

POINCARÉ INEQUALITIES AND SOBOLEV SPACES

PAUL MACMANUS

Abstract

Our understanding of the interplay between Poincaré inequalities, Sobolev inequalities and the geometry of the underlying space has changed considerably in recent years. These changes have simultaneously provided new insights into the classical theory and allowed much of that theory to be extended to a wide variety of different settings. This paper reviews some of these new results and techniques and concludes with an example on the preservation of Sobolev spaces by the maximal function.

Introduction

The Sobolev embedding theorem is one of the fundamental results of harmonic analysis and the theory of partial differential equations. The Sobolev-Poincaré form states that if $f \in W^{1,p}(\mathbb{R}^n)$ and $1 \leq p < n$ then

$$\left(\frac{1}{|B|} \int_B |f - f_B|^{np/(n-p)} dx \right)^{(n-p)/np} \leq C(n, p) r_B \left(\frac{1}{|B|} \int_B |\nabla f|^p dx \right)^{1/p}.$$

Here B is any ball in \mathbb{R}^n , r_B is its radius, and f_B is the average value of f on B . If the function f has compact support, then the preceding inequality yields the usual Sobolev form:

$$\left(\int_{\mathbb{R}^n} |f - f_B|^{np/(n-p)} dx \right)^{(n-p)/np} \leq C(n, p) \left(\int_{\mathbb{R}^n} |\nabla f|^p dx \right)^{1/p}.$$

These theorems are crucial for proving regularity properties, such as Harnack's inequality and Hölder continuity, for solutions to non-linear degenerate elliptic equations. A related, but apparently weaker, inequality is the Poincaré inequality

$$\frac{\int_B |f - f_B| dx}{|B|} \leq C(n) r_B \frac{\int_B |\nabla f| dx}{|B|},$$

2000 *Mathematics Subject Classification.* 46E35 (42B25).

Key words. Poincaré inequalities, Sobolev inequalities, metric spaces, doubling measures.

which holds for functions f in $W^{1,1}(\mathbb{R}^n)$. This is often called the $(1, 1)$ Poincaré inequality, in reference to the exponents. Another common type of Poincaré inequality is the $(2, 2)$ version. Moser's famous iteration technique combines this Poincaré inequality with the improved integrability estimate of Sobolev's theorem to obtain the regularity results mentioned above. The improved integrability need only be for any power greater than p ; there is no need to go all the way to $np/(n-p)$. This technique was first used by Moser to give a new proof of earlier regularity results of De Giorgi and Nash and has turned out to be a very powerful tool. It works just as well on Riemannian manifolds or on sub-Riemannian manifolds, such as the Heisenberg group, provided that the measure is doubling and both the Poincaré and Sobolev-Poincaré inequalities hold on the space. See [HKM], for example. The relevant measure is doubling and a Poincaré inequality holds on Riemannian manifolds with non-negative curvature [Bus], on Carnot groups [Va], [VSC], and, more generally, on Carnot-Carathéodory spaces where the vector fields satisfy the Hörmander condition [Je]. Obtaining the Sobolev-Poincaré inequality proved more elusive but through the work of numerous authors over the past ten years we now know that once doubling and Poincaré hold on a space then the Sobolev-Poincaré inequality is *automatically* satisfied. Thus all the regularity results follow once doubling and Poincaré are established.

The process of going from doubling and Poincaré to Sobolev-Poincaré is now just an abstract process that has little to do with either the class of functions being considered or with the geometry of the space in question. The procedure works on any space of homogeneous type, these being the spaces with the minimal structure required for doing harmonic analysis. On the other hand, the Poincaré inequality is intimately connected with the geometry of the space. For example, it is well known that on Riemannian manifolds the validity of the Poincaré inequality is equivalent to the relative isoperimetric property holding on the manifold.

Along with this development of the Sobolev embedding theorem in a wider setting has come a new understanding of what it means for a function to belong to a Sobolev class. Sobolev functions can now be defined on any metric space and a very complete Sobolev theory has been obtained on those metric spaces which support some type of Poincaré inequality.

1. Some definitions

Throughout this article X will be a metric space equipped with a regular, Borel, doubling measure μ of dimension d . We need the measure to be regular in order to have such standard results as Lebesgue’s differentiation theorem. The latter depends on continuous functions’ being dense in $L^p(X)$, and regularity of the measure is a standard way to achieve this. Recall that doubling means that $\mu(2B) \leq C_D\mu(B)$ for every ball B and that the dimension $d = \log_2 C_D$. It follows that

$$\frac{\mu(B_2)}{\mu(B_1)} \leq C \left(\frac{r_2}{r_1}\right)^d$$

whenever B_1 and B_2 are balls of radii r_1 and r_2 respectively with $B_1 \subset B_2$.

In the final section only we will need the following definitions. A metric space X is a length space (in the sense of Gromov) if the distance between any two points is the infimum of the lengths of all paths joining the points. X is a geodesic space if this infimum is always attained. Any space that is both complete (as a topological space) and a length space must be a geodesic space.

By analogy with the Fundamental Theorem of Calculus we say that a non-negative function g is an *upper gradient* for a function $f: X \rightarrow \mathbb{R}$ if $|f(b) - f(a)| \leq \int_{\gamma} g ds$ for all rectifiable paths γ in X that join a and b . We then define the Sobolev space to be

$$W^{1,p}(X) = \{f \in L^p(X) : f \text{ has an upper gradient in } L^p(X)\}.$$

The associated norm is $\|f\|_{1,p} = \|f\|_p + \inf_g \|g\|_p$ where the infimum is taken over all upper gradients of f . A weak upper gradient of f is the same as an upper gradient except that the inequality is allowed to fail for a family of paths of p -modulus zero. This extra freedom in the definition is quite useful and corresponds to the usual ACL definition of Sobolev functions in \mathbb{R}^n . It is easy to see that that the Sobolev space does not change if we replace upper gradients by weak upper gradients [KM]. It follows from the weak upper gradient version that these Sobolev spaces agree with the usual ones on any Riemannian manifold. The following result is due to Shanmugalingam [Sh]:

Theorem 1. $W^{1,p}(X)$ is a Banach space.

In order for this definition to be useful, X must have many rectifiable curves joining any pair of points in the space. One way to quantify

“many” is to require that the modulus of any annulus in the space satisfies a certain lower bound. This type of condition is called a Loewner condition and is crucial to extending the classical theory of quasiconformal maps to metric spaces (see [HeK]). It turns out that such a condition on X is equivalent to requiring that the X satisfy a certain Poincaré inequality, and this is the point of view that we will take here. We will say that X is Poincaré if for any continuous function f and any upper gradient g of f we have

$$\frac{1}{\mu(B)} \int_B |f(x) - f_B| d\mu \leq C \frac{r_B}{\mu(B)} \int_B g d\mu$$

for every ball $B \subset X$.

A huge number of metric spaces are Poincaré. For example:

- i) \mathbb{R}^n or, more generally, any Lipschitz domain in \mathbb{R}^n .
- ii) Riemannian manifolds with non-negative Ricci curvature [Bus].
- iii) Carnot groups, including the Heisenberg group [Va], [HaK2].
- iv) Carnot-Carathéodory spaces where the defining vector fields are smooth and satisfy Hörmander’s condition [Je].
- v) Carnot-Carathéodory spaces associated to certain non-smooth vector fields, including those of Grushin type [FL1], [FL2], [FGW].
- vi) Linearly locally contractable spaces equipped with doubling measures [Se].

In the case of ii), Buser also showed that the Riemannian measure is doubling. That the natural measure is doubling in iv) is a result of [NSW]. As well as Lipschitz domains, any domain in \mathbb{R}^n that is uniform or John, or that satisfies certain chaining conditions such as Boman’s condition, are Poincaré. These all follow from the Poincaré inequality for balls and a well-known procedure for converting inequalities for balls into inequalities for more general domains. See [Boj], [Chua], [FPW], [HaK1], [HaK2].

There are various other ways to define Sobolev functions on metric spaces. Some references are [Ha], [Sh], [KS], [KM], [Ch]. The strongest one, in that it implies all the others is that of Hajłasz. We denote this space by $M^{1,p}(X)$. It consists of those function $u \in L^p(X)$ for which there exists $g \in L^p(X)$ (called a gradient of u) with

$$|u(x) - u(y)| \leq C|x - y|(g(x) + g(y)) \text{ for a.e. } x \text{ and } y.$$

The norm on $M^{1,p}(X)$ is $\|f\|_{M^{1,p}} = \|f\|_p + \inf_g \|g\|_p$ where the infimum is taken over all gradients of f . It is not difficult to show that

$M^{1,p}(X) \subset W^{1,p}(X)$ and that the norms are comparable, see [Sh]. A big difference between the Hajlasz space and the others is that there is a Poincaré inequality built into the former definition. Simply integrate the inequality above in x and y to get the Poincaré inequality. In general, the various Sobolev spaces are not equal. Nevertheless we have the following theorem:

Theorem 2. *If X is Poincaré and $p > 1$, then $M^{1,p}(X) = W^{1,p}(X)$.*

We will prove this shortly. In fact, all of the Sobolev spaces mentioned above are the same when $p > 1$ and X is Poincaré, see [KM], [Sh].

A simple but important technique that is used time and time again is *chaining*. This is an old idea whose importance lies in the fact that it allows us to convert integral estimates into pointwise estimates. Early references for its use in recent developments are [HaK1] and [FLW2]. An example of chaining is the following representation formula. Suppose that B is a ball. Then for a.e. $x \in B/2$ we have

$$|f(x) - f_B| \leq \sum_{k=0}^{\infty} |f_{B_k} - f_{B_{k+1}}| \leq C \sum_{k=0}^{\infty} \frac{1}{\mu(B_k)} \int_{B_k} |f(x) - f_{B_k}| d\mu.$$

Here the B_k are any sequence of balls containing x , starting with B , whose radii decrease geometrically. The first inequality is a consequence of Lebesgue’s differentiation theorem while the second is simple algebra. We can now prove the theorem.

Proof: Suppose that $f \in W^{1,p}(X)$ and that $x, y \in X$. Let B be a ball containing x and y and of radius comparable to $|x - y|$. Then

$$\begin{aligned} |f(x) - f_{2B}| &\leq C \sum_{k=0}^{\infty} \frac{1}{\mu(B_k)} \int_{B_k} |f(x) - f_{B_k}| d\mu \\ &\leq C \sum_{k=0}^{\infty} r_k \frac{1}{\mu(B_k)} \int_{B_k} g d\mu \\ &\leq C|x - y| Mg(x). \end{aligned}$$

The first inequality uses chaining, the second uses the Poincaré inequality, and the final one follows from the definition of the maximal function and the fact that the r_k decrease geometrically.

Similarly, $|f(y) - f_{2B}| \leq C|x - y|Mg(y)$ and so we see that $f \in M^{1,p}(X)$ with Mg as a gradient. \square

Because of the appearance of the maximal function in the proof it is crucial that we have $p > 1$. The two spaces are not equal, not even in \mathbb{R}^n , when $p = 1$.

2. Sobolev-Poincaré from Poincaré

By a generalized Poincaré inequality we mean an inequality of the form

$$(1) \quad \frac{1}{\mu(B)} \int_B |f(x) - f_B| d\mu \leq a(B)$$

that holds for a function f and for all balls B in X . Thus we are controlling the oscillation of f by the functional a . If $a(B) \equiv 1$, then this control translates, by the John-Nirenberg theorem, into f 's being exponentially integrable. We want to start with a weaker hypothesis and obtain a weaker conclusion; that f be integrable to some higher power. If we allow $a(B)$ to become large, but only for balls of decreasing size centered on a small set, then we can get what we want. We will say that a functional a satisfies the D_r condition if for each ball B and any family $\{B_i\}$ of pairwise disjoint sub-balls of B ,

$$\sum_i a(B_i)^r \mu(B_i) \leq C^r a(B)^r \mu(B).$$

The best constant C on the right is denoted $\|a\|$. It is always at least one. Examples of such functionals are

- i) $a(B) = C r_B \left(\frac{1}{\mu(B)} \int_B |\nabla f|^2 d\mu \right)^{1/2}$ and X is a Riemannian manifold.
- ii) $a(B) = C r_B \left(\frac{1}{\mu(B)} \int_B g^p d\mu \right)^{1/p}$ where g is an upper gradient of f .
- iii) $a(B) = r_B^\alpha \left(\frac{\nu(B)}{\mu(B)} \right)^{1/p}$ where $\alpha > 0, p > 0, \nu$ is a measure.

These satisfy condition D_r for $r = \frac{2n}{n-2}$, $r = \frac{pn}{n-p}$, and $r = \frac{dp}{d-\alpha p}$ respectively. Case i) was considered by Saloff-Coste [S-C] in one of the earliest results in the area. The second case was examined in the important works [HaK1] and [HaK2] of Hajlasz and Koskela. Their ideas had a strong influence on the whole field. The D_r condition and the idea of a generalised Poincaré inequality appeared in [FPW]. A closely related approach was treated in [BCLS].

The Sobolev embedding theorem almost follows from the generalised Poincaré inequality (1) when a satisfies D_r . However, the best that can be obtained in general is a weak version of the theorem. If the functional a is of type ii) above or if X is connected, then the classical proof of the embedding theorem via Riesz potentials can be imitated. See [HaK2] for the former case and [MP2] for latter case when $p = d$ (corresponding to Trudinger’s inequality). The proofs involve chaining to bound $|f(x) - f_B|$ by a series (corresponding to a Riesz integral classically), then splitting the series into two terms (from 1 to N and N to infinity) that are bounded in different ways to obtain a bound that depends on N , and finally minimizing over N to bound $|f(x) - f_B|$ by an appropriate power of some maximal function. In order to deal with general a and general X it is necessary to take a different approach. This involves using the well-known good- λ technique of Burkholder and Gundy and completely avoids representation formulas and chaining.

A somewhat weaker version of the next result was obtained by Franchi, Pérez and Wheeden [FPW]. Their proof relied heavily on a complicated result of Sawyer and Wheeden on the existence of certain nested dyadic-type coverings of spaces of homogeneous type. The formulation below is taken from MacManus and Pérez [MP1]. As in [FPW] the proof is via a good- λ inequality but it uses only elementary geometric properties of spaces of homogeneous type.

Theorem 3. *Let B_0 be a ball in X and let $\delta > 0$ be given. Set $\widehat{B}_0 = (1 + \delta)B_0$. Suppose that the functional a satisfies the D_r condition for some $0 < r < \infty$. If f is a locally integrable function on \widehat{B}_0 for which there exist constants $\tau \geq 1$ and M_0 such that for all balls B with $B \subset \widehat{B}_0$*

$$(2) \quad \frac{1}{\mu(B)} \int_B |f - f_B| d\mu \leq M_0 a(\tau B),$$

then there exists a constant C independent of f and B_0 such that

$$(3) \quad \sup_{\lambda > 0} \lambda \left(\frac{\mu(\{x \in B_0 : |f(x) - f_{B_0}| > \lambda\})}{\mu(B_0)} \right)^{1/r} \leq C M_0 \|a\| a(\tau \widehat{B}_0).$$

Allowing the factor of τ on right in (2) is important because when one actually proves that a space X supports a Poincaré inequality there is typically a larger ball on the right. However, if all balls in X satisfy a chain condition, such as the Boman condition, there is a by now standard procedure which shows that Poincaré or Sobolev-Poincaré inequalities which have a larger ball on the right actually imply the same inequality but with the larger ball replaced by the usual ball. The point

is that these inequalities hold for all sub-balls of any given ball and that a chain condition allows us to combine the weaker inequality on balls belonging to some Whitney-type decomposition to obtain the improved inequality. See [Boj], [Je], [Bom], [Chua], [FPW]. Examples of domains satisfying chain conditions are Lipschitz domains, uniform domains, and John domains. See [HaK2] for yet more examples and other types of chain conditions.

The theorem can be strengthened in a number of ways. X does not have to be a metric space; it can be a space of homogeneous type, i.e., a space equipped with a quasimetric and a doubling measure. In this case, we pick up an extra factor of K , the quasimetric constant, in \widehat{B}_0 . Also, in the conclusion, the measure μ can be replaced by any measure w that is A_∞ related to μ , provided that the functional a satisfies the D_r condition but with μ replaced by w .

Sketch of proof: Fix a ball B_0 . We can assume that $f_{B_0} = 0$. A key object is the family of balls

$$(4) \quad \mathcal{B} = \{B : x_B \in B_0 \text{ and } r_B \leq \delta r_{B_0}\}.$$

The maximal function and the sharp maximal function associated to \mathcal{B} are

$$M_{\mathcal{B}}g(x) = \sup_{\substack{B:x \in B \\ B \in \mathcal{B}}} \frac{1}{\mu(B)} \int_B |f| d\mu \quad \text{and}$$

$$M_{\mathcal{B}}^{\#}g(x) = \sup_{\substack{B:x \in B \\ B \in \mathcal{B}}} \frac{1}{\mu(B)} \int_B |f - f_B| d\mu,$$

where we understand that the supremum is zero if x is not contained in any element of \mathcal{B} . Both of these functions are zero outside \widehat{B}_0 .

The heart of the proof is showing that, as in \mathbb{R}^n , $M^{\#}f$ and Mf are related by a good- λ inequality. The proof is similar to the classical one in that it involves a Calderón-Zygmund, or stopping-time, argument. Once we have this, condition D_r implies that $M^{\#}f$ lies in weak L^r and then the good- λ inequality implies that Mf lies in weak L^r , along with suitable estimates, of course. This yields the theorem. \square

As mentioned previously, the weak result is the best result in general. See [MP1] for an example. Under certain circumstances the weak norm can be replaced by the strong norm; this requires the functional a to depend on f and to have certain stability properties with respect to truncation of f . Given a non-negative function g , the truncation $\tau_\lambda(g)$

is defined by

$$\tau_\lambda(g) = \min\{g, 2\lambda\} - \min\{g, \lambda\} = \begin{cases} 0 & \text{if } g(x) \leq \lambda \\ g(x) - \lambda & \text{if } \lambda < g(x) \leq 2\lambda \\ \lambda & \text{if } g(x) > 2\lambda. \end{cases}$$

We now replace a by $a(B, f)$ and require the following properties:

- $a(B, f) = a(B, f + \lambda)$ for all $f \in \mathcal{F}$ and $\lambda \in \mathbb{R}$.
- $a(B, |f|) \leq a(B, f)$ for all $f \in \mathcal{F}$.
- There exist $r \geq 1$ and a constant C such that for any nonnegative $f \in \mathcal{F}$, any ball B and any sequence λ_k of the form $\{\lambda_k = 2^k \lambda\}$, $k = 1, 2, \dots$, $\lambda > 0$, we have

$$(5) \quad \sum_{k=1}^{\infty} a(B, \tau_{\lambda_k}(f))^r \leq c a(B, f)^r.$$

The complicated looking inequality (5) merely says that when we use truncations to slice up a function, then the sum of the norms of the slices is bounded by the norm of the original function. Typical examples of such functionals are $a(B, f) = C r_B^\alpha \left(\frac{1}{\mu(B)} \int_B g^p d\mu \right)^{1/p}$, where g is an upper gradient of f and X is Poincaré. Another example is $a(B, f) = \left(\frac{\nu_f(B)}{\mu(B)} \right)^{1/2}$, where ν_f is the energy measure of f given by a strongly local, irreducible, Dirichlet form.

Under these hypotheses we obtain the same result as before except with the strong norm replacing the weak norm, and so we have a complete extension of the Sobolev embedding theorem [MP1], [FPW]. To obtain the strong norm we simply cut the given function into slices, apply the weak estimate to each of these truncations, and then use the properties above to obtain the strong estimate for the original function. This technique has been rediscovered several times and has been used by many authors. Some references are [Ma], [LN], [Se], [BCLS], [FPW], [HaK2]. The importance of truncation methods in potential theory dates back to Beurling’s 1949 paper [Beu]. The ideas in that paper were exploited in the 1950s by Beurling and Deny in their development of the theory of Dirichlet spaces.

An alternative to the Poincaré inequality as a starting point for the development of a Sobolev theory is the idea of a representation formula.

An example of such a formula is

$$|f(x) - f_B| \leq C \int_{\tau B} g(y) \frac{|x - y|}{\mu(B(x, |x - y|))} d\mu(y)$$

where g is an upper gradient of f . In \mathbb{R}^n this is just the well known formula that relates f to the Riesz integral of its gradient. This approach was taken in, for example, [FGW] and [FLW1] in the context of metric spaces determined by Hörmander or Grushin type vector fields. It is easy to see that a representation formula implies a $(1, 1)$ -Poincaré inequality, simply use Fubini's theorem. The converse implication is true in many situations, see [FLW2], [FW], [LW]. A hint of how the proof goes can be found in the proof in Theorem 2, which gives a basic example of how an integral estimate can be converted into a pointwise estimate.

The theory of Sobolev functions has been reduced to its essentials. The key to obtaining a good theory is the Poincaré inequality and the existence of such an inequality depends crucially on the specific geometric nature of the space. There is an automatic self-improving process, requiring only the minimum amount of structure on the space, that converts the Poincaré inequality into the Sobolev embedding theorem. The robustness of this approach is clear from the the generality of the results already described and is further revealed by the fact that other standard results such as Trudinger's inequality, the Rellich-Kondrachov Theorem on compactness of Sobolev embeddings, and trace versions of the embedding theorem all hold in this context and can be proven using the techniques outlined here (see [HaK2], [Ka], [MP2]). In the examples just mentioned we sometimes need some extra hypothesis on X , such as connectivity or local connectivity, to get the full result. In the next section we show how making some natural geometric assumptions on X allows us to prove that the maximal function preserves smoothness of Sobolev functions.

3. An example

The following result of Kinnunen appeared in 1997 ([Ki1], also [Ki2]), a related result on the preservation of the BMO space can be found in Section 5.7 of [BeS].

Theorem 4. *If $f \in W^{1,p}(\mathbb{R}^n)$ and $p > 1$, then $Mf \in W^{1,p}(\mathbb{R}^n)$, $\|Mf\|_{1,p} \leq C \|f\|_{1,p}$, and $|\nabla Mf| \leq M|\nabla f|$.*

Here Mf is the usual Hardy-Littlewood maximal function.

This result can be extended to the setting of metric spaces using the techniques discussed in this paper. In order to get good estimates on

the oscillation of the maximal function we need to be sure that very thin annuli in X do not have a lot of mass. The precise condition we need is the following, which we call the annulus condition.

There is a constant C with the following property. Suppose that A is an annulus of outer radius R and thickness h and that B is any ball of radius no more than $2R$ that intersects A . Then

$$\frac{\mu(B \cap A)}{\mu(B)} \leq C \frac{h}{r_B}.$$

Theorem 5. *Let X satisfy the annulus condition. If $p > 1$ and $u \in M^{1,p}$, then $Mu \in M^{1,p}$ and $\|Mu\|_{1,p} \leq C \|u\|_{1,p}$.*

In a geodesic space we obtain the complete \mathbb{R}^n result:

Theorem 6. *Suppose that X satisfies the annulus condition and is also a geodesic space. If $p > 1$, $u \in M^{1,p}$, and g is a gradient of u , then $Mu \in M^{1,p}$, $\|Mu\|_{1,p} \leq C \|u\|_{1,p}$, and Mg is a gradient of u .*

We will outline the proof of the second result below. The full details, and the similar proof of the first result, can be found in [MM].

Some comments on the annulus condition are in order. An annulus is simply the region contained between two balls having a common centre. The thickness of the annulus is just the difference between the radii. Note that the thickness is not necessarily the distance between the inner and outer spheres of the annulus. One consequence of the condition is that balls can have no more than polynomial volume growth and so it fails on spaces such as the hyperbolic disc, which has constant negative curvature. It obviously holds on \mathbb{R}^n and it is easy to see that it holds on many other Riemannian manifolds of non-negative curvature.

In the Heisenberg group the situation is somewhat more subtle. There are a variety of different metrics on the group and balls have the same polynomial volume for all of these metrics. However, the shapes of the balls in the different metrics are very different. The annulus condition fails for any of the usual norm metrics but holds for the usual Carnot-Carathéodory metric on the group. To see that the annulus condition fails for the norm metrics, take A to be an annulus centered at the origin, of inner radius $1 - h^2$ and outer radius $1 + h^2$, and take B to be a ball of radius h centered on the t -axis at a distance 1 from the origin. Then A contains B and this violates the annulus condition. Verifying the annulus condition for the Carnot-Carathéodory metric is an extended calculation and can be found in [MM]. It reduces to showing that the surface formed by intersecting any sphere of large radius with the unit ball is a Euclidean Lipschitz surface with uniform Lipschitz constant. For many

reasons, the Carnot-Carathéodory metric is the natural metric to put on the Heisenberg group. In particular, it gives the group the structure of a geodesic space. Unfortunately, there is no explicit formula for the metric. This makes it difficult to work with and so the norm metrics are often used as replacements.

Suppose that $u \in M^{1,p}(X)$ and that g is a derivative of u . We can assume that $u \geq 0$. For Theorem 6 it is sufficient to show that $|Mu(x) - Mu(y)| \leq C|x - y|(Mg(x) + Mg(y))$ for a.e. x and y . This will follow from the uniform estimate below, where the constant C is independent of g :

$$(6) \quad \left| \frac{1}{\mu(B(x, R))} \int_{B(x, R)} u \, d\mu - \frac{1}{\mu(B(y, R))} \int_{B(y, R)} u \, d\mu \right| \leq C|x - y|(Mg(x) + Mg(y))$$

for all x and y and all $R > 0$.

Set $h = |x - y|$, $B_x = B(x, R/8)$, $B_y = B(y, R/8)$, $A_x =$ annulus centered at x of radius R and thickness h , and $A_y =$ annulus centered at y of radius R and thickness h . We only need to consider the case $R \gg 3h$ as the other case follows immediately from integrating the $M^{1,p}$ inequality. Some simple manipulation yields

$$\begin{aligned} & \left| \frac{1}{\mu(B(x, R))} \int_{B(x, R)} u \, d\mu - \frac{1}{\mu(B(y, R))} \int_{B(y, R)} u \, d\mu \right| \\ & \leq C \left(\frac{1}{\mu(B_x)} \int_{A_x} |u - u_{8B_x}| \, d\mu + \frac{1}{\mu(B_y)} \int_{A_y} |u - u_{8B_y}| \, d\mu \right. \\ & \quad \left. + \frac{\mu(A_y)}{\mu(B_y)} |u_{8B_x} - u_{8B_y}| \right). \end{aligned}$$

The first two terms are symmetric in x and y ; we will bound the first. The final term is easily bounded by $ChMg(x)$ using the annulus condition and the Poincaré inequality. In the same way $\frac{\mu(A_x)}{\mu(B_x)} |u_{8B_x} - u_{8B_y}|$ can be bounded by $ChMg(x)$, and thus it suffices to show that

$$\frac{1}{\mu(B_x)} \int_{A_x} |u - u_{B_x}| \, d\mu \leq ChMg(x).$$

For each $a \in A_x$ choose a shortest path γ_a joining x to a and define the cone C_a to be

$$\bigcup_{z \in \gamma_a} B\left(z, \frac{d(z, a)}{4}\right).$$

Each of these cones contains the ball B_x and each one is contained in $16B_x$. A simple chaining argument using balls centered on γ_a gives the following representation formula.

Lemma 7. *For a.e. $a \in A_x$ we have*

$$|u(a) - u_{B_x}| \leq C \int_{C_a} K(z, a)g(z) d\mu(z)$$

where

$$K(z, a) = \frac{d(z, a)}{\mu(B(z, d(z, a)))}.$$

The relationship between Poincaré inequalities and representation formulas such as that in the lemma is well known. See [FLW1], [FLW2], [FW].

As a consequence of the lemma we obtain the formula

$$\begin{aligned} & \frac{1}{\mu(B_x)} \int_{A_x} |u - u_{B_x}| d\mu \\ & \leq C \frac{1}{\mu(16B_x)} \int_{16B_x} g(z) \left(\int_{A_x} K(z, a) \chi_{C_a}(z) d\mu(a) \right) d\mu(z). \end{aligned}$$

There is a technical difficulty to be dealt with here: the function $\chi_{C_a}(z)$ may not be measurable. This is because there are no relationships among the geodesics joining x to the various points of A_x . This can be overcome by using any number of approximation methods. However, in the interests of brevity and clarity we ignore this difficulty.

Some straightforward, but tedious, calculations using the definitions and the fact that γ_a is a geodesic show that if $a \in A_x$ and $z \in C_a$, then $a \in B(z, r)$ for $r = 2(h + |R - d(z, x)|)$. Thus

$$\int_{A_x} K(z, a) \chi_{C_a}(z) d\mu(a) \leq \int_{A_x \cap B(z, r)} K(z, a) d\mu(a).$$

The integral of the kernel $K(z, a)$ over any ball centered at z is no more than a uniform multiple of the radius. Thus if $r \leq 8h$, the preceding integral is bounded by Ch . On the other hand, if $r > 8h$, then a calculation reveals that $B(z, r/4)$ does not intersect A_x . Therefore

$$\int_{A_x \cap B(z, r)} K(z, a) d\mu(a) \leq \int_{A_x \cap (B(z, r) \setminus B(z, r/4))} K(z, a) d\mu(a).$$

On the annulus the kernel is comparable to $r/\mu(B(z,r))$. It follows immediately from this and the annulus condition that once again the integral is bounded by Ch . We have now shown that

$$\int_{A_x} K(z,a)\chi_{C_a}(z) d\mu(a) \leq Ch$$

for every z in $16B_x$ and so

$$\begin{aligned} \frac{1}{\mu(B_x)} \int_{A_x} |u - u_{B_x}| d\mu &\leq C \frac{1}{\mu(2B_x)} \int_{2B_x} g(z) \left(\int_{A_x} K(z,a)\chi_{C_a}(z) d\mu(a) \right) d\mu(z) \\ &\leq Ch \frac{1}{\mu(2B_x)} \int_{2B_x} g(z) d\mu(z) \\ &\leq ChMg(x) \end{aligned}$$

as required.

References

- [BCLS] D. BAKRY, T. COULHON, M. LEDOUX AND L. SALOFF-COSTE, Sobolev inequalities in disguise, *Indiana Univ. Math. J.* **44(4)** (1995), 1033–1074.
- [BeS] C. BENNETT AND R. SHARPLEY, “*Interpolation of operators*”, Pure and Applied Mathematics **129**, Academic Press, Inc., Boston, MA, 1988.
- [Beu] A. BEURLING, On the spectral synthesis of bounded functions, *Acta Math.* **81** (1948), 14 pp.
- [Boj] B. BOJARSKI, Remarks on Sobolev imbedding inequalities, in: “*Complex analysis*” (Joensuu, 1987), Lecture Notes in Math. **1351**, Springer, Berlin, 1988, pp. 52–68.
- [Bom] J. BOMAN, L^p -estimates for very strongly elliptic systems, Report no. 29, Department of Mathematics, University of Stockholm, Sweden (1982).
- [Bus] P. BUSER, A note on the isoperimetric constant, *Ann. Sci. École Norm. Sup. (4)* **15(2)** (1982), 213–230.
- [Ch] J. CHEEGER, Differentiability of Lipschitz functions on metric measure spaces, *Geom. Funct. Anal.* **9(3)** (1999), 428–517.
- [Chua] S.-K. CHUA, Weighted Sobolev inequalities on domains satisfying the chain condition, *Proc. Amer. Math. Soc.* **117(2)** (1993), 449–457.

- [FGW] B. FRANCHI, C. E. GUTIÉRREZ AND R. L. WHEEDEN, Weighted Sobolev-Poincaré inequalities for Grushin type operators, *Comm. Partial Differential Equations* **19(3–4)** (1994), 523–604.
- [FL1] B. FRANCHI AND E. LANCONELLI, Hölder regularity theorem for a class of linear nonuniformly elliptic operators with measurable coefficients, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **10(4)** (1983), 523–541.
- [FL2] B. FRANCHI AND E. LANCONELLI, An embedding theorem for Sobolev spaces related to nonsmooth vector fields and Harnack inequality, *Comm. Partial Differential Equations* **9(13)** (1984), 1237–1264.
- [FLW1] B. FRANCHI, G. LU AND R. L. WHEEDEN, Representation formulas and weighted Poincaré inequalities for Hörmander vector fields, *Ann. Inst. Fourier (Grenoble)* **45(2)** (1995), 577–604.
- [FLW2] B. FRANCHI, G. LU AND R. L. WHEEDEN, A relationship between Poincaré-type inequalities and representation formulas in spaces of homogeneous type, *Internat. Math. Res. Notices* **1996(1)** (1996), 1–14.
- [FPW] B. FRANCHI, C. PÉREZ AND R. L. WHEEDEN, Self-improving properties of John-Nirenberg and Poincaré inequalities on spaces of homogeneous type, *J. Funct. Anal.* **153(1)** (1998), 108–146.
- [FW] B. FRANCHI AND R. L. WHEEDEN, Some remarks about Poincaré type inequalities and representation formulas in metric spaces of homogeneous type, *J. Inequal. Appl.* **3(1)** (1999), 65–89.
- [Ha] P. HAJLASZ, Sobolev spaces on an arbitrary metric space, *Potential Anal.* **5(4)** (1996), 403–415.
- [HaK1] P. HAJLASZ AND P. KOSKELA, Sobolev meets Poincaré, *C. R. Acad. Sci. Paris Sér. I Math.* **320(10)** (1995), 1211–1215.
- [HaK2] P. HAJLASZ AND P. KOSKELA, Sobolev met Poincaré, *Mem. Amer. Math. Soc.* **145(688)** (2000), 101 pp.
- [HKM] J. HEINONEN, T. KILPELÄINEN AND O. MARTIO, “*Nonlinear potential theory of degenerate elliptic equations*”, Oxford Mathematical Monographs, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1993.
- [HeK] J. HEINONEN AND P. KOSKELA, Quasiconformal maps in metric spaces with controlled geometry, *Acta Math.* **181(1)** (1998), 1–61.
- [Je] D. JERISON, The Poincaré inequality for vector fields satisfying Hörmander’s condition, *Duke Math. J.* **53(2)** (1986), 503–523.

- [Ka] A. KAŁAMAJSKA, On compactness of embedding for Sobolev spaces defined on metric spaces, *Ann. Acad. Sci. Fenn. Math.* **24(1)** (1999), 123–132.
- [Ki1] J. KINNUNEN, The Hardy-Littlewood maximal function of a Sobolev function, *Israel J. Math.* **100** (1997), 117–124.
- [Ki2] J. KINNUNEN AND P. LINDQVIST, The derivative of the maximal function, *J. Reine angew. Math.* **503** (1998), 161–167.
- [KS] N. J. KOREVAAR AND R. M. SCHOEN, Sobolev spaces and harmonic maps for metric space targets, *Comm. Anal. Geom.* **1(3–4)** (1993), 561–659.
- [KM] P. KOSKELA AND P. MACMANUS, Quasiconformal mappings and Sobolev spaces, *Studia Math.* **131(1)** (1998), 1–17.
- [LN] R. L. LONG AND F. S. NIE, Weighted Sobolev inequality and eigenvalue estimates of Schrödinger operators, in: “*Harmonic analysis*” (Tianjin, 1988), Lecture Notes in Math. **1494**, Springer, Berlin, 1991, pp. 131–141.
- [LoW] E. P. LOTKOWSKI AND R. L. WHEEDEN, The equivalence of various Lipschitz conditions on the weighted mean oscillation of a function, *Proc. Amer. Math. Soc.* **61(2)** (1976), 323–328.
- [Lu] G. LU, The sharp Poincaré inequality for free vector fields: an endpoint result, *Rev. Mat. Iberoamericana* **10(2)** (1994), 453–466.
- [LW] G. LU AND R. L. WHEEDEN, High order representation formulas and embedding theorems on stratified groups and generalizations, *Studia Math.* **142(2)** (2000), 101–133.
- [MM] P. MACMANUS, The maximal function and Sobolev spaces, Preprint.
- [MP1] P. MACMANUS AND C. PÉREZ, Generalized Poincaré inequalities: sharp self-improving properties, *Internat. Math. Res. Notices* **1998(2)** (1998), 101–116.
- [MP2] P. MACMANUS AND C. PÉREZ, Trudinger inequalities without derivatives, *Trans. Amer. Math. Soc.* **354(5)** (2002), 1997–2012.
- [Ma] V. G. MAZ’JA, “*Sobolev spaces*”, Translated from the Russian by T. O. Shaposhnikova, Springer Series in Soviet Mathematics, Springer-Verlag, Berlin, 1985.
- [NSW] A. NAGEL, E. M. STEIN AND S. WAINGER, Balls and metrics defined by vector fields. I. Basic properties, *Acta Math.* **155(1–2)** (1985), 103–147.
- [S-C] L. SALOFF-COSTE, A note on Poincaré, Sobolev, and Harnack inequalities, *Internat. Math. Res. Notices* **1992(2)** (1992), 27–38.

- [Se] S. SEMMES, Finding curves on general spaces through quantitative topology, with applications to Sobolev and Poincaré inequalities, *Selecta Math. (N.S.)* **2(2)** (1996), 155–295.
- [Sh] N. SHANMUGALINGAM, Newtonian spaces: an extension of Sobolev spaces to metric measure spaces, *Rev. Mat. Iberoamericana* **16(2)** (2000), 243–279.
- [Va] N. TH. VAROPOULOS, Fonctions harmoniques sur les groupes de Lie, *C. R. Acad. Sci. Paris Sér. I Math.* **304(17)** (1987), 519–521.
- [VSC] N. TH. VAROPOULOS, L. SALOFF-COSTE AND T. COULHON, “*Analysis and geometry on groups*”, Cambridge Tracts in Mathematics **100**, Cambridge University Press, Cambridge, 1992.

Phillips Exeter Academy
Exeter, NH 03833
U.S.A.
E-mail address: pmacmanus@exeter.edu