean distance.

In this paper we compare these two different definitions in the setting of metric spaces and show that the Poincaré inequality for pairs of functions and upper gradients plays a key role in the subject. See Section 2 below for the definitions. More precisely, we shall use a fairly new self-improving property of the right hand side of a Poincaré inequality (see Theorem 5), instead of the known self-improving property of the left hand side (see Theorem 3).

The central examples of metric spaces we have in mind are given by the so-called Carnot–Carathéodory metrics associated with a family of Lipschitz continuous vector fields. As there is a natural way to define the Sobolev classes in terms of a family of vector fields identified with first order differential operators, a crucial test for our definitions of Sobolev spaces associated with a metric is to check compatibility with this definition. It has been inquired by Garofalo and Strichartz whether the Sobolev space defined by the pointwise inequality (1) above for the Carnot–Carathéodory metric associated with a system of vector fields satisfying Hörmander's condition coincides with the space obtained as the closure of smooth functions in the Sobolev norm generated by the family of vector fields. Theorems 1, 11 (see the discussion preceding this result) and 12 and Corollary 13 below give a complete answer to this question. Some partial results have been obtained earlier in [18], [26], [30], [36], [38].

Acknowledgement. The authors wish to thank Juha Heinonen for his comments.

2. Sobolev spaces on a metric space. Our notation is fairly standard. By L-Lipschitz functions we mean Lipschitz functions with the Lipschitz constant L. The average value will be denoted by

$$u_B = \fint_B u \, d\mu = \frac{1}{\mu(B)} \int_B u \, d\mu.$$

The space $L^p_{\rm loc}$ is the space of functions L^p-summable on every ball. The characteristic function of a set E will be denoted by χ_E. Balls will be denoted by B. The ball concentric with B and with the radius λ times that of B will be denoted by λB. General constants will be denoted by C. The value of C may change even in the same string of estimates. By Borel measure we will mean nonnegative Borel-regular measure.

Fix a triple (S, d, μ), where (S, d) is a metric space and μ is a Borel measure that is finite on every ball. For $1 \le p < \infty$, Hajłasz [26], defines the Sobolev space $M^{1,p}(S, d, \mu)$ as the set of all $u \in L^p(S)$ for which there exists $0 \le g \in L^p$ such that the inequality

$$|u(x) - u(y)| \le d(x, y)(g(x) + g(y)) \tag{2}$$

holds μ-a.e. From now on, we shall write a.e. instead of μ-a.e.

The space $M^{1,p}(S,d,\mu)$ is equipped with the Banach norm

$$\|u\|_{M^{1,p}} = \|u\|_p + \inf_g \|g\|_p,$$

where the infimum is taken over the set of all functions $0 \le g \in L^p$ that satisfy (2) a.e.

If $\Omega = \mathbb{R}^n$ or if $\Omega \subset \mathbb{R}^n$ is an open bounded set with smooth boundary, $|\cdot|$ is the Euclidean metric, and H^n is the Lebesgue measure, then for $1 < p < \infty$

$$W^{1,p}(\Omega) = M^{1,p}(\Omega, |\cdot|, H^n), \tag{3}$$

see [26] and also [31]. However, these spaces are different for $p = 1$, see [27] or the example below. The inclusion $W^{1,p} \subset M^{1,p}$ for $1 < p < \infty$ follows from the well known pointwise inequality

$$|u(x) - u(y)| \le C|x-y|(M|\nabla u|(x) + M|\nabla u|(y)), \tag{4}$$

where $Mg(x) = \sup_{r>0} \fint_{B(x,r)} |g(z)|\, dz$ denotes the Hardy–Littlewood maximal operator. The opposite inclusion requires another argument that we shall not discuss now. Since the maximal operator fails to be bounded on L^1, it is not surprising that in general $W^{1,1} \not\subset M^{1,1}$. The following example is taken from [27].

Let $\Omega = (-1/2, 1/2)$, and $u(x) = -x/(|x| \log |x|)$. Then $u \in W^{1,1}(\Omega)$, since $u'(x) = |x|^{-1}(\log |x|)^{-2} \in L^1(-1/2, 1/2)$. Suppose that there exists $g \in L^1(-1/2, 1/2)$, such that (2) holds with $d(x,y) = |x-y|$. Then for $0 < x < 1/2$, we would have $|u(x) - u(-x)| \le 2x(g(x) + g(-x))$ and hence

$$\frac{-2}{\log x} \le 2x(g(x) + g(-x)).$$

Thus

$$\int_{-1/2}^{1/2} g(x)\, dx = \int_0^{1/2} (g(x) + g(-x))\, dx \ge \int_0^{1/2} \frac{-dx}{x \log x} = \infty.$$

This contradicts the summability of g, proving that $W^{1,1} \not\subset M^{1,1}$, and hence that the equivalence (3) fails for $p = 1$.

Roughly speaking, the function g in (2) corresponds to the maximal function of the gradient. It does not reflect pointwise properties of the gradient, but it has good properties in average — the L^p norms are comparable i.e. $\|\nabla u\|_p \approx \inf_g \|g\|_p$, for $1 < p < \infty$.

In the remaining part of this paper we assume that S is a metric space of homogeneous type, i.e. that the measure μ is Borel, finite on every ball, and that it satisfies a doubling condition. This means that there exists a constant $C_d > 0$ such that for every ball B

$$\mu(2B) \le C_d \mu(B).$$

Integrating the pointwise inequality (2) with respect to x and y, we obtain the Poincaré type inequality

$$\fint_B |u - u_B|\, d\mu \le Cr \fint_B g\, d\mu. \tag{5}$$

This inequality seems weaker than (2), but, as we will see later on, the two inequalities are almost equivalent.

Theorem 1 *Let $1 < p < \infty$. The following conditions are equivalent.*

1. *$u \in M^{1,p}(S, d, \mu)$.*
2. *$u \in L^p(S)$ and there exist $C > 0$, $\lambda \ge 1$, $0 \le g \in L^p(S)$, and $1 \le q < p$ such that the Poincaré inequality*

$$\fint_B |u - u_B|\, d\mu \le C\, r \left(\fint_{\lambda B} g^q\, d\mu \right)^{1/q} \tag{6}$$

holds on every ball B of radius r.

This result is due to Hajłasz and Koskela [30]. Moreover, under some additional assumptions a version of it for $q = 1$ has been proved by Franchi, Lu, and Wheeden using a representation formula, [18], [22]. For further generalizations, see [28], [33], [38].

Implication 1. ⇒ 2. follows form (5). The opposite implication 2. ⇒ 1. is a consequence of the following proposition [30, Theorem 3.2] and the maximal theorem.

Proposition 2 *Assume that the pair (u, g), $u \in L^1_{\rm loc}$, $0 \le g \in L^q_{\rm loc}$, $q \ge 1$, satisfies inequality (6) for every ball B of radius r. Then*

$$|u(x) - u(y)| \le C d(x, y) \left((Mg^q(x))^{1/q} + (Mg^q(y))^{1/q} \right)$$

for almost every $x, y \in S$, where $Mh(x) = \sup_{r>0} \fint_{B(x,r)} |h|\, d\mu$.

The maximal theorem states that for $1 < p \le \infty$, $\|Mh\|_{L^p(S)} \le C\|h\|_{L^p(S)}$. In the limiting case $p = 1$ we obtain the following weak type estimate

$$\mu(\{Mh > t\}) \le \frac{C}{t} \int |h|\, d\mu. \tag{7}$$

The maximal theorem for doubling measures is due to Coifman and Weiss [8]; the proof follows almost the same argument as in the Euclidean case.

As we already noticed, the function g in (2) corresponds to the maximal function of the gradient, while the functions g in (6) looks more like the norm of the gradient.

Let $u \in L^1_{\rm loc}(S)$ and $g : S \to [0,\infty]$ be Borel measurable functions. We say that the pair (u,g) satisfies a (q,p)-Poincaré inequality, $p,q \geq 1$ if there exist $C>0$ and $\lambda \geq 1$ such that the inequality

$$\left(\fint_B |u-u_B|^q\, d\mu\right)^{1/q} \leq Cr\left(\fint_{\lambda B} g^p\, d\mu\right)^{1/p} \tag{8}$$

holds on every ball B of radius r. We do not put any integrability conditions upon g here, so that we can take for example $g \equiv \infty$.

By $P^{q,p}_{\rm loc}$ we denote the set of all functions $u \in L^q_{\rm loc}$ such that there exist $0 \leq g \in L^p$, $C>0$ and $\lambda \geq 1$ which make the pair (u,g) satisfy the (q,p)-Poincaré inequality (8) on every ball B of radius r.

Obviously, inequality (8) with $\lambda = 1$ implies (8) with $\lambda > 1$. However, in general the converse implication does not hold (i.e. $\lambda > 1$ cannot be replaced by $\lambda = 1$), see [30]. However, if the metric space satisfies some additional geometric properties, then one can replace $\lambda > 1$ in (8) by $\lambda = 1$ in the sense that if the pair (u,g) satisfies (8) with $\lambda > 1$ on every ball, then there exists a bigger constant C, such that (u,g) satisfies (8) with $\lambda = 1$ on every ball. A sufficient geometric condition for the replacement of $\lambda > 1$ by $\lambda = 1$ is that bounded and closed sets are compact and the distance between any pair of points equals the infimum of lengths of curves joining the two points. For details, see [34], [15, Section 5], [17, Section 3], [25], [29], [30].

The following imbedding theorem is due to Hajłasz and Koskela [29], [30, Theorem 5.1] (see also [19]).

Theorem 3 *Let (S,d,μ) be a metric space with μ doubling. Assume that the pair (u,g) satisfies a $(1,p)$-Poincaré inequality. Then there exists $q_0 > p$ such that for every $1 \leq q < q_0$ the pair (u,g) satisfies (q,p)-Poincaré inequality. The exponent q_0 depends only on the doubling constant and on p.*

The constants C and λ for the $(1,p)$ and (q,p)-Poincaré inequalities in the above theorem may be different.

Thus for the given range $1 \leq q < q_0$, the class $P^{q,p}_{\rm loc}$ is the same as the class $P^{1,p}_{\rm loc}$. For that reason we restrict our attention to the class $P^{1,p}$ only. We denote all L^p-integrable functions in $P^{1,p}_{\rm loc}$ by $P^{1,p}(S,d,\mu)$ or simply by $P^{1,p}(S)$ or by $P^{1,p}$.

Theorem 1 states that $u \in M^{1,p}$, $p>1$, if and only if $u \in L^p$ and there exist $g \in L^p$ and $1 \leq q < p$ such that the pair (u,g) satisfies $(1,q)$-Poincaré inequality. This suggests the following question: Is it true that $M^{1,p}(S,d,\mu) = P^{1,p}(S,d,\mu)$ for $1 \leq p < \infty$? In the case $p=1$ the answer is negative (see remark after Theorem 4). In the case $p>1$ the answer is positive provided we assume that in addition the space supports a $(1,q)$-Poincaré inequality for some $1 \leq q < p$ (see the definition below). We do not know if any additional condition is necessary. The positive answer is due to Koskela and MacManus [38].

A related question was also raised by Hajłasz and Koskela, [30]: Is it true that if the pair (u, g) satisfies $(1, p)$-Poincaré inequality, $1 < p < \infty$, then there exists $1 \le q < p$ such that the pair (u, g) satisfies $(1, q)$-Poincaré inequality? Note that the positive answer to that question together with Theorem 1 would imply $M^{1,p} = P^{1,p}$. Below we give a positive answer to the question under the same additional assumption as before: the space supports a $(1, q)$-Poincaré inequality for some $1 \le q < p$.

First we need some definitions.

Let (S, d, μ) be a triple as above. Following Heinonen and Koskela [33] we say that a Borel function $g : S \to [0, \infty]$ is an *upper gradient* of another Borel function $u : S \to \mathbb{R}$ if for every 1-Lipschitz curve $\gamma : [0, T] \to S$ we have

$$|u(\gamma(0)) - u(\gamma(T))| \le \int_0^T g(\gamma(t))\, dt$$

(remember that a curve γ is called 1-Lipschitz if $d(\gamma(\beta), \gamma(\alpha)) \le |\beta - \alpha|$ for all $0 \le \alpha < \beta \le T$). Moreover, we stress the fact that we could define upper gradients using the class of rectifiable curves, due to the fact that every rectifiable curve admits an arc-length parametrization, (see. [3], [45]).

We say that the space (S, d, μ) *supports a* $(1, p)$*-Poincaré inequality*, $1 \le p < \infty$, if there exist $C > 0$ and $\lambda \ge 1$ such that if u is a continuous function and g is an upper gradient of u, then the pair (u, g) satisfies a $(1, p)$-Poincaré inequality with given C and λ.

The above notions have been developed in [33], [2], [30], [38], [42], [44].

The theorem of Koskela and MacManus [38] reads as follows.

Theorem 4 *Let $1 < p < \infty$. If the space supports a $(1, q)$-Poincaré inequality for some $1 \le q < p$, then $P^{1,p}(S, d, \mu) = M^{1,p}(S, d, \mu)$.*

The most important example is $\mathbb{R}^n$ with the Euclidean metric $|\cdot|$ and the Lebesgue measure H^n. The space supports a $(1, 1)$-Poincaré inequality and hence

$$P^{1,p}(\mathbb{R}^n, |\cdot|, H^n) = M^{1,p}(\mathbb{R}^n, |\cdot|, H^n), \tag{9}$$

for all $1 < p < \infty$. As we already noted both spaces coincide with $W^{1,p}(\mathbb{R}^n)$, see (3). Later we will generalize this result to the case $p = 1$ and prove that $W^{1,1}(\mathbb{R}^n) = P^{1,1}(\mathbb{R}^n, |\cdot|, H^n)$. As $W^{1,1} \ne M^{1,1}$ by the above example, we will also obtain that $P^{1,1} \ne M^{1,1}$.

Theorem 4 is a consequence of the following stronger result and of Theorem 1.

Theorem 5 *Let $1 \le p < \infty$ and let the space support a $(1, q)$-Poincaré inequality for some $1 \le q \le p$ with given $\lambda \ge 1$. Let $\tau \ge 1$ and assume that the pair (u, g),*

$u \in L^1_{\rm loc}(S)$, $0 \le g \in L^p(S)$, *satisfies the family*

$$\fint_B |u - u_B|\, d\mu \le Cr \left(\fint_{\tau B} g^p\, d\mu \right)^{1/p} \tag{10}$$

of Poincaré inequalities on every ball B with radius r. Then there exists another constant $C' > 0$ such that for every ball B of radius r

$$\fint_B |u - u_B|\, d\mu \le C' r \left(\fint_{\lambda B} g^q\, d\mu \right)^{1/q}. \tag{11}$$

Remarks. 1) Compare the case $q = p$ with the discussion preceding Theorem 3. The novelty here is that λ might be smaller than τ. 2) The idea of the proof is to approximate u by "convolutions". The approximating sequence satisfies a $(1, q)$-Poincaré inequality and by passing to the limit we obtain (11). Similar techniques of approximation were employed in [42], [21], [38]. The case $q = p$ requires new ideas.

Proof of Theorem 5. We start with a construction of an approximating sequence. Fix $\varepsilon > 0$ and let $\{B_i^*\}$ be a covering of S by balls with radii ε and the property that the balls $\frac{1}{5}B_i^*$ are pairwise disjoint. Put now $B_i = 2B_i^*$: the doubling property implies that there is a constant C such that no point of S belongs to more than C balls B_i. Let $\{\varphi_i\}$ be a Lipschitz partition of unity associated to the given family of balls i.e., $\sum_i \varphi_i = 1$, $0 \le \varphi_i \le 1$, $\operatorname{supp} \varphi_i \subset B_i$ and all the functions φ_i are Lipschitz with the same constant $C\varepsilon^{-1}$. To this end it is enough to choose

$$\varphi_i = \frac{\psi\left(\frac{d(x_i, x)}{\varepsilon}\right)}{\sum_k \psi\left(\frac{d(x_k, x)}{\varepsilon}\right)},$$

where ψ is a smooth function, $\psi \equiv 1$ on $[0, 1]$, $\psi \equiv 0$ on $[3/2, \infty)$, $0 \le \psi \le 1$, and x_i is the center of B_i for $i = 1, 2, \ldots$ We can define now $u_\varepsilon = \sum_i \varphi_i u_{B_i}$. Then $\int_B |u - u_\varepsilon|\, d\mu \to 0$ as $\varepsilon \to 0$ on every ball B. Indeed, this is obvious when u is continuous and the general case follows by approximating u by continuous functions in the L^1 norm. For the following lemma, see [38, Lemma 4.7].

Lemma 6 *Let u be an arbitrary locally integrable function. If $d(b, a) < \varepsilon$, then*

$$|u_\varepsilon(b) - u_\varepsilon(a)| \le C d(b, a) h_\varepsilon(a),$$

where

$$h_\varepsilon = \sum_i \left(\fint_{4B_i} \fint_{4B_i} \frac{|u(y) - u(x)|}{\varepsilon}\, d\mu(y)\, d\mu(x) \right) \chi_{B_i}.$$

We do not prove this lemma. Later we will prove a similar result (Lemma 14), but in a different setting. The proof given there may be easily modified to cover Lemma 6.

As was noticed in [38], the above lemma implies that Ch_ε is an upper gradient for u_ε. Indeed, let $\gamma : [0,T] \to S$ be a 1-Lipschitz curve. Then for $0 \le \alpha < t < \beta \le T$, $\beta - \alpha < \varepsilon$, we have

$$\begin{aligned} |u_\varepsilon(\gamma(\beta)) - u_\varepsilon(\gamma(\alpha))| &\le |u_\varepsilon(\gamma(\beta)) - u_\varepsilon(\gamma(t))| + |u_\varepsilon(\gamma(\alpha)) - u_\varepsilon(\gamma(t)| \\ &\le C|\beta - \alpha| h_\varepsilon(\gamma(t)). \end{aligned} \tag{12}$$

The last inequality follows from Lemma 6 as γ is 1-Lipschitz. Since $\alpha < t < \beta$ was arbitrary, taking the average of (12) over t we obtain

$$|u_\varepsilon(\gamma(\beta)) - u_\varepsilon(\gamma(\alpha))| \le C \int_\alpha^\beta h_\varepsilon(\gamma(t))\, dt.$$

Let now $t_0 = 0 < t_1 < \ldots < t_N = T$ be a partition of $[0,T]$ into N congruent intervals of length less than ε. Then

$$|u(\gamma(T)) - u(\gamma(0))| \le \sum_{j=1}^N |u(\gamma(t_j)) - u(\gamma(t_{j-1}))| \le \int_0^T h_\varepsilon(\gamma(t))\, dt,$$

and the assertion is proved.

If the space supports a $(1,q)$-Poincaré inequality, we conclude that the pair $(u_\varepsilon, h_\varepsilon)$ satisfies a $(1,q)$-Poincaré inequality.

Assume now that the pair (u,g) satisfies the assumptions of Theorem 5. It remains to prove that if we pass to the limit as $\varepsilon \to 0$ in the $(1,q)$-Poincaré inequality for $(u_\varepsilon, h_\varepsilon)$ then we arrive at the desired inequality (11).

As a direct consequence of the definition of h_ε and the $(1,p)$-Poincaré inequality for (u,g) we obtain the following result.

Lemma 7 *If the pair (u,g) satisfies the family (10) of $(1,p)$-Poincaré inequalities, then*

$$h_\varepsilon \le C \sum_i \left(\fint_{4\tau B_i} g^p \, d\mu \right)^{1/p} \chi_{B_i}.$$

The following lemma that seems to be of independent interest is the main new ingredient in our argument.

Lemma 8 *Let Y be a metric space equipped with a doubling measure ν. Let $0 \le g \in L^p(Y)$, $1 \le p < \infty$, and suppose $\sigma \ge 1$. To every $\varepsilon > 0$ we associate a covering $\{B_i\}_i$ as above. Let*

$$g_\varepsilon = \sum_i \left(\fint_{\sigma B_i} g^p \, d\nu \right)^{1/p} \chi_{B_i}.$$

Then $\limsup_{\varepsilon\to 0} g_\varepsilon \le Cg$ *a.e. Moreover, for every ball* B *and each* $1 \le q \le p$, *the family* $\{g_\varepsilon^q\}_\varepsilon$ *is uniformly integrable and*

$$\limsup_{\varepsilon\to 0} \int_B g_\varepsilon^q \, d\nu \le C \int_B g^q \, d\nu. \tag{13}$$

Here the constant C *depends only on* q *and on the doubling constant.*

We recall that the uniform integrability of the family $\{g_\varepsilon^q\}_\varepsilon$ on B means that for every $\eta > 0$ there exists $\delta > 0$ such that if $A \subset B$, $\mu(A) < \delta$, then $\sup_\varepsilon \int_A g_\varepsilon^q < \eta$.

Note that Theorem 5 is a direct consequence of the above two lemmas and the fact that $(u_\varepsilon, h_\varepsilon)$ satisfies the $(1,q)$-Poincaré inequality.

Proof of Lemma 8. First note that $\limsup_{\varepsilon\to 0} g_\varepsilon(x) \le Cg(x)$ whenever x is a Lebesgue point of g^p. Indeed, if $x \in B_i$, then $\sigma B_i \subseteq B(x, 2(1+\sigma)\varepsilon)$, and hence

$$\limsup_{\varepsilon\to 0} g_\varepsilon(x) \le \limsup_{\varepsilon\to 0} C \left(\fint_{B(x,2(1+\sigma)\varepsilon)} g^p \, d\nu \right)^{1/p} = Cg(x).$$

The constant C is independent of ε due to the fact that both the number of balls B_i such that $x \in B_i$ and the ratio $\nu(B(x, 2(1+\sigma)\varepsilon))/\nu(\sigma B_i)$ can be bounded by a constant depending only on the doubling constant.

Let us show now that the family $\{g_\varepsilon^q\}_\varepsilon$ is uniformly integrable on B. Since the sum in the definition of g_ε is locally finite, we have

$$g_\varepsilon^p \le C \sum_i \left(\fint_{\sigma B_i} g^p \, d\nu \right) \chi_{B_i}$$

and hence $\sup_\varepsilon \int_B g_\varepsilon^p \, d\nu \le C \int_Y g^p \, d\nu$. This and the Hölder inequality imply uniform integrability when $1 \le q < p$, so that we can restrict ourselves to the case $q = p$. If the family failed to be uniformly integrable on B, then there would exist $\eta > 0$, a sequence of sets $K_n \subset B$ and a sequence ε_n such that

$$\nu(K_n) \to 0 \qquad \text{and} \qquad \int_{K_n} g_{\varepsilon_n}^p \, d\nu > \eta$$

Then we would have

$$\begin{aligned} \eta < \int_{K_n} g_{\varepsilon_n}^p \, d\nu &\le C \sum_i \int_{K_n \cap B_i} \left(\fint_{\sigma B_i} g^p \, d\nu \right) d\nu \\ &= C \sum_{K_n \cap B_i \ne \emptyset} \frac{\nu(K_n \cap B_i)}{\nu(\sigma B_i)} \int_{\sigma B_i} g^p \, d\nu = A_n. \end{aligned} \tag{14}$$

Given $\varepsilon > 0$ we can find $\delta > 0$ such that

$$\nu(E) < \delta \qquad \Longrightarrow \qquad \int_E g^p \, d\nu < \varepsilon.$$

By the doubling property, there exists a constant C' such that $\sum_i \chi_{B_i} \le \sum_i \chi_{\sigma B_i} \le C'$. Fix a positive integer m and choose n so large that $\nu(K_n) < \delta/(C'm)$. Divide now the set of indices i such that $K_n \cap B_i \neq \emptyset$ into two classes: the class I_1 consists of all those i such that $\nu(K_n \cap B_i)/\nu(\sigma B_i) > 1/m$, whereas the class I_2 consists of all the remaining indices. We have

$$\nu(\bigcup_{i \in I_1} \sigma B_i) < m \sum_{i \in I_1} \nu(K_n \cap B_i) = m \sum_{i \in I_1} \int_{K_n} \chi_{B_i} d\nu \le mC'\nu(K_n) < \delta.$$

Hence

$$A_n = \sum_{i \in I_1} + \sum_{i \in I_2} \le \sum_{i \in I_1} \int_{\sigma B_i} g^p \, d\nu + \sum_{i \in I_2} \frac{1}{m} \int_{\sigma B_i} g^p \, d\nu \le C' \left(\varepsilon + \frac{1}{m} \int_Y g^p \, d\nu \right).$$

Since we can choose an ε arbitrary small and an m arbitrary large we arrive to a contradiction with (14). This completes the proof of the uniform integrability.

We now proceed to prove (13). Fix $1 \le q \le p$ and a ball B. It is enough to prove that for every sequence $\varepsilon_n \to 0$ for which the limit on the left hand side of (13) exists, we have

$$\lim_{n \to \infty} \int_B g_{\varepsilon_n}^q \, d\nu \le C \int_B g^q \, d\nu$$

with some constant C depending on q and the doubling constant only. Fix such a sequence $\{\varepsilon_n\}_n$. We need the following theorem of Dunford and Pettis, see [10] or [11].

Theorem 9 *Let Z be a measurable space equipped with a finite measure ν and let $f_n \in L^1(Z, \nu)$. Then the sequence $\{f_n\}_n$ is weakly relatively compact in $L^1(Z, \nu)$ if and only if the family $\{|f_n|\}_n$ is uniformly integrable.*

Due to the above theorem we can find a subsequence of $g_{\varepsilon_n}^q$ (also denoted by $g_{\varepsilon_n}^q$) and $h \in L^1(B)$ such that $g_{\varepsilon_n}^q \rightharpoonup h$ weakly in $L^1(B)$. Then due to Mazur's lemma a sequence of convex combinations of $g_{\varepsilon_n}^q$ converges to h a.e.

Since $\limsup_{n \to \infty} g_{\varepsilon_n} \le Cg$ a.e. we conclude that $h \le Cg^q$ a.e. and hence (13) follows.

This completes the proof of Lemma 8 and hence those of Theorem 5 and Theorem 4.

In the case $1 \le q < p$ of Lemma 8 we could provide a more direct proof. Namely we could avoid the proof of the uniform integrability of the family $\{g_\varepsilon^q\}_\varepsilon$, and replace Dunford–Pettis' theorem by the reflexivity of the space $L^{p/q}$ and the fact that the sequence $g_{\varepsilon_n}^q$ is bounded in $L^{p/q}$.

The case $1 \le q < p$ of Lemma 8 implies Theorem 5 for $1 \le q < p$ and hence it is sufficient for the proof of Theorem 4. The case $p = q$ of Lemma 8 will be used in the next section.

3. Sobolev spaces arising from vector fields. One of the central applications of the theory of Sobolev spaces on metric spaces concerns the theory of Sobolev spaces associated with a family of vector fields that we next describe.

Let $X = (X_1, \ldots, X_k)$ be a family of vector fields in $\Omega \subset \mathbb{R}^n$ with real valued, locally Lipschitz continuous coefficients. One can define the Sobolev space $W^{1,p}_X(\Omega)$, $1 \leq p \leq \infty$, associated with the family X as the space of all the functions with finite norm $\|u\|_{W^{1,p}_X} = \|u\|_p + \|Xu\|_p$, where $|Xu|^2 = \sum |X_j u|^2$ and the derivatives $X_j u$ are understood in the sense of distributions.

Another way to define the space for $1 \leq p < \infty$ is to take the closure of C^∞ functions in the above norm. As in the Euclidean case, the two approaches are equivalent. This was obtained independently in [21] and [25]. The method goes, however, back to Friedrichs [23].

For the sake of simplicity, we assume from now on that $\Omega = \mathbb{R}^n$.

It is well known that we can canonically associate with X a metric (the so-called *Carnot–Carathéodory metric*, or *control metric*) as follows: we say that an absolutely continuous curve $\gamma : [a, b] \to \mathbb{R}^n$ is admissible if there exist measurable functions $c_i(t)$, $a \leq t \leq b$, satisfying $\sum_{j=1}^k c_j(t)^2 \leq 1$ and $\dot\gamma(t) = \sum_{j=1}^k c_j(t) X_j(\gamma(t))$ a.e.

Then we can define the distance $\rho(x, y)$ between $x, y \in \mathbb{R}^n$ as the infimum of those $T > 0$ for which there exists an admissible curve $\gamma : [0, T] \to \mathbb{R}^n$ with $\gamma(0) = x$, $\gamma(T) = y$. If there is no admissible curve joining x to y, then we set $\rho(x, y) = \infty$.

In general ρ may not be a metric, since it need not be finite. However, in many important situations ρ is finite for every pair of points and hence it is a metric: for instance, this happens when the family X satisfies Hörmander's condition (i.e., when the rank of the Lie algebra generated by X equals n at any point) [41], or when X is a system of Grushin type vector fields like those in [16], [14], and [15]. In what follows we assume in addition that the identity map induces a homeomorphism between $\mathbb{R}^n$ endowed with the Euclidean topology and $\mathbb{R}^n$ endowed with the Carnot–Caratéodory metric. This assumption excludes pathological situations like typically ∂_x, $x_+\partial_y$ in $\mathbb{R}^2$.

To avoid misunderstandings, by $\widetilde{B}$ we will denote balls with respect to the Carnot–Carathéodory metric and we will call them *metric balls*; Lipschitz functions with respect to ρ will be called *metric Lipschitz*.

It was proved independently by Garofalo and Nhieu [24, Theorem 1.3] and by Franchi, Serapioni and Serra Cassano [21, Proposition 2.9] that if u is metric Lipschitz, then $X_j u \in L^\infty_{\rm loc}$ for $j = 1, 2, \ldots, k$, where $X_j u$ is understood in the sense of distributions. A careful examination of the estimates given in [21] and [24] leads, however, to a stronger result.

Theorem 10 *If u is metric L-Lipschitz, then $|Xu| \leq L$ a.e.*

Proof. By [21], [24] we know that $X_j u \in L^\infty_{\rm loc}$ for $j = 1, 2, \ldots, k$. Fix any point x where $Xu(x)$ is defined. We can assume that $|Xu(x)| > 0$, otherwise the inequality is obvious. Since $Xu = X(u - \text{const.})$ we can assume that $u(x) = 0$.

Let $Y = \sum_{j=1}^k c_j X_j$, where $c_j = X_j u(x)/|Xu(x)|$ and let $B(x, \varepsilon)$ denote the Euclidean ball. Since $u(x) = 0$, $\sup_{B(x,\varepsilon)} |u| \le L \operatorname{diam}_\rho(B(x, \varepsilon))$. Now the estimates in [21] imply that for every $\varphi \in C_0^\infty(B(x, \varepsilon))$

$$\begin{aligned} |\langle Yu, \varphi\rangle| &\le CL \operatorname{diam}_\rho(B(x,\varepsilon))\|\varphi\|_{L^1} \\ &+ \limsup_{t\to 0+} \int_{B(x,\varepsilon)} \left| \frac{u(z) - u(\exp_z(-tY))}{t} \right| |\varphi(z)|\, dz, \end{aligned} \tag{15}$$

where $t \mapsto \exp_z(-tY)$ denotes the integral curve of $-Y$ passing through z at $t = 0$.

Note that $t \mapsto \exp_z(-tY)$ is an admissible curve and hence $|u(z) - u(\exp_z(-tY))| \le L\rho(z, \exp_z(-tY)) \le L|t|$, so that

$$|\langle Yu, \varphi\rangle| \le (CL\operatorname{diam}_\rho(B(x,\varepsilon)) + L)\|\varphi\|_{L^1}.$$

This implies that

$$\sup_{B(x,\varepsilon)} |Yu| \le CL\operatorname{diam}_\rho(B(x,\varepsilon)) + L.$$

Note that $\operatorname{diam}_\rho(B(x,\varepsilon)) \to 0$ as $\varepsilon \to 0$ (because of the assumption that the identity map is a homeomorphism between ρ and the Euclidean metric), so that, taking the limit as $\varepsilon \to 0$, we get $|Yu| \le L$ a.e.

In [34], D. Jerison proved that if the vector fields satisfy Hörmander's condition then the following version of the Poincaré inequality holds

$$\left(\fint_{\widetilde{B}} |u - u_{\widetilde{B}}|^p \, dx \right)^{1/p} \le Cr \left(\fint_{\widetilde{B}} |Xu|^p \, dx \right)^{1/p} \tag{16}$$

for any $1 \le p < \infty$. Here we integrate with respect to the Lebesgue measure. A similar inequality for Grushin type vector fields has been obtained earlier by Franchi and Lanconelli, [16] (see also [14] and [15]). After those papers many other results have been obtained. The references in the subject include [1], [4], [5], [17], [18], [19], [22], [25], [29], [30], [39], [40].

We shall formulate our results in an abstract setting that does not rely on any specific smoothness or structure assumption on X. As e.g. in [24], [25], [21], [30] we just assume that the vector fields are such that for every locally metric Lipschitz function u, the pair $(u, |Xu|)$ satisfies a kind of $(1,p)$-Poincaré inequality. More precisely, we assume that there is a Borel measure μ, doubling with respect to ρ, $\lambda \ge 1$, $C > 0$, and $1 \le p < \infty$ such that for every metric ball $\widetilde{B}$ of radius r

$$\fint_{\widetilde{B}} |u - u_{\widetilde{B}}| \, d\mu \le Cr \left(\fint_{\lambda\widetilde{B}} |Xu|^p \, d\mu \right)^{1/p}. \tag{17}$$

Note that, as we pointed out above, without loss of generality we may assume $\lambda = 1$. However, this will not play any role in our proofs.

As examples show, [30], [33], even in the Euclidean setting it sometimes happens that a $(1,p)$-Poincaré inequality holds for some $p > 1$ but the $(1,1)$-Poincaré inequality fails.

Let $d\mu = \omega\, dx$, $\omega > 0$ a.e., $\omega \in L^1_{\rm loc}$. We define the Sobolev spaces $H^{1,p}_X(\mathbb{R}^n,\mu)$, $1 \le p \le \infty$, associated with the family of vector fields as a completion of locally metric Lipschitz functions in the norm

$$\|u\|_{H^{1,p}_X(\mathbb{R}^n,\mu)} = \|u\|_{L^p(\mu)} + \|Xu\|_{L^p(\mu)}.$$

If $1 \le p < \infty$, then every metric Lipschitz function can be approximated by C^∞ functions in the Sobolev norm, so equivalently $H^{1,p}_X(\mathbb{R}^n,\mu)$ can be defined as the closure of C^∞ functions. Indeed, let u be metric Lipschitz with compact support. Then by the argument of Friedrichs, [23], [21], [24], the usual convolution approximation $u_\varepsilon = \varphi_\varepsilon * u$ satisfies $u_\varepsilon \to u$ uniformly, $Xu_\varepsilon \to Xu$ in L^p (with respect to the Lebesgue measure) and Xu_ε is uniformly bounded, as we can see since

$$X_j u_\varepsilon = [X_j(u * \phi_\varepsilon) - (X_j u) * \phi_\varepsilon] + (X_j u) * \phi_\varepsilon.$$

Indeed, the last term is bounded since $X_j u$ is bounded (again by [21] and [24]), whereas the first term can be bounded by writing it explicitly as in the proof of Proposition 1.4 of [21]. This easily implies that $u_\varepsilon \to u$ in $H^{1,p}_X(\mathbb{R}^n,\mu)$.

Under an additional assumption on μ all the above definitions are equivalent with the distributional definition, see [21], [24]. However in the case of general weights it is more appropriate to define the Sobolev space as a closure of locally metric Lipschitz functions.

Recently, N. Garofalo and R. Strichartz independently raised the following question: *does the Sobolev space $H^{1,q}_X(\mathbb{R}^n)$ associated with a system of vector fields satisfying (for instance) Hörmander's condition coincide with the Sobolev space defined using the Carnot-Carathéodory distance as in definition (2)?*

As we have seen, even in the classical Euclidean setting the answer is no when $q = 1$, so that we assume in the question that $q > 1$.

If Poincaré inequality (17) holds for some $p \ge 1$, then Theorem 1 implies the inclusion $H^{1,q}_X(\mathbb{R}^n,\mu) \subset M^{1,q}(\mathbb{R}^n,\mu) \subset P^{1,q}(\mathbb{R}^n,\mu)$ for $p < q < \infty$. Thus the question concerns the opposite inclusions.

In the following theorems we give an affirmative answer. Moreover we give a "metric" characterization of the Sobolev space even for $q = 1$ which is a more striking result. Let as start with the following abstract result.

Theorem 11 *Let (S, d, μ) be a metric space equipped with a doubling measure and let N be a positive integer. Suppose that there is a linear operator which associates with each locally Lipschitz function u a measurable function $Du : S \to \mathbb{R}^N$ in such a way that*

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

Let $H^{1,p}(S)$ be the Banach space defined as the closure of the set of locally Lipschitz functions with finite norm $\|u\| = \|u\|_p + \|Du\|_p$. Then $P^{1,p}(S) \subset H^{1,p}(S)$ for $1 \leq p < \infty$.

It seems that in general there may be a problem with the definition of Du for a given $u \in H^{1,p}(S)$. Namely, suppose that u_k and v_k are two sequences of locally Lipschitz functions such that both sequences converge to u in L^p, but $Du_k \to g$ in L^p, $Dv_k \to h$ in L^p, $g \neq h$. Then (u, g) and (u, h) represent two different elements in $H^{1,p}(S)$, which means that the gradient is not uniquely determined (for related examples, see [12, p. 91], [7]). This makes the situation very unpleasant. Fortunately, for a reasonable class of spaces we have the uniqueness of the gradient.

We say that the *uniqueness of the gradient* holds if the following condition is satisfied: if u_n is a sequence of locally Lipschitz functions such that $u_n \to 0$ in L^p and $Du \to g$ in L^p, then $g = 0$. In such a situation we can associate a unique Du obtained by taking the limit of 'gradients' of the approximating sequence of locally Lipschitz functions to each $u \in H^{1,p}(S)$.

Theorem 12 *Let (S, d, μ) be a metric space equipped with a doubling measure and let N be a positive integer. Suppose that there is a linear operator which associates with each locally Lipschitz function u a measurable function $Du : S \to \mathbb{R}^N$ in such a way that*

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

Let $1 \leq p < \infty$. Assume that there exist $C > 0$ and $\lambda \geq 1$ such that for every locally Lipschitz function u, the pair $(u, |Du|)$ satisfies a $(1, p)$-Poincaré inequality with given C and λ. Define $H^{1,p}(S)$ as in Theorem 11. Then $H^{1,p}(S) = P^{1,p}(S)$, the uniqueness of the gradient holds and $|Du| \leq Cg$ a.e., whenever (u, g) satisfies the $(1, p)$-Poincaré inequality.

Corollary 13 *Assume that the system X of vector fields on $\mathbb{R}^n$ is such that the identity map gives a homeomorphism between the Carnot–Carathéodory metric ρ and the Euclidean metric. Let μ be doubling with respect to the metric ρ and such that $d\mu = \omega\, dx$, $\omega > 0$ a.e., $\omega \in L^1_{\rm loc}$. Let $1 \leq p < \infty$. Assume that there exist $C > 0$ and $\lambda \geq 1$ such that for every locally metric Lipschitz function u*

$$\fint_{\widetilde{B}} |u - u_{\widetilde{B}}|\, d\mu \leq Cr \left(\fint_{\lambda \widetilde{B}} |Xu|^p\, d\mu \right)^{1/p} \tag{18}$$

for all metric balls. Define $H^{1,p}_X(\mathbb{R}^n, \mu)$ as before (completion of the space of all locally metric Lipschitz functions). Then $H^{1,p}_X(\mathbb{R}^n, \mu) = P^{1,p}(\mathbb{R}^n, \rho, \mu)$, the uniqueness of the gradient holds and $|Xu| \leq Cg$ whenever (u, g) satisfies a (1,p)-Poincaré inequality (with constants which may be different from C and λ in (18)).

The assumptions of the corollary are satisfied for instance by a system of vector fields satisfying Hörmander's condition, by Grushin-type vector fields like those in [16], [14], and [15] or by the general vector fields considered in [24], [25], [9], [30].

Proof of Theorem 11. Assume that $u \in P^{1,p}$ i.e., there exists $0 \leq g \in L^p$ and $C > 0$, $\lambda \geq 1$ such that the $(1,p)$-Poincaré inequality

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

holds on every ball B of radius r.

Fix $\varepsilon > 0$ and define the covering $\{B_i\}$, the Lipschitz partition of unity $\{\varphi_i\}$, and u_ε as in the proof of Theorem 5. First we show that $u_\varepsilon \to u$ in $L^p(S)$ as $\varepsilon \to 0$.

Due to Theorem 3 there exists $\tau \geq 1$ such that

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

on every ball B of radius r. Here the constant C is not necessarily the same as in the first inequality.

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

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

and hence

$$\int_S |u - u_\varepsilon|^p\, d\mu \leq C \sum_i \int_{B_i} |u - u_{B_i}|^p\, d\mu \leq C\varepsilon \sum_i \int_{\tau B_i} g^p\, d\mu \leq C\varepsilon \int_S g^p\, d\mu.$$

Thus $u_\varepsilon \to u$ in $L^p(S)$ as $\varepsilon \to 0$.

The following lemma is a variant of Lemmas 6 and 7. For the sake of completeness we provide a proof.

Lemma 14 *Assume that the pair* (u,g) *is as above. Then*

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

Proof. For $x \in S$ fix a ball in the covering that contains x. Denote the ball by B_0. Then

$$\begin{aligned}
|Du_\varepsilon(x)| &= |D(u_\varepsilon - u_{B_0})(x)| \le \sum_{i:\, x\in B_i} |D\varphi_i(x)||u_{B_i} - u_{B_0}| \\
&\le C\varepsilon^{-1} \sum_{i:\, x\in B_i} \fint_{B_0}\fint_{B_i} |u(y)-u(z)| \, d\mu(y)\, d\mu(z) \\
&\le C\varepsilon^{-1} \fint_{3B_0} |u(y) - u_{3B_0}| \, d\mu(y) \\
&\le C \left(\fint_{3\lambda B_0} g^p \, d\mu \right)^{1/p}.
\end{aligned}$$

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.

We claim that we can subtract a subsequence of Du_{ε_n} weakly convergent in L^p. Assume for a moment that we have already proved this claim and we show how to complete the proof of the theorem: indeed, by Mazur's lemma a sequence of convex combinations of u_{ε_n} is a Cauchy sequence for the norm in the space $H^{1,p}$ and this sequence converges to u. Thus $u \in H^{1,p}$ and Theorem 11 is proved once the claim is proved. Thus we are left with the proof of the claim.

Assume first that $p > 1$. By Lemma 14, $\sup_\varepsilon \int_S |Du_\varepsilon|^p \le C \int_S g^p$, so that the sequence Du_ε is bounded in L^p and the claim follows from the reflexivity of the space. The case $p = 1$ requires more effort. By Lemma 8 and Lemma 14 the family $|Du_{\varepsilon_n}|$ is uniformly integrable and $\limsup_{n\to\infty} |Du_{\varepsilon_n}| \le Cg$. Hence, by the Dunford–Pettis theorem, we find a function $h \in L^1_{\rm loc}$ and a subsequence of Du_{ε_n} (also denoted by Du_{ε_n}) that converges weakly to h in L^1 on every bounded set. Now it suffices to prove that the given subsequence converges weakly in $L^1(S)$.

As in the last step of the proof of Lemma 8, Mazur's lemma implies that $|h| \le Cg$ and hence $h \in L^1$. We have to prove that for every $\varphi \in L^\infty$,

$$\int_S Du_{\varepsilon_n}\varphi \, d\mu \to \int_S h\varphi \, d\mu. \tag{19}$$

We know that this property holds for φ with bounded support. By $B(R)$ we will denote the ball with radius R centred at a fixed point. Fix $\varphi \in L^\infty$. We have

$$\left| \int_S (Du_{\varepsilon_n} - h)\varphi \right| \le \left| \int_{B(R)} (Du_{\varepsilon_n} - h)\varphi \right| + \|\varphi\|_\infty \int_{S\setminus B(R)} |h| + \|\varphi\|_\infty \int_{S\setminus B(R)} |Du_{\varepsilon_n}|$$

The first term on the right hand side goes to 0 as $n \to \infty$. The second term is very small provided R is sufficiently large. To estimate the last term we apply Lemma 14,

$$\int_{S\setminus B(R)} |Du_{\varepsilon_n}| \leq C \int_{S\setminus B(R-8\lambda\varepsilon_n)} g.$$

This term is very small (independently on n) provided R is large. The estimates imply convergence (19). The proof of Theorem 11 is complete.

Proof of Theorem 12. First we prove the uniqueness of the gradient by modifying the argument of Semmes [32]. 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 $g = 0$.

By selecting a subsequence we may assume that

$$\int_S (|u_{n+1} - u_n|^p + |Du_{n+1} - Du_n|^p \, d\mu) \leq 10^{-np}. \tag{20}$$

This implies that the sequences u_n and Du_n converge a.e. Let $u_{n+1} - u_n = v_n$. Since by assumption $(u, |Du|)$ satisfies the $(1,p)$-Poincaré inequality, by Proposition 2 we have

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

Hence for $l \geq k \geq k_0$

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

where

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

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

$$|u_k(x) - u_k(y)| \leq Cd(x,y)(g_{k_0}(x) + g_{k_0}(y)), \tag{21}$$

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

$$\begin{aligned} \mu(\{g_{k_0} > t\}) &\leq \sum_{n=k_0}^{\infty} \mu\left(\left\{ (M|Dv_n|^p)^{1/p} > \frac{t}{2^{n-k_0+1}} \right\} \right) \\ &\leq \sum_{n=k_0}^{\infty} C \frac{2^{(n-k_0+1)p}}{t^p} \int_S |Dv_n|^p \, d\mu \\ &\leq C' t^{-p} 10^{-k_0 p}. \end{aligned}$$

In the middle inequality we used the weak type estimate for the maximal function (7), while in the last inequality we invoked (20).

Let $E_{k_0,t} = \{g_{k_0} > t\}$. Note that (21) implies that $u_k|_{S\setminus E_{k_0,t}}$ is Lipschitz with the Lipschitz constant Ct.

Observe now that if u is locally Lipschitz and $u|_F$ is Lipschitz with the Lipschitz constant L, then $|Du| \leq CL$ almost everywhere in F. Indeed, $u|_F$ can be extended to a globally Lipschitz function $\bar{u}$ on S with the Lipschitz constant L (see, [43, Theorem 5.1], [13, Section 2.10.4]). 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 $S \setminus E_{k_0,t}$, and hence $|g| \leq Ct$ a.e. in $S \setminus E_{k_0,t}$. Thus $\mu(\{|g| > Ct\}) \leq \mu(E_{k_0,t}) \to 0$ as $k_0 \to \infty$. Since $t > 0$ can be arbitrary small we conclude that $g = 0$ a.e. and the uniqueness of the gradient follows.

By Theorem 11 we know that $P^{1,p}(S) \subset H^{1,p}(S)$. The converse inclusion follows from the definition of $H^{1,p}(S)$ and the fact that we have a $(1,p)$-Poincaré inequality for locally Lipschitz functions.

Thus it remains to prove that if the pair (u, g) satisfies $(1,p)$-Poincaré inequality, then $|Du| \leq Cg$ a.e.

As in the proof of Theorem 11 we find a sequence u_{ε_n} such that $u_{\varepsilon_n} \to u$ in L^p and Du_{ε_n} is weakly convergent in L^p. Then by Mazur's lemma a sequence of convex combinations of u_{ε_n} is a Cauchy sequence in the norm of $H^{1,p}$.

By Lemma 8 and Lemma 14, $\limsup_{n\to\infty} |Du_{\varepsilon_n}| \leq Cg$ a.e. Since convex combinations of Du_{ε_n} converge to Du in L^p, we conclude that $|Du| \leq Cg$ a.e. This completes the proof.

Proof of Corollary 13. According to Theorem 12, we need only to prove the following lemma which is of independent interest.

Lemma 15 *Assume that the system of vector fields on $\mathbb{R}^n$ is such that the identity map gives a homeomorphism between the Carnot–Carathéodory metric ρ and the Euclidean metric. If u is locally metric Lipschitz and it is constant in a measurable set E, then $Xu = 0$ a.e. in E.*

Proof. Obviously, we may assume that $u = 0$ in E. Let $x \in E$ be simultaneously a Lebesgue point of Xu and a density point of E, both with respect to the Lebesgue measure.

Let $\varphi \in C_0^\infty(B(0,1))$, $\varphi \geq 0$, $\int \varphi(z)\,dz = 1$, and $\varphi_\varepsilon(z) = \varepsilon^{-n}\varphi(z/\varepsilon)$. Inequality (15) holds with Y replaced by any of the X_j's, so that we get

$$\begin{aligned} |(X_j u * \varphi_\varepsilon)(x)| &\leq CL\,\mathrm{diam}_\rho(B(x,\varepsilon)) \\ &+ C \limsup_{t\to 0+} \fint_{B(x,\varepsilon)} \left| \frac{u(z) - u(\exp_z(-tX_j))}{t} \right| dz \end{aligned} \tag{22}$$

Now $(X_j u * \varphi_\varepsilon)(x) \to X_j u(x)$ as $\varepsilon \to 0$, and thus it remains to show that the right hand side of the above inequality tends to 0 as $\varepsilon \to 0$. This is obviously true for the first term on the right hand side, so we are left with the second one.

Let $E^1_{\varepsilon,t}$ and $E^2_{\varepsilon,t}$ denote the sets of all $z \in B(x,\varepsilon)$ with $u(z) \neq 0$ or $u(\exp_z(-tX_j)) \neq 0$ respectively. Since the integrand in (22) is bounded by the Lipschitz constant of u in a neighborhood of x (cf. the proof of Theorem 10), it suffices to prove that to every $\varepsilon > 0$, there is $t(\varepsilon) > 0$ such that $\sup_{0<t<t(\varepsilon)} |E^i_{\varepsilon,t}|/|B(x,\varepsilon)| \to 0$ as $\varepsilon \to 0$ for $i = 1,2$. This is obvious for $i = 1$, as x is a density point of E and $u = 0$ on E. Now it remains to show that $\sup_{0<t<t(\varepsilon)} |E^2_{\varepsilon,t}| = o(\varepsilon^n)$.

Assume that $t > 0$ is sufficiently small. Let $\Phi_t(z) = \exp_z(tX_j)$. The inverse mapping is $\Phi_{-t}(z) = \exp_z(-tX_j)$. By [21, p. 101], the mapping $z \mapsto \Phi_t(z)$ is bi-Lipschitz on $B(x,\varepsilon)$. Moreover for $T > 0$ small, the Jacobian of the mapping, which is defined a.e. (by Rademacher's theorem), satisfies $J(z,t) = 1 + J_1(z,t)$, $|J_1(z,t)| \leq Ct$ for $0 \leq t \leq T$, $z \in B(x,\varepsilon)$, with the constant C not depending on x and t.

Note that $|E \cap B(x,\varepsilon-\varepsilon^2)| = |B(x,\varepsilon)| + o(\varepsilon^n)$ as $\varepsilon \to 0$. Indeed,

$$|E \cap B(x,\varepsilon-\varepsilon^2)| = |E \cap B(x,\varepsilon)| - |E \cap (B(x,\varepsilon) \setminus B(x,\varepsilon-\varepsilon^2))|.$$

Now since x is a density point of E, $|E \cap B(x,\varepsilon)| = |B(x,\varepsilon)| + o(\varepsilon^n)$, and $|E \cap (B(x,\varepsilon) \setminus B(x,\varepsilon-\varepsilon^2))| \leq |B(x,\varepsilon) \setminus B(x,\varepsilon-\varepsilon^2)| = o(\varepsilon^n)$. For $\varepsilon > 0$ we can find $t(\varepsilon)$ such that $t(\varepsilon) \to 0$ as $\varepsilon \to 0$ and such that $\Phi_t(B(x,\varepsilon-\varepsilon^2)) \subset B(x,\varepsilon)$ for $0 < t < t(\varepsilon)$. Hence by the change of variables formula

$$|\Phi_t(E \cap B(x,\varepsilon-\varepsilon^2)| \geq \int_{E\cap B(x,\varepsilon-\varepsilon^2)} 1 - Ct(\varepsilon) = |B(x,\varepsilon)| + o(\varepsilon^n). \tag{23}$$

Observe now that if $z \in \Phi_t(E \cap B(x,\varepsilon-\varepsilon^2))$, then $\exp_z(-tX_j) = \Phi_{-t}(z) \in E \cap B(x,\varepsilon-\varepsilon^2)$, so $u(\exp_z(-tX_j)) = 0$. Hence $\Phi_t(E \cap B(x,\varepsilon-\varepsilon^2)) \subset B(x,\varepsilon) \setminus E^2_{t,\varepsilon}$, and then by (23), $|E^2_{t,\varepsilon}| = o(\varepsilon^n)$. This ends the proof of the lemma and hence that for the corollary.

In the metric setting a good counterpart of the length of the gradient would be for example

$$Du(x) = \limsup_{y\to x} |u(y) - u(x)|/d(y,x).$$

Note that Du is an upper gradient of a given metric Lipschitz function u. However, this operator is not linear, and thus it is not covered by the above theorem. Thus it seems that the following modification of the above theorem would be more suitable for the general metric setting. Because this operator D is not linear, we cannot use Mazur's lemma to turn a sequence u_k convergent in L^p with Du_k weakly convergent in L^p into a Cauchy sequence in $W^{1,p}$. Thus we replace in the assumption that our space be a Banach space by the property of being closed under a kind of weak convergence.

Theorem 16 *Let (S,d,μ) be a metric space equipped with a doubling measure. Suppose that with each locally Lipschitz function u we can associate a nonnegative measurable function Du (called the length of the gradient) in such a way that*

1. *$D(u+v) \leq C(Du + Dv)$ and $D(\lambda u) \leq C|\lambda| Du$ a.e. whenever u,v are locally Lipschitz and λ is a real.*

2. *If u is L-Lipschitz, then $Du \le CL$ a.e.*

3. *If u is locally Lipschitz and constant on an open set $\Omega \subset S$, then $Du = 0$ a.e. in Ω.*

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

1. *If $u \in L^p(S)$ is locally Lipschitz and such that $Du \in L^p(S)$ then $u \in W^{1,p}(S)$ and $\|u\| \le C(\|u\|_p + \|Du\|_p)$.*

2. *If $u_k \in W^{1,p}$ converges in L^p to w and Du_k is a sequence weakly convergent in L^p then $w \in W^{1,p}$.*

Then $P^{1,p}(S) \subset W^{1,p}(S)$.

As the proof is almost the same as that for Theorem 11, we leave it to the reader.

References

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

[2] Bourdon, M., Pajot, H.: Poincaré inequalities and quasiconformal structure on the boundary of some hyperbolic buildings, (preprint).

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

[4] Capogna, L., Danielli, D., Garofalo, N.: An imbedding theorem and the Harnack inequality for nonlinear subelliptic equations. *Comm. Partial Diff. Equations* 18 (1993), 1765–1794.

[5] Capogna, L., Danielli, D., Garofalo, N.: Subelliptic mollifiers and a basic pointwise estimate of Poincaré type. *Math. Z.* 226 (1997), 147–154.

[6] Capogna, L., Danielli, D., Garofalo, N.: Capacitary estimates and the local behavior of solutions to nonlinear subelliptic equations. *Amer. J. Math.* 118 (1997), 1153–1196.

[7] Chiadò Piat, V., Serra Cassano, F.: Some remarks about the density of smooth functions in weighted Sobolev spaces. *J. Conv. Anal.* 1 (1994), 135–142.

[8] Coifman, R., Weiss, G.: *Analyse harmonique sur certains espaces homogenes*, Lecture Notes in Math. 242, Springer 1971.

[9] Danielli, D., Garofalo, N. Nhieu, D. M.: Trace inequalities for Carnot–Carathéodory spaces and applications to quasilinear subelliptic equations, (preprint).

[10] Dellacherie, C., Meyer, P. A.: *Probabilities and potential.* North-Holland Mathematics Studies, 29. North-Holland Publishing Co., 1978.

[11] Dunford, N., Schwartz, J.: *Linear Operators*, Part 1, John Wiley & Sons Inc., New York 1988.

[12] Fabes, E. B., Kenig, C. E., Serapioni, R. P.: The local regularity of solutions of degenerate elliptic equations. *Comm. Partial Differential Equations* 7 (1982), 77–116.

[13] Federer, H.: *Geometric Measure Theory.* Springer, 1969.

[14] Franchi, B.: Weighted Sobolev-Poincaré inequalities and pointwise estimates for a class of degenerate elliptic equations, *Trans. Amer. Math. Soc.* 327 (1991), 125–158.

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

[16] Franchi, B., Lanconelli, E.: Hölder regularity theorem for a class of linear nonuniformly elliptic operators with measurable coefficients. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) 10 (1983), 523–541.

[17] Franchi, B., Lu, G., Wheeden, R. L.: A relationship between Poincaré type inequalities and representation formulas in spaces of homogeneous type. *Ann. Inst. Fourier (Grenoble)* 45 (1995), 577-604.

[18] Franchi, B., Lu, G., Wheeden, R. L.: Representation formulas and weighted Poincaré inequalities for Hörmander vector fields, *Int. Mat. Res. Notices* (1996), 1–14.

[19] Franchi, B., Pérez, C., Wheeden R.L.: Self–improving iroperties of John–Nirenberg and Poincaré inequalities on spaces of homogeneous type. *J. Funct. Anal.* (to appear).

[20] Franchi, B., Serapioni R., Serra Cassano, F: Meyers–Serrin type theorems and relaxation of variational integrals depending on vector fields. *Houston Math. J.* 22 (1996), 859–889.

[21] Franchi, B., Serapioni R., Serra Cassano, F: Approximation and imbedding theorems for weighted Sobolev spaces associated with Lipschitz continuous vector fields *Boll. Un. Mat. Ital.* (7) 11–B (1997), 83–117.

[22] Franchi, B., Wheeden, R. L.: Some remarks about Poincaré type inequalities and representation formulas in metric spaces of homogeneous type. *J. Inequalities and Appl.* (to appear).

[23] Friedrichs, K. O.: The identity of weak and strong extensions of differential operators. *Trans. Amer. Math. Soc.* 55 (1944), 132–151.

[24] Garofalo, N., Nhieu, D. M.: Lipschitz continuity, global smooth approximations and extension theorems for Sobolev functions in Carnot–Carathéodory spaces, (preprint).

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

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

[27] Hajłasz, P.: Geometric approach to Sobolev spaces and badly degenerated elliptic equations. *The Proceedings of Banach Center Minisemester: Nonlinear Analysis and Applications.* (N.Kenmochi, M.Niezgódka, P.Strzelecki, eds.) GAKUTO International Series; Mathematical Sciences and Applications, vol. 7 (1995), 141–168.

[28] Hajłasz, P., Kinnunen, J.: Hölder quasicontinuity of Sobolev functions on metric spaces. *Rev. Mat. Iberoamericana*, (to appear).

[29] Hajłasz, P., Koskela, P.: Sobolev meets Poincaré. *C. R. Acad. Sci. Paris* 320 (1995), 1211–1215.

[30] Hajłasz, P., Koskela, P.: Sobolev met Poincaré, (preprint).

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

[32] Heinonen, J., Koskela, P.: Weighted Sobolev and Poincaré inequalities and quasiregular mappings of polynomial type. *Math. Scand.* 77 (1995), 251–271.

[33] Heinonen, J., Koskela, P.: Quasiconformal maps on metric spaces with controlled geometry. *Acta Math.* (to appear).

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

[35] Kałamajska, A.: On compactness of embedding for Sobolev spaces defined on metric spaces. *Ann. Acad. Sci. Fenn. Math.* (to appear).

[36] Kilpeläinen, T.: Smooth approximation in weighted Sobolev spaces. *Comment. Math. Univ. Carolinae* 38 (1997) 29–35.

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

[38] Koskela, P., MacManus, P.: Quasiconformal mappings and Sobolev spaces. *Studia Math.* (to appear).

[39] Lu, G.: Weighted Poincaré and Sobolev inequalities for vector fields satisfying Hörmander's condition and applications. *Rev. Mat. Iberoamericana*, 8 (1992), 367–439.

[40] Lu, G.: The sharp Poincaré inequality for free vector fields: An endpoint result, *Revista Mat. Iberoamericana*, 10 (1994), 453–466.

[41] Nagel, A., Stein, E. M. and Wainger, S.: Balls and metrics defined by vector fields I: Basic properties. *Acta Math.* 155 (1985), 103–147.

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

[43] Simon, L.: *Lectures on Geomteric Measure Theory.* Proc. of Centre for Math. Anal. Austral. Nat. Univ. Vol. 3, 1983.

[44] Tyson, J.: Quasiconformality and quasisymmetry in metric measure spaces. *Ann. Acad. Sci. Fenn. Math.* (to appear).

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

[46] Varopoulos, T., Saloff-Coste, L., Coulhon, T.:, *Analysis and geometry on groups* (Cambridge Univ. Press 1992).

Bruno Franchi: Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato, 5, I-40127 Bologna, Italy; E-mail: `franchib@dm.unibo.it`

Piotr Hajłasz: *(Permanent address)* Institute of Mathematics, Warsaw University, ul. Banacha 2 02–097 Warszawa, Poland; E-mail: `hajlasz@ghost.mimuw.edu.pl`

and

(Current address) Max-Planck Institute for Mathematics in the Sciences, Inselstr. 22–26, D-04103 Leipzig, Germany; E-mail: `hajlasz@mis.mpg.de`

Pekka Koskela: University of Jyväskylä, Department of Mathematics, P.O. Box 35, FIN-40351 Jyväskylä, Finland; E-mail: `pkoskela@math.jyu.fi`