Canad. J. Math. Vol. **66** (4), 2014 pp. 721–742
http://dx.doi.org/10.4153/CJM-2012-064-x

On Whitney-type Characterization of Approximate Differentiability on Metric Measure Spaces

E. Durand-Cartagena, L. Ihnatsyeva, R. Korte, and M. Szumańska

Abstract. We study approximately differentiable functions on metric measure spaces admitting a Cheeger differentiable structure. The main result is a Whitney-type characterization of approximately differentiable functions in this setting. As an application, we prove a Stepanov-type theorem and consider approximate differentiability of Sobolev, BV, and maximal functions.

1 Introduction

A classical theorem of Luzin states that a measurable function which is finite almost everywhere coincides with a continuous function outside a set of arbitrary small measure. A function with such a property is said to satisfy the *Luzin property of order zero*. The reverse implication in Luzin's theorem also holds true and thus the Luzin property actually characterizes measurable functions. With the aid of the Lebesgue differentiation theorem, one can see that a function defined on $\mathbb{R}^n$ has the Luzin property of order zero if and only if it is approximately continuous almost everywhere. This characterization is known as Denjoy–Luzin theorem; see [13, 33].

For more regular functions, it is natural to expect Luzin properties of higher order. Indeed, Whitney [41] proved that approximately differentiable functions are precisely the functions that have the Luzin property of order one, in the sense that they are smooth on "nearly" all of their domain.

The concept of approximate continuity makes perfect sense for functions defined on arbitrary metric measure spaces. The same reasoning as in the Euclidean case shows that the Denjoy–Luzin theorem holds true for metric spaces equipped with a doubling measure; see Theorem 2.4. Our aim is to extend the Whitney theorem to a more general setting. Recently, there has been intensive research, where a first order differential calculus has been developed on metric measure spaces. For a general introduction to the subject, we mention here the survey works by Heinonen [21, 22], Heinonen–Koskela [20], Ambrosio–Tilli [3], Hajłasz–Koskela [18], Semmes [38],

Received by the editors September 18, 2012; revised December 21, 2012.
Published electronically February 8, 2013.
The first author was partially supported by grant MTM2009-07848 (Spain). The second, the third and the fourth authors were supported by the Academy of Finland (grants 252293, 250403 and 138738). The fourth author was partially supported by MNiSW Grant no N N201 397737, *Nonlinear partial differential equations: geometric and variational problems.*
AMS subject classification: **26B05**, 28A15, 28A75, 46E35.
Keywords: approximate differentiability, metric space, strong measurable differentiable structure, Whitney theorem.

and Björn–Björn [7]. The standard assumptions, which allow the first order differential calculus, include that the measure is doubling and that the space supports a p-Poincaré inequality.

Cheeger [12] constructed a measurable differentiable structure for the above mentioned class of metric spaces (see also Keith [23]) in such a way that Lipschitz functions can be differentiated almost everywhere with respect to this differentiable structure (Rademacher's theorem). Cheeger's differentiable structure provides a means to study approximate differentiability in metric measure spaces. The concept of approximate differentiability in this setting has been already considered by Keith [24] and by Bate and Speight [6]. See also Basalaev and Vodopyanov [5] for the study of approximate differentiability and Whitney-type theorems in the sub-Riemannian setting.

Here, we consider the class of approximately differentiable functions in spaces that admit a Cheeger differentiable structure. The main result of this paper, Theorem 3.1, is a Whitney-type characterization of approximate differentiability in the metric setting. The Whitney theorem is interesting in its own right as a classical result of real analysis, but the characterization of the Luzin property has also been used, for example, to prove regularity properties of different function spaces, especially when differentiability is not always guaranteed. This is the case, for example, for Sobolev and BV functions. For approximate differentiability properties of Sobolev and BV functions in the Euclidean case, one can consult [15].

We apply our main result in three different directions. The first one is related to Stepanov's theorem, and the other two are in connection with differentiability properties of Sobolev functions and the discrete Hardy–Littlewood maximal function.

The Stepanov theorem [40] states that a function is differentiable on the set of points where a certain local growth condition holds. We prove an approximate version of Stepanov's theorem (Theorem 3.3) in the metric setting and use it to give another characterization of approximate differentiability (Corollary 3.4) and an alternative proof for Stepanov's theorem by Balogh–Rogovin–Zürcher [4]. Our methods follow the lines of proof of the classical approximate Stepanov theorem in [16].

The obtained characterizations allow us to give a simple proof of the approximate differentiability for Sobolev functions (Hajłasz–Sobolev spaces [17] and Newtonian spaces [39]) and BV functions (Miranda [36]) in the metric setting; see Corollary 4.1. Notice that approximate differentiability properties can be also deduced from existing results. See Björn [9] and Ranjbar–Motlagh [37].

To finish, we use the Whitney type characterization to show in Theorem 5.1 that the notion of approximate differentiability in metric spaces is preserved under the action of the discrete maximal operator. The analogous statement for the regular Hardy–Littlewood maximal operator in the Euclidean setting was proved earlier by Hajłasz–Malý in [19]. Buckley [11] has shown that for a metric space with a doubling measure, the maximal operator may not preserve Lipschitz and Hölder spaces. Therefore some Lipschitz-type estimates that were used to prove the approximate continuity in Euclidean spaces do not hold in more general spaces. In order to have a maximal function that preserves, for example, the Sobolev spaces on metric spaces, Kinnunen and Latvala [26] used the discrete maximal function. Notice that in many applications the Hardy–Littlewood maximal operator can be replaced by the discrete maximal operator, as they are comparable by two-sided estimates [26].

Luzin properties of order k for $k > 1$ have been studied by Bojarski [10], Liu [29], and Liu–Tai [30, 31] in the Euclidean setting. See also [15]. In this paper, we only consider Luzin properties of order 1, since the theory for higher order derivatives has not been developed yet in the metric setting. However, it would be interesting to extend these results to higher order cases at least for lower dimensional subsets of $\mathbb{R}^n$.

The paper is organized as follows. In Section 2, we first briefly recall the concepts of approximate continuity and approximate differentiability in the Euclidean setting. After that we give some standard notation and relevant notions regarding metric spaces supporting a doubling measure that enable us to define approximate differentiability in this more general context. Section 3 contains the main result of this paper: a Whitney-type characterization of approximately differentiable functions in this setting as well as a Stepanov-type characterization. In Section 4, we use the obtained characterizations to show the approximate differentiability for Sobolev and BV functions. In the final Section 5, we prove that approximate differentiability a.e. is preserved under the action of the discrete maximal operator.

2 Preliminaries

2.1 Approximate Differentiability in $\mathbb{R}^n$

We say that $l \in \mathbb{R}$ is the *approximate limit* of a function $f \colon \mathbb{R}^n \to \mathbb{R}$ as $y \to x$, and write

$$\operatorname{ap}\lim_{y\to x} f(y) = l,$$

if for every $\varepsilon > 0$, $x \in \mathbb{R}^n$ is a density point for the set $\{y : |f(y) - l|) < \varepsilon\}$.

Observe that equivalently we can formulate the definition in the following way. There exists $A \subset \mathbb{R}^n$ with x a point of density for A such that

$$\lim_{\substack{y\to x\\ y\in A}} |f(y) - l| = 0.$$

If the approximate limit l exists and $f(x) = l$, then we say that f is *approximately continuous* at x.

Using the notion of approximate limit one can define the approximate differential.

Definition 2.1 Let $E \subset \mathbb{R}^n$ and $f \colon E \to \mathbb{R}$. We say that f is *approximately differentiable* at $x \in E$ if there exists a vector $L = (L_1, \dots, L_n)$ such that

$$\operatorname{ap}\lim_{y\to x} \frac{|f(y) - f(x) - L\cdot(y-x)|}{|y-x|} = 0.$$

Approximate differentiability is a much weaker notion than differentiability. The function $f \colon [0,1] \to \mathbb{R}$, $f(x) = 1$ if $x \in \mathbb{R}\backslash\mathbb{Q}$ and $f(x) = 0$ if $x \in \mathbb{Q}$ is approximately differentiable almost everywhere but nowhere differentiable. On the other hand, even a continuous function might be approximately differentiable almost nowhere; see, for example, [35].

The following characterization of approximate differentiability was given by Whitney in [41]. See also [16, Theorem 3.1.8].

Theorem 2.2 *Let $E \subset \mathbb{R}^n$ be a Lebesgue measurable set and let $f\colon E \to \mathbb{R}$ be a $\mathscr{L}^n$-measurable function. Then the following conditions are equivalent:*

(i) *f is approximately differentiable $\mathscr{L}^n$-a.e.;*

(ii) *f has a Lipschitz Luzin approximation, that is, for any $\varepsilon > 0$ there is a closed set $F \subset E$ and a locally Lipschitz function $g\colon \mathbb{R}^n \to \mathbb{R}$ such that $f_{|F} = g_{|F}$ and $\mathscr{L}^n(E \setminus F) < \varepsilon$;*

(iii) *f has a smooth Luzin approximation, that is, for any $\varepsilon > 0$ there is a closed set $F \subset E$ and a function $g \in \mathcal{C}^1(\mathbb{R}^n)$ such that $f_{|F} = g_{|F}$ and $\mathscr{L}^n(E \setminus F) < \varepsilon$;*

(iv) *f induces the decomposition*

$$E = \bigcup_{i=1}^{\infty} E_i \cup Z,$$

where E_i are pairwise disjoint closed sets, $f_{|E_i}$ is Lipschitz continuous, and Z has measure zero.

2.2 Approximate Differentiability in Metric Measure Spaces

Our main aim is to extend the statement of Whitney's theorem (Theorem 2.2) to the more general setting of a metric measure space. To formulate a definition of approximate differentiability in such a setting, we employ the ideas of Cheeger [12], who extended the fundamental notions of first order differential calculus to a general class of metric spaces. We start with several standard definitions.

Throughout the paper (X, d, μ) refers to a *metric measure space*, where (X, d) is a *separable* metric space and μ is a Borel regular measure such that $0 < \mu(B) < \infty$ for every ball $B \subset X$.

For $x \in X$ and $r > 0$ we denote by $B(x, r) := \{y \in X : d(x, y) < r\}$ the open ball of radius r centered at x.

One of the natural assumptions imposed on the measure is the doubling condition.

Definition 2.3 A measure μ on X is called *doubling* if there is a positive constant C_μ such that

$$\mu\big(B(x, 2r)\big) \leq C_\mu\, \mu\big(B(x, r)\big)$$

for each $x \in X$ and $r > 0$.

Recall that a point $x \in X$ is a *density point* for a μ-measurable set $A \subset X$ if

$$\lim_{r\to 0} \frac{\mu(B(x, r) \cap A)}{\mu(B(x, r))} = 1.$$

The following theorem gives a characterization of approximate continuity in the metric setting and gives an interpretation of the notion of "0-smoothness". For a proof of the theorem in the Euclidean setting, see [16] or [15]. See also [10] for a nice discussion of the role of Luzin–Denjoy theorem.

https://doi.org/10.4153/CJM-2012-064-x Published online by Cambridge University Press

Theorem 2.4 *Let (X, d, μ) be a metric measure space with μ-doubling. Let $E \subset X$ be a bounded μ-measurable set and $f: X \to \mathbb{R}$. The following conditions are equivalent:*

(i) *f is μ-measurable on E;*

(ii) *f is approximately continuous μ-a.e. in E;*

(iii) *f is quasicontinuous, that is, for each $\varepsilon > 0$ there is a closed set $F \subset E$ with $\mu(E \setminus F) < \varepsilon$ and $f_{|F}$ is continuous. In other words, f has a Luzin approximation of order zero;*

(iv) *f induces a (zero order) Luzin decomposition of E, that is,*

$$E = \bigcup_{i=1}^{\infty} E_i \cup Z,$$

where E_i are closed sets such that $f_{|E_i}$ is continuous and Z has measure zero.

We do not give a complete proof, since it follows the lines of the classical setting. One just needs to have in mind that the Lebesgue differentiation theorem holds in spaces equipped with a doubling measures (see [21, 1.8]). The equivalence of (i) and (ii) is shown by Federer [16, Theorem 2.9.13]. By Luzin's theorem, (i) implies (iii); see [16, Theorem 2.3.5]. The implications (iii) $\Rightarrow$ (iv) and (iv) $\Rightarrow$ (ii) can be shown following the lines of the second and the third parts of the proof of Theorem 3.1 with certain modifications; for instance, to show that (iv) implies (ii), one should apply Tietze's extension theorem instead of McShane's.

The structure of metric spaces endowed with a doubling measure has turned out to be too weak to develop a first order differential calculus involving derivatives, and therefore extra conditions are needed. The following Poincaré inequality creates a link between the measure, the metric, and the upper-gradient, and is ubiquitous in analysis on metric spaces. We recall that a non-negative Borel function g on X is an *upper gradient* of an extended real-valued function f on X if $|f(\gamma(a)) - f(\gamma(b))| \leq \int_\gamma g\, ds$ for all rectifiable curves $\gamma\colon [a, b] \to X$. We interpret the above inequality as also requiring that $\int_\gamma g\, ds = \infty$ whenever at least one of $f(\gamma(a)), f(\gamma(b))$ is not finite.

Definition 2.5 Let $1 \leq p \leq \infty$. We say that (X, d, μ) supports a *weak p-Poincaré inequality* if there are constants $\lambda_p \geq 1$ and $C_p > 0$ such that when $f: X \to \mathbb{R} \cup \{-\infty, \infty\}$ is a measurable function, $g: X \to [0, \infty]$ is an upper gradient of f, and $B(x, r)$ is a ball in X,

$$\text{(2.1)} \qquad \fint_{B(x,r)} |f - f_{B(x,r)}|\, d\mu \leq C_p\, r \Big(\fint_{B(x,\lambda_p r)} g^p\, d\mu \Big)^{1/p}$$

if $1 \leq p < \infty$, and

$$\fint_{B(x,r)} |f - f_{B(x,r)}|\, d\mu \leq C_\infty\, r \|g\|_{L^\infty(B(x,\lambda_\infty r))}$$

if $p = \infty$.

Here and everywhere below we write

$$f_A = \fint_A f := \frac{1}{\mu(A)} \int_A f\, d\mu,$$

https://doi.org/10.4153/CJM-2012-064-x Published online by Cambridge University Press

where $A \subset X$ and $0 < \mu(A) < \infty$.

We now recall the following theorem of Cheeger [12], which states that a metric space equipped with a doubling measure and having a p-Poincaré inequality admits a certain differentiable structure for which Lipschitz functions are differentiable μ-a.e.

Theorem 2.6 *Let X be a metric space with a doubling measure μ, and suppose that X supports a weak p-Poincaré inequality for some $1 \le p < \infty$. Then there exists a countable collection $\{(X_\alpha, \mathbf{x}_\alpha)\}_{\alpha \in \mathbf{\Lambda}}$ of measurable sets $X_\alpha \subset X$ and Lipschitz coordinates*

$$\mathbf{x}_\alpha = (x_\alpha^1, \dots, x_\alpha^{N(\alpha)}) \colon X \longrightarrow \mathbb{R}^{N(\alpha)}$$

with the following properties:

(i) $\mu(X \setminus \bigcup_\alpha X_\alpha) = 0$;
(ii) *There exists $N \ge 0$ such that $N(\alpha) \le N$ for each $(X_\alpha, \mathbf{x}_\alpha)$;*
(iii) *If $f \colon X \to \mathbb{R}$ is Lipschitz, then for each $(X_\alpha, \mathbf{x}_\alpha)$ there exists a unique (up to a set of zero measure) measurable bounded vector valued function $d^\alpha f \colon X_\alpha \to \mathbb{R}^{N(\alpha)}$ such that*

$$\lim_{\substack{y \to x \\ y \neq x}} \frac{|f(y) - f(x) - d^\alpha f(x) \cdot (\mathbf{x}_\alpha(y) - \mathbf{x}_\alpha(x))|}{d(y,x)} = 0 \tag{2.2}$$

for μ-a.e. $x \in X_\alpha$.

If a metric measure space (X, d, μ) satisfies the conclusion of Theorem 2.6, we say that the space admits a *strong measurable differentiable structure*. In particular, $\{(X_\alpha, \mathbf{x}_\alpha)\}_{\alpha \in \mathbf{\Lambda}}$ is said to be a strong measurable differentiable structure for (X, d, μ).

Notice that although the exponent p is present in the hypothesis of this result, it has no role in the conclusions. Keith weakened the hypotheses using the Lip-lip condition, formulated as follows. There exists a constant $K \ge 1$ such that

$$\operatorname{Lip} f(x) \le K \operatorname{lip} f(x)$$

for all Lipschitz functions $f \colon X \to \mathbb{R}$ and for μ-almost every $x \in X$, where $\operatorname{Lip} f$ and $\operatorname{lip} f$ denote the upper and lower scaled oscillation functions respectively. This Lip-lip condition is satisfied by any complete metric space endowed with a doubling measure that admits a p-Poincaré inequality for some $1 \le p < \infty$; see [23].

See [27] for an accessible introduction to the basis of the theory of differentiable structures.

The existence of the differentiable structure allows us to consider the following notion of differentiability of a function.

Definition 2.7 A function $f \colon X \to \mathbb{R}$ is *Cheeger differentiable* at a point $x \in X_\alpha$ with respect to the strong measurable differentiable structure $\{(X_\alpha, \mathbf{x}_\alpha)\}_{\alpha \in \mathbf{\Lambda}}$ if there exists a unique vector (*Cheeger differential*) $d^\alpha f(x) \in \mathbb{R}^{N(\alpha)}$ such that (2.2) holds for f at x.

Notice that the definition of the differentiable structure implies that the uniqueness of the Cheeger differential can be inferred from its existence almost everywhere on X. The exceptional set depends only on the differentiable structure.

Analogously, one can introduce a notion of approximate differentiability of a function defined on a metric space. See [6, 24].

Definition 2.8 If a metric measure space (X, d, μ) satisfies the conclusion of Theorem 2.6, where the limit in (2.2) is replaced with the approximate limit, it is said that the space admits *an approximate differentiable structure* (or *a measurable differentiable structure*).

Recently, Bate and Speight [6] have proved that if a metric measure space admits a strong measurable differentiable structure, then the measure is pointwise doubling almost everywhere. They also gave an example showing that if one only requires an approximate differentiable structure, the measure does not need to be pointwise doubling.

Definition 2.9 Let (X, d, μ) be a metric measure space that supports an approximate differentiable structure $\{(X_\alpha, \mathbf{x}_\alpha)\}_{\alpha\in\mathbf{\Lambda}}$. A function $f\colon X \to \mathbb{R}$ is *approximately differentiable* at $x \in X_\alpha$ with respect to $(X_\alpha, \mathbf{x}_\alpha)$ if there exists a vector $L^\alpha f(x) \in \mathbb{R}^{N(\alpha)}$ *(approximate differential)* such that

$$\text{(2.3)} \qquad \operatorname*{ap\,lim}_{y\to x} \frac{|f(y) - f(x) - L^\alpha f(x)(\mathbf{x}_\alpha(y) - \mathbf{x}_\alpha(x))|}{d(x, y)} = 0;$$

i.e., for every $\varepsilon > 0$ the set

$$\text{(2.4)} \qquad A_{x,\varepsilon} = \Big\{ y : \frac{|f(y) - f(x) - L^\alpha f(x) \cdot (\mathbf{x}_\alpha(y) - \mathbf{x}_\alpha(x))|}{d(x, y)} < \varepsilon \Big\}$$

has x as a density point.

The following lemma shows that the approximate differential is well defined, in the sense that if there exists such vector $L^\alpha f(x)$ satisfying (2.3) then it is unique for almost all points $x \in X_\alpha$. Thus, redefining (if necessary) the given measurable differentiable structure on a set of measure zero, we get the structure with respect to which the approximate differential is always unique.

Lemma 2.10 *Let $\{(X_\alpha, \mathbf{x}_\alpha)\}_{\alpha\in\mathbf{\Lambda}}$ be an approximate differentiable structure defined on a metric measure space (X, d, μ). Then for every $\alpha \in \mathbf{\Lambda}$ one can choose a subset $\widetilde{X}_\alpha \subset X_\alpha$ such that $\mu(X_\alpha \setminus \widetilde{X}_\alpha) = 0$ and for any function $f\colon X \to \mathbb{R}$ and every $x \in \widetilde{X}_\alpha$ the following statement is true: if there exist vectors $L_1^\alpha f(x)$, $L_2^\alpha f(x) \in \mathbb{R}^{N(\alpha)}$ such that*

$$\text{(2.5)} \qquad \operatorname*{ap\,lim}_{y\to x} \frac{|f(y) - f(x) - L_i^\alpha f(x) \cdot (\mathbf{x}_\alpha(y) - \mathbf{x}_\alpha(x))|}{d(x, y)} = 0, \quad i = 1, 2,$$

then $L_1^\alpha f(x) = L_2^\alpha f(x)$.

https://doi.org/10.4153/CJM-2012-064-x Published online by Cambridge University Press

Proof The definition of the approximate differentiable structure implies that the function $g \equiv 0$ has a unique approximate differential on a set $\widetilde{X}_\alpha$ satisfying $\mu(X_\alpha \setminus \widetilde{X}_\alpha) = 0$.

Assume that there is a function $f\colon X \to \mathbb{R}$ and a point $x \in \widetilde{X}_\alpha$ such that two different vectors $L_1^\alpha f(x)$ and $L_2^\alpha f(x)$ satisfy (2.5). By the definition of the approximate limit, there exist sets $A_1, A_2 \subset X_\alpha$ for which x is a density point and

$$\lim_{\substack{y\to x\\ y\in A_i}} \frac{|f(y) - f(x) - L_i^\alpha f(x)\cdot(\mathbf{x}_\alpha(y) - \mathbf{x}_\alpha(x))|}{d(y,x)} = 0,$$

for $i = 1, 2$. By the triangle inequality, we have that

$$\lim_{\substack{y\to x\\ y\in A_1\cap A_2}} \frac{|(L_1^\alpha f - L_2^\alpha f)(x)\cdot(\mathbf{x}_\alpha(y) - \mathbf{x}_\alpha(x))|}{d(y,x)} = 0.$$

Since $\widetilde{X}_\alpha$ is a set where $g \equiv 0$ has a unique approximate differential, we have

$$(L_1^\alpha f - L_2^\alpha f)(x) = 0,$$

as required. ■

In what follows, we will prove that the approximate differential is a measurable function. We will need the following technical lemma. Recall that $\operatorname{ap\,lim\,sup}_{y\to x} F(y)$ is the infimum of the set of numbers $a \in \mathbb{R}$ for which the set $\{y \in X : F(y) > a\}$ has density zero at the point x.

Lemma 2.11 *Let (X, d, μ) be a metric measure space and let $g\colon X \times X \to \mathbb{R}$ be a $\mu \otimes \mu$-measurable function. Then $x \mapsto \operatorname{ap\,lim\,sup}_{y\to x} g(x,y)$ is μ-measurable.*

Proof The proof is an immediate adaptation of the proof of the analogous statement in the Euclidean case [16, 3.1.3(2)] and is based on the fact that for any $\mu \otimes \mu$-measurable set S and any fixed $\varepsilon, \delta > 0$, the set

$$\bigcap_{0<r<\delta} \Big\{ x \in X \mid \mu\big(\{y \mid (x,y) \in S,\ y \in B(x,r)\}\big) < \varepsilon\mu\big(B(x,r)\big) \Big\} \tag{2.6}$$

is μ-measurable. To prove the measurability of the set defined above, one needs to use the fact that for every $r > 0$ the function $f(x) = \mu(B(x,r))$ is lower semicontinuous and, hence, measurable.

To obtain the measurability of the function $x \mapsto \operatorname{ap\,lim\,sup}_{y\to x} g(x,y)$, it will be enough to use the above observation for the set $S := \{(x,y) \mid g(x,y) > t\}$ for any $t \in \mathbb{R}$. ■

Now we are ready to prove the measurability of the approximate differential.

Lemma 2.12 *Let (X, d, μ) be a metric measure space that supports an approximate differentiable structure. If $f\colon X \to \mathbb{R}$ is a measurable function that is approximately differentiable at μ-almost every $x \in X_\alpha$, then the approximate differential $L^\alpha f\colon X_\alpha \to \mathbb{R}^{N(\alpha)}$ is μ-measurable on X_α.*

Here the value $L^\alpha f(x)$ is given by Definition 2.9, if x is a point of approximate differentiability of f, and $L^\alpha f(x) = \mathbf{0}$ otherwise.

Proof To prove that the function $l = L^\alpha f$ is measurable, we show that $l^{-1}(K)$ is a measurable set for each compact $K \subset \mathbb{R}^{N(\alpha)}$. Let K be a compact set. Denote by

$$A_x(\lambda) = \operatorname{ap}\limsup_{y\to x} \frac{|f(y) - f(x) - \lambda\cdot(\mathbf{x}_\alpha(y) - \mathbf{x}_\alpha(x))|}{d(x,y)}, \quad \lambda \in \mathbb{R}^{N(\alpha)}.$$

Observe that, for every x, the function $\lambda \mapsto A_x(\lambda)$ is continuous. Indeed, since

$$\operatorname{ap}\limsup_{y\to x} |g(y) + h(y)| \le \operatorname{ap}\limsup_{y\to x} |g(y)| + \operatorname{ap}\limsup_{y\to x} |h(y)|,$$

we have

$$|A_x(\lambda) - A_x(\lambda')| \le \operatorname{ap}\limsup_{y\to x} \frac{|(\lambda - \lambda')\cdot(\mathbf{x}_\alpha(y) - \mathbf{x}_\alpha(x))|}{d(x,y)} \le C|\lambda - \lambda'|$$

for any $\lambda,\ \lambda' \in \mathbb{R}^{N(\alpha)}$. Set

$$E = \{x \in X_\alpha : \exists \lambda \in K \text{ such that } A_x(\lambda) = 0\}.$$

Note that since the approximate differential is unique, E coincides with $l^{-1}(K)$. To check that E is measurable, fix a dense countable subset K' of K. Then by the continuity of A_x and the density of K' in K, we have

$$E = \Big\{ x \in X_\alpha : \exists (\lambda_n)_{n\in\mathbb{N}} \subset K',\ \lambda \in K \text{ such that } \lambda_n \to \lambda \text{ and } \lim_{n\to\infty} A_x(\lambda_n) = 0 \Big\}.$$

Consequently, we can write E as

$$E = \bigcap_{n\in\mathbb{N}} \bigcup_{\lambda\in K'} \Big\{ x \in X_\alpha : A_x(\lambda) < \frac{1}{n} \Big\}.$$

To finish it remains to check that the function $x \mapsto A_x(\lambda)$ is measurable for each $\lambda \in \mathbb{R}^{N(\alpha)}$. This follows from Lemma 2.11. ■

Next observe that the notion of approximate differentiability does not depend on the choice of the approximate differentiable structure.

Lemma 2.13 *Let (X, d, μ) be a metric measure space that admits an approximate differentiable structure $\{(X_\alpha, \mathbf{x}_\alpha)\}_{\alpha\in\mathbf{\Lambda}}$. If $f\colon X \to \mathbb{R}$ is approximately differentiable μ-a.e. on X with respect to $\{(X_\alpha, \mathbf{x}_\alpha)\}_{\alpha\in\mathbf{\Lambda}}$, then it is approximately differentiable almost everywhere with respect to any approximate differentiable structure defined on X.*

https://doi.org/10.4153/CJM-2012-064-x Published online by Cambridge University Press

Proof Let $\{(X_\alpha, \mathbf{x}_\alpha)\}_{\alpha\in\boldsymbol{\Lambda}}$ and $\{(Y_\beta, \mathbf{y}_\beta)\}_{\beta\in\mathbf{B}}$ be two approximate differentiable structures defined on (X, d, μ). We will write $L^\alpha_{\mathbf{x}} f$ for the approximate differential of f with respect to $\{(X_\alpha, \mathbf{x}_\alpha)\}_{\alpha\in\boldsymbol{\Lambda}}$ at $x \in X_\alpha$.

First we notice that for fixed x the real valued function $g_x(\,\cdot\,) = L^\alpha_{\mathbf{x}} f(x)\cdot \mathbf{x}_\alpha(\,\cdot\,)$ is Lipschitz continuous on X, and thus it is approximately differentiable μ-a.e. with respect to $\{(Y_\beta, \mathbf{y}_\beta)\}_{\beta\in\mathbf{B}}$. Moreover the set of points where g_x is approximately differentiable does not depend on the choice of x.

Let $\beta \in \mathbf{B}$. For almost every $x \in X_\beta$ we can choose $\alpha \in \boldsymbol{\Lambda}$ such that $x \in X_\alpha$ and x is a point of approximate differentiability of f with respect to $(X_\alpha, \mathbf{x}_\alpha)$. Thus for μ-a.e. $x \in X_\beta$, we have

$$\begin{aligned}\big| f(y) - f(x) - L^\beta_{\mathbf{y}} g_x(x)(\mathbf{y}_\beta(y) - \mathbf{y}_\beta(x))\big| &\le \big| f(y) - f(x) - L^\alpha_{\mathbf{x}} f(x)(\mathbf{x}_\alpha(y) - \mathbf{x}_\alpha(x))\big| \\ &\quad + \big| g_x(y) - g_x(x) - L^\beta_{\mathbf{y}} g_x(x)(\mathbf{y}_\beta(y) - \mathbf{y}_\beta(x))\big|.\end{aligned}$$

Obviously the set

$$A_{x,\varepsilon} := \big\{\, y \in B(x, r) : \big| f(y) - f(x) - L^\beta_{\mathbf{y}} g_x(x)(\mathbf{y}_\beta(y) - \mathbf{y}_\beta(x))\big| < \varepsilon d(x, y)\big\}$$

contains the intersection of the sets

$$\Big\{\, y \in B(x, r) : \big| f(y) - f(x) - L^\alpha_{\mathbf{x}} f(x)(\mathbf{x}_\alpha(y) - \mathbf{x}_\alpha(x))\big| < \frac{\varepsilon}{2} d(x, y)\Big\}$$

and

$$\Big\{\, y \in B(x, r) : \big| g_x(y) - g_x(x) - L^\beta_{\mathbf{y}} g_x(x)(\mathbf{y}_\beta(y) - \mathbf{y}_\beta(x))\big| < \frac{\varepsilon}{2} d(x, y)\Big\}.$$

Therefore each x that is a point of approximate differentiability of f with respect to $(X_\alpha, \mathbf{x}_\alpha)$ and a point of approximate differentiability of $L^\alpha_{\mathbf{x}} f(x)\cdot \mathbf{x}_\alpha$ with respect to $(Y_\beta, \mathbf{y}_\beta)$ is a point of density of $A_{x,\varepsilon}$. We conclude that f is approximately differentiable a.e. on X with respect to $\{(Y_\beta, \mathbf{y}_\beta)\}_{\beta\in\mathbf{B}}$. ∎

Remark 2.14 Definition 2.9 of approximate differentiability makes sense whenever the underlying space supports an approximate differentiable structure. If we additionally assume that the measure μ is doubling, then an approximate differentiable structure turns out to be a strong measurable differentiable structure as well; see [24, Prop. 3.5]. Since the results presented later on are formulated under the assumption that the measure μ is doubling, we in fact deal with a strong measurable differentiable structure.

3 Characterization of Approximate Differentiability

3.1 Whitney-type Characterization of Approximate Differentiability

The proof of the following theorem is strongly inspired by the original proof of Whitney for the Euclidean case; see [41, Theorem 1].

https://doi.org/10.4153/CJM-2012-064-x Published online by Cambridge University Press

Theorem 3.1 *Let (X, d, μ) be a complete metric measure space, where μ is a doubling measure and let $\{(X_\alpha, \mathbf{x}_\alpha)\}_{\alpha\in\mathbf{\Lambda}}$ be an approximate differentiable structure on (X, d, μ). Suppose that $E \subset X$ and $f\colon E \to \mathbb{R}$ is a μ-measurable function. Then the following conditions are equivalent:*

(i) *f is approximately differentiable μ-a.e. in E;*
(ii) *for any $\varepsilon > 0$ there is a closed set $F \subset E$ such that $\mu(E \setminus F) < \varepsilon$ and $f_{|F}$ is locally Lipschitz;*
(iii) *f induces a Luzin decomposition of E, that is,*

$$E = \bigcup_{j=1}^{\infty} E_j \cup Z, \tag{3.1}$$

where E_i are pairwise disjoint measurable sets, $f_{|E_i}$ are Lipschitz functions and Z has measure zero.

Remark 3.2 When $E \subset X$ is a bounded set, condition (ii) can be replaced by:

(ii′) for any $\varepsilon > 0$ there exists a closed set $F \subset E$ and a Lipschitz function $g\colon X \to \mathbb{R}$ such that $\mu(E \setminus F) < \varepsilon$ and $f_{|F} = g$.

To show that the function f is globally Lipschitz on F, one needs to notice that, since X is proper, the set F is compact. Then one can extend f to the whole space by standard arguments; see, for example, [21, Theorem 6.2].

Proof of Theorem 3.1 Without loss of generality, we can consider all coordinate functions $\mathbf{x}_\alpha$ to be Lipschitz continuous with a Lipschitz constant equal to one, since clearly $\{(X_\alpha, \frac{\mathbf{x}_\alpha}{\mathrm{LIP}(\mathbf{x}_\alpha)})\}_{\alpha\in\mathbf{\Lambda}}$ is an approximate differentiable structure on X, and f is approximately differentiable with respect to the structure. Here $\mathrm{LIP}(\mathbf{x}_\alpha)$ denotes the Lipschitz constant of $\mathbf{x}_\alpha$.

Let f be approximately differentiable μ-a.e. in E. We can assume that the sets X_α are pairwise disjoint and extend $L^\alpha f$ by zero outside X_α. Denote by N the bound on the dimension given by Theorem 2.6. Consider $L^\alpha f(x)$ as vectors in $\mathbb{R}^N$ (we extend the vector with zeros when necessary) and let $Lf = \sum_\alpha L^\alpha f$. If a f function is Cheeger differentiable μ-a.e. on X, the analogue construction would give a "gradient" for f. This construction is quite standard in the literature; see *e.g.*, [8, 9].

(i) $\Longrightarrow$ (ii) First, assume that $E \subset X$ is a bounded set. Define

$$D = \{x \in E : f \text{ is approximately differentiable at } x\}.$$

First we show that for any $\varepsilon > 0$ there exists a closed set $F = F_\varepsilon \subset D$, $\delta > 0$ and $L > 0$ such that $\mu(D \setminus F) < \varepsilon$ and

$$|f(x) - f(y)| \le L\, d(x, y) \text{ for each } x, y \in F,\ d(x, y) < \delta.$$

Since μ is a doubling measure, we have for any $r > 0$, $x, y \in X$ such that $d(x, y) \le r/2$ that

$$\mu\big(B(x, r) \cap B(y, r)\big) \ge \mu\big(B(x, r/2)\big) \ge 2\, a\, \mu\big(B(x, r)\big), \tag{3.2}$$

https://doi.org/10.4153/CJM-2012-064-x Published online by Cambridge University Press

where $a = 1/2C_\mu$, and C_μ denotes the doubling constant.

For each $\eta > 0$ define the following sequence of functions:

$$\psi_i^\eta(x) = \mu\big(B(x,1/i) \setminus A_{x,\eta}\big) \quad x \in D, \quad i \in \mathbb{N},$$

where $A_{x,\eta}$ is given by formula (2.4). It is clear that for each $i \in \mathbb{N}$, the function $\psi_i^\eta(x)$ is measurable in x for fixed η. Moreover, for each $\eta > 0$ and $x \in D$ one has

$$\phi_i^\eta(x) = \frac{\psi_i^\eta(x)}{\mu(B(x,1/i))} \to 0 \quad \text{as} \quad i \to \infty.$$

Next set $\eta = 1$. By Luzin's and Egoroff's theorems there exists a closed set $F \subset E$ such that

(a) $\mu(E \setminus F) = \mu(D \setminus F) < \varepsilon$;
(b) $Lf_{|F}$ is continuous, moreover, since X is proper, $Lf_{|F}$ is bounded in F, *i.e.*, $|Lf_{|F}| \le C$;
(c) $\phi_i^1 \to 0$ uniformly on F.

Now choose i_0 such that

$$\phi_i^1(x) < \frac{a}{C_\mu}, \quad x \in F, \quad i \ge i_0, \tag{3.3}$$

where $C_\mu \ge 1$ is the doubling constant.

Fix $x, y \in F$ such that $d(x,y) < 1/(2i_0)$. Points x and y may belong to two different charts, thus we write $x \in X_\alpha$ and $y \in X_\beta$. Choose $i \ge i_0$ such that

$$1/(2i+2) < d(x,y) \le 1/(2i).$$

For such i we have that $i \ge i_0$ and by (3.3), we get that

$$\psi_i^1(y) < \frac{a}{C_\mu}\mu\big(B(y,1/i)\big) \le \frac{a}{C_\mu}\mu\big(B(x,2/i)\big) \le a\mu\big(B(x,1/i)\big).$$

Hence

$$\psi_i^1(x) < a\mu\big(B(x,1/i)\big) \quad \text{and} \quad \psi_i^1(y) < a\mu\big(B(x,1/i)\big). \tag{3.4}$$

Combining (3.2) and (3.4), we deduce that there exists a point $z \in B(x,1/i) \cap B(y,1/i)$ that does not belong to the union of $B(x,1/i) \setminus A_{x,1}$ and $B(y,1/i) \setminus A_{y,1}$. For such point z, we have that $d(x,z) < 1/i$, $d(y,z) < 1/i$,

$$\frac{|f(z) - f(x) - L^\alpha f(x) \cdot (\mathbf{x}_\alpha(z) - \mathbf{x}_\alpha(x))|}{d(z,x)} < 1 \tag{3.5}$$

and

$$\frac{|f(z) - f(y) - L^\beta f(y) \cdot (\mathbf{x}_\beta(z) - \mathbf{x}_\beta(y))|}{d(z,y)} < 1. \tag{3.6}$$

By combining (3.5), (3.6), $d(y,z) < 4d(x,y)$, and $d(x,z) < 4d(x,y)$ we obtain the inequality

$$\begin{aligned}|f(y)-f(x)| \le & \,\big| f(z)-f(x)-L^{\alpha} f(x)\cdot(\mathbf{x}_\alpha(z)-\mathbf{x}_\alpha(x))\big| \\ & + \big| f(z)-f(y)-L^{\beta} f(y)\cdot(\mathbf{x}_\beta(z)-\mathbf{x}_\beta(y))\big| \\ & + \big| L^{\alpha} f(x)\cdot(\mathbf{x}_\alpha(z)-\mathbf{x}_\alpha(x))\big| + \big| L^{\beta} f(y)\cdot(\mathbf{x}_\beta(z)-\mathbf{x}_\beta(y))\big| \\ \le & \,8d(x,y)+Cd(x,y),\end{aligned}$$

which shows that $f_{|F}$ is a locally Lipschitz function and finishes the proof of the implication for the case in which $E \subset X$ is a bounded set.

Let now E be an arbitrary subset of X. Fix any point $x_0 \in X$ and consider a family of open balls $B_j = B(x_0, j)$, $j = 1, 2, \dots$, covering X.

Apply now the above reasoning to get closed sets $F_j \subset E \cap B_j$ such that

$$\mu\big((E\cap B_j)\setminus F_j\big) \le 2^{-j}\varepsilon$$

and $f_{|F_j}$ are locally Lipschitz functions. Set

$$F = X \setminus \bigcup_{j=1}^{\infty}(B_j \setminus F_j),$$

then

$$\mu(E\setminus F) = \mu(E\cap F^c) = \mu\Big(E\cap \bigcup_{j=1}^{\infty}(B_j\setminus F_j)\Big) \le \sum_{j=1}^{\infty}\mu\big(E\cap(B_j\setminus F_j)\big) \le \varepsilon.$$

It is easy to see that F is a closed set and $F\cap B_j \subset F_j$ for every $j = 1, 2, \dots$. Hence, $F \subset E$ and $f_{|F}$ is locally Lipschitz. The last observation follows from the fact that for every point of F we can find a neighborhood of x contained in some B_j and $f_{|F_j}$ is locally Lipschitz.

(ii) $\Longrightarrow$ (iii) For each $i \in \mathbb{N}$ there exists a closed set F_i such that $\mu(E\setminus F_i) \le 1/i$ and $f_{|F_i}$ is locally Lipschitz. Setting $\widetilde{F}_k := \bigcup_{i=1}^{k} F_i$ we obtain an ascending family of closed sets, such that f is locally Lipschitz on each of its members. We define $Z := \bigcap_{k=1}^{\infty}(E\setminus \widetilde{F}_k)$. Then

$$\mu(Z) = \lim_{k\to\infty}\mu(E\setminus\widetilde{F}_k) = 0.$$

Let $\{B_i\}_{i=1}^{\infty}$ denote a countable family of balls covering X. Then

$$E = Z \cup \bigcup_{i=1}^{\infty}\bigcup_{k=1}^{\infty}\widetilde{F}_k\cap B_i = Z\cup\bigcup_{j=1}^{\infty}F'_j,$$

where the family $\{F'_j\}_{j=1}^{\infty}$ is obtained just by renumerating the family $\{F_k\cap B_i\}$. Observe that $f_{|F'_j}$ is Lipschitz. To finish the proof we take the disjoint sets as follows: $E_1 := F'_1$ and $E_j := F'_j \setminus \bigcup_{m<j} E_m$, $j > 1$.

(iii) $\Longrightarrow$ (i) If decomposition (3.1) holds, then the restriction $f_{|E_i}$ is Lipschitz for every i. Using McShane's theorem, we can extend $f_{|E_i}$ to a Lipschitz function $\widetilde{f_i}$ defined on the whole space X. By the definition of the approximate differentiable structure, $\widetilde{f_i}$ is μ-a.e. approximately differentiable on X. Since E_i is measurable, almost every point of E_i is one of its points of density. Therefore it follows that $\widetilde{f_i}$ (hence also f) is μ-a.e. approximately differentiable on E_i. ∎

3.2 Stepanov-type Characterization

The following Stepanov-type theorem shows that an approximate local growth condition on a function guarantees its approximation by Lipschitz functions in Luzin's sense.

Theorem 3.3 *Let μ be a doubling measure. A μ-measurable function $f\colon E \to \mathbb{R}$ defined on a measurable subset $E \subset X$ satisfies the condition*

$$\operatorname{ap}\limsup_{y\to x} \frac{|f(y)-f(x)|}{d(x,y)} < \infty \tag{3.7}$$

for μ-a.e. x in E if and only if for any $\varepsilon > 0$ there is a closed set $G \subset E$ such that $\mu(E \setminus G) \le \varepsilon$ and f is locally Lipschitz on G.

The proof is an adaptation of arguments in [16, Theorem 3.1.8] to the metric setting.

Proof First assume that condition (3.7) holds. Define for each positive integer j the set

$$Q(u,r,j) = B(u,r) \cap \left\{ x : x \notin E \text{ or } |f(x)-f(u)| > jd(x,u) \right\},$$

whenever $u \in E$ and $r > 0$. Define also the set

$$A_j = E \cap \left\{ u : \mu\big(Q(u,r,j)\big) < a\mu\big(B(u,r)\big) \text{ for } 0 < r < 1/j \right\},$$

where $a > 0$ is some constant for which (3.2) holds. Then each set A_j is measurable, which follows from the measurability of the sets defined by (2.6), and

$$\mu\Big(E \setminus \bigcup_{j=1}^{\infty} A_j\Big) = 0. \tag{3.8}$$

Observe that if $u, v \in A_j$ and $d(u,v) < 1/2j$, then

$$\big| f(u) - f(v) \big| \le 4jd(u,v).$$

Indeed, set $r = 2d(u,v)$, then by the definition of the sets Q and by inequality (3.2)

$$\mu\big(Q(u,r,j) \cup Q(v,r,j)\big) < a\big(\mu(B(u,r)) + \mu(B(v,r))\big) \le \mu\big(B(u,r) \cap B(v,r)\big).$$

Thus, we can choose a point $x \in \big(B(u,r) \cap B(v,r)\big) \setminus \big(Q(u,r,j) \cup Q(v,r,j)\big)$, and we have

$$\begin{aligned} \big|f(u) - f(v)\big| &\le \big|f(u) - f(x)\big| + \big|f(v) - f(x)\big| \\ &\le j\big(d(x,u) + d(x,v)\big) \le 2jr = 4jd(u,v). \end{aligned}$$

It follows from the last inequality that f is locally Lipschitz on every A_j. Since the sequence of sets A_j, $j = 1, 2, \dots$ is increasing, the measure μ is Borel regular and equality (3.8) holds, for any $\varepsilon > 0$ we can choose a closed set $G \subset E$ such that $\mu(E \setminus G) \le \varepsilon$ and $f|_G$ is a locally Lipschitz function.

Let us show the reverse implication. Let G be a closed set such that $\mu(E \setminus G) \le \varepsilon$ and $f|_G$ is a locally Lipschitz function. Then $X \setminus G$ has density zero at μ-almost every point of G. Thus, (3.7) holds μ-a.e. in G and the fact that ε can be chosen arbitrary small finishes the proof. ∎

As a corollary of Theorems 3.1 and 3.3, we get the following characterization of approximate differentiability.

Corollary 3.4 *Under the hypothesis of Theorem 3.1, a function $f\colon X \to \mathbb{R}$ is approximately differentiable μ-a.e. in a bounded measurable subset $E \subset X$ if and only if*

$$\operatorname{ap}\limsup_{y \to x} \frac{|f(y) - f(x)|}{d(x,y)} < \infty, \quad \mu\text{-a.e. in } E.$$

A similar integral local growth condition is used in [37] to guarantee L^p-differentiability of a function. It is also mentioned in [37, Remark 3.4] that the technique used in [37, Theorem 3.3] can be adapted for the notion of approximate differentiability.

As mentioned before, approximate differentiability is a much weaker property than differentiability. However, if it is the case that both the approximate differential and the Cheeger differential exist almost everywhere, they should coincide. Therefore it is interesting to search for additional conditions of global and infinitesimal character that imply Cheeger differentiability almost everywhere.

The following Stepanov differentiability theorem in metric measure spaces was proved by Balogh–Rogovin–Zürcher in [4].

Theorem 3.5 **[4]** *Let (X,d,μ) be a metric space endowed with a doubling measure μ. Assume that there is a strong measurable differentiable structure for (X,d,μ). Then a function $f\colon X \to \mathbb{R}$ is μ-a.e. Cheeger differentiable in the set*

$$\Big\{x\colon \limsup_{y \to x} \frac{|f(y) - f(x)|}{d(x,y)} < \infty\Big\}.$$

The proof of Theorem 3.5 is based on Malý's proof of Stepanov's theorem in the Euclidean case; see [34]. Note that the Stepanov differentiability theorem in $\mathbb{R}^n$ can also be derived from its approximate analogue; see *e.g.*, [16, Theorem 3.1.9]. The same arguments work in metric spaces. Thus, one can obtain an alternative proof for Theorem 3.5 combining Corollary 3.4 and the version of [16, Lemma 3.1.5] adapted to the metric measure setting.

4 Differentiability Properties of Sobolev Functions

In this section, we show that the approximate differentiability of Sobolev functions and BV functions follows easily from the Stepanov-type characterization of approximate differentiability. The results in this section are basically known, but our approach gives another point of view.

First let us notice that if we have a Lipschitz-type pointwise estimate for a function, then we have an approximate local growth condition on f as in Theorem 3.3. Namely, if $f\colon X \to \mathbb{R}$ is a μ-measurable function for which there exists a μ-measurable function $g\colon X \to \mathbb{R}$ and a set N of measure zero with

$$|f(x) - f(y)| \leq d(x,y)\big(g(x) + g(y)\big) \qquad \text{for every } x, y \in X \setminus N, \tag{4.1}$$

then

$$\operatorname{ap}\limsup_{y\to x} \frac{|f(y) - f(x)|}{d(x,y)} < \infty, \quad \mu\text{-a.e. in } X. \tag{4.2}$$

Indeed, notice first that by Theorem 2.4, g is approximately continuous μ-a.e. Divide both sides of the inequality (4.1) by $d(x,y)$ and take approximate supremum limits when $y \to x$ to get (4.2).

There are several generalizations of classical Sobolev spaces to the setting of arbitrary metric measure spaces.

Hajłasz–Sobolev spaces $M^{1,p}(X)$ were defined in [17] as the functions $f \in L^p(X)$ for which there exists a positive function $g \in L^p(X)$ satisfying inequality (4.1). It follows from the discussion above that under the hypothesis of Theorem 3.1, Hajłasz–Sobolev functions are approximately differentiable almost everywhere.

Using the notion of upper gradient (and more generally weak upper gradient), Shanmugalingam [39] introduced Newtonian spaces $N^{1,p}(X)$ for $1 \leq p \leq \infty$. A non-negative Borel function g on X is a *p-weak upper gradient* of an extended real-valued function f on X if $|f(\gamma(a)) - f(\gamma(b))| \leq \int_\gamma g$ for all rectifiable curves $\gamma : [a,b] \to X$ except for a family of zero p-modulus. See [39] for the definition of modulus of a family of curves.

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

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

where the infimum is taken over all p-weak upper gradients g of f. Now, we define in $\widetilde{N}^{1,p}(X,d,\mu)$ an equivalence relation by $f_1 \sim f_2$ if and only if $\|f_1 - f_2\|_{\widetilde{N}^{1,p}} = 0$. Then the space $N^{1,p}(X,d,\mu) = N^{1,p}(X)$ is defined as the quotient $\widetilde{N}^{1,p}(X,d,\mu)/\sim$.

As shown by Hajłasz and Koskela [18, Theorem 3.2], if we have a pair of functions (f,g) that satisfies a p-Poincaré inequality (2.1), then we have the following pointwise estimate

$$|f(x) - f(y)| \leq Cd(x,y)\Big[\big(M_{2\sigma d(x,y)} g^p(x)\big)^{1/p} + (M_{2\sigma d(x,y)} g^p(y))^{1/p}\Big], \tag{4.3}$$

for μ-a.e. $x, y \in X$ and for some constants $C, \sigma > 0$. Here $M_R f$ is defined by

$$M_R f(x) := \sup_{0<r\leq R} \fint_{B(x,r)} |f(y)| d\mu(y),$$

and notice that $M_R f(x) \leq Mf(x)$, where Mf is the standard Hardy–Littlewood maximal function. Actually, if the space supports a doubling measure and a p-Poincaré inequality, with $p > 1$, Newtonian spaces $N^{1,p}(X)$ are characterized by (4.3). Moreover, under these hypotheses, Newtonian spaces coincide with Hajłasz–Sobolev spaces; see [39]. If X is only known to be bounded, complete and supports an ∞-Poincaré inequality, the space of Lipschitz functions coincides with $N^{1,\infty}(X)$ (see [14, Theorem 4.7]).

Stepanov-type characterization can also be used to prove that BV functions on metric spaces are approximately differentiable almost everywhere. See the work by Miranda [36] for the corresponding definition of BV functions. Very recently, Lahti and Tuominen [28] have shown that a similar pointwise estimate as in (4.3) holds for BV functions assuming that the space is complete and supports a 1-Poincaré inequality. Namely, if $f \in BV(X)$, there exists a constant $\sigma \geq 1$ such that

$$\left| f(x) - f(y) \right| \leq C d(x,y) \left[M_{2\sigma d(x,y)} \|Df\|(x) + M_{2\sigma d(x,y)} \|Df\|(y) \right],$$

for μ-a.e. $x, y \in X$, where C is a constant depending only on the doubling constant and the constants involved in the Poincaré inequality. Here $M_{2\sigma d(x,y)}\|Df\|$ denotes the restricted maximal function of the measure $\|Df\|$; that is,

$$M_R \|Df\|(x) := \sup_{0<r\leq R} \frac{\|Df\|(B(x,r))}{\mu(B(x,r))},$$

where $\|Df\|$ denotes the total variation of the measure μ.

Corollary 4.1 *Assume that (X, d, μ) is complete doubling metric measure space. Then a function f is approximately differentiable μ-a.e. in each of the following cases:*

- *X supports an approximate differentiable structure and $f \in M^{1,p}(X)$, for some $p \geq 1$;*
- *X supports the p-Poincaré inequality and $f \in N^{1,p}(X)$ for some $p \geq 1$;*
- *X supports the 1-Poincaré inequality and $f \in BV(X)$.*

Notice that the assumption of a Poincaré inequality cannot be dropped from the hypothesis in the Newtonian case and in the BV case. For example, if the space has no rectifiable curves, except for the constant ones, then $N^{1,p}(X) = L^p(X)$, and therefore it could happen that a function in $N^{1,p}(X)$ is nowhere differentiable, nor approximately differentiable. On the other hand, when one uses Hajłasz approach, it is enough to assume that the space admits a differentiable structure to reach the conclusion.

The results in Corollary 4.1 can be also deduced from existing results in literature. Björn [9] has shown that if (X, d, μ) is doubling and supports a p-Poincaré inequality, then for each function $f \in N^{1,p}(X)$ and μ-a.e. $x \in X$,

$$\limsup_{r\to 0} \frac{1}{r} \fint_{B(x,r)} \left| f(y) - f(x) - df^{\alpha}(x) \cdot \left(\mathbf{x}_{\alpha}(y) - \mathbf{x}_{\alpha}(x) \right) \right| d\mu(y) = 0;$$

https://doi.org/10.4153/CJM-2012-064-x Published online by Cambridge University Press

in other words, f is L^1-differentiable. For uniformly perfect spaces equipped with a doubling measure, L^1-differentiability implies approximate differentiability. For a proof of this fact, see [24, Prop. 3.4]. Notice that a space supporting a Poincaré inequality is connected and thus also uniformly perfect.

In [37] it is proved that BV functions are L^1-differentiable μ-a.e., if the space supports a 1-Poincaré inequality. As a direct consequence we deduce as well that BV functions are approximately differentiable μ-a.e. Observe that the approximate differentiability is a weaker notion than L^p-differentiability. In particular, the definition of the approximate differentiability does not involve any integrability assumptions.

5 Approximate Differentiability of the Maximal Function

Hajłasz and Malý proved in [19] that, in the case of $X = \mathbb{R}^n$, approximate differentiability is preserved under the action of the *Hardy–Littlewood maximal operator*

$$Mf(x) := \sup_{r>0} \fint_{B(x,r)} |f(y)|\, dy, \quad x \in X.$$

It was recently shown by H. Luiro [32, Corollary 1.5] that, in the Euclidean case, differentiability almost everywhere is also preserved under the action of the maximal function.

On the other hand, in the setting of metric spaces endowed with a doubling measure, the maximal operator does not preserve the regularity of a function in the same manner as in the Euclidean case. Kinnunen proved in [25, Theorem 1.4] that the Hardy–Littlewood maximal operator is bounded in $W^{1,p}(\mathbb{R}^n)$ for $1 < p \le \infty$. Notice that the case $p = \infty$ corresponds to the space of Lipschitz functions. On the other hand, Buckley [11, Example 1.4] has shown that for a metric space with a doubling measure, the maximal operator may not preserve Lipschitz and Hölder spaces. In order to have a maximal function that preserves, for example, the Sobolev spaces on metric spaces, Kinnunen and Latvala [26] constructed a maximal function based on discrete convolution (see also [1, 2]).

In the next theorem, we will show that the discrete maximal operator also preserves approximate differentiability. First, we define the discrete maximal operator at scale $r > 0$. Let $B_i = B(x_i, r)$, $i \in \mathbb{N}$, be a collection of balls such that they cover X and the balls $B(x_i, r/2)$, $i \in \mathbb{N}$, are pairwise disjoint. Let ψ_i be a partition of unity subordinate to the covering B_i, $i = 1, 2, \dots$, *i.e.*, $0 \le \psi_i \le 1$, $\operatorname{supp} \psi_i \subset B(x_i, 6r)$, $\psi_i \ge 1/C$ in $B(x_i, 3r)$, ψ_i is Lipschitz with constant L/r and $\sum_i \psi_i = 1$. Then we define the discrete convolution of $f \in L^1_{\mathrm{loc}}(X)$ by setting

$$f_r(x) = \sum_{i=1}^{\infty} \psi_i(x) f_{B(x_i,3r)}. \tag{5.1}$$

Let r_j be an enumeration of the positive rationals. We define the discrete maximal function (which depends on the chosen covering)

$$M^* f(x) = \sup_j |f|_{r_j}(x).$$

Now we can state our theorem. Notice that the theorem holds for arbitrarily chosen coverings defining $M^* f$.

Theorem 5.1 *Let (X, d, μ) be a complete metric space equipped with a doubling measure μ. Assume also that X supports an approximate differentiable structure. If $f \in L^1(X)$ is approximately differentiable μ-a.e., then $M^* f$ is approximately differentiable μ-a.e.*

The proof follows the ideas used in [19]. First, we consider the restricted maximal function $M^*_\varepsilon f$, $\varepsilon > 0$, defined by the formula

$$M^*_\varepsilon f(x) := \sup_{r_j > \varepsilon} |f|_{r_j}(x).$$

Lemma 5.2 *Let (X, d) be a complete metric measure space equipped with a doubling measure μ. Assume also that X supports an approximate differentiable structure. If $f \in L^1(X)$, then $M^*_\varepsilon f$, $\varepsilon > 0$, is approximately differentiable μ-a.e. in X.*

Proof We start by proving that for some constant $\widetilde{Q}$, which depends only on the doubling constant, the following inequality holds:
(5.2)

$$\left| M^*_\varepsilon f(x) - M^*_\varepsilon f(y) \right| \le \frac{\widetilde{Q}}{\varepsilon} d(x, y)\big(M^*_\varepsilon f(x) + M^*_\varepsilon f(y)\big) \qquad \text{for } \mu\text{-a.e. } x, y \in X.$$

Notice first, that the claim clearly holds if $d(x, y) \ge \varepsilon$, so we may assume that $d(x, y) < \varepsilon$. Fix a rational number $r > \varepsilon$. Let $\{B(x_i, r)\}_i$ be the covering used to define the discrete convolution $|f|_r$ and let I denote the set of indexes i such that x or y belong to $B(x_i, 6r)$. The doubling property implies that $|I| \le C$ with a constant that only depends on the doubling constant. For every $i \in I$, there exists a point $\widetilde{x} \in B(x_i, 4r)$ such that the ball $B(\widetilde{x}, 4r)$ belongs to the covering used to construct $|f|_{4r}$ as in (5.1). Since $B(x_i, 3r) \subset B(\widetilde{x}, 12r)$ and $x \in B(\widetilde{x}, 12r)$, we have

$$|f|_{B(x_i,3r)} \le C|f|_{B(\widetilde{x},3\cdot 4r)} \le C|f|_{4r}(x) \le CM^*_\varepsilon f(x)$$

with a constant depending only on the doubling constant; see, for example, the proof of [26, Lemma 3.1]. Thus we can conclude that

$$\begin{aligned}\big| |f|_r(x) - |f|_r(y) \big| &= \Big| \sum_{i \in I} \big(\psi_i(x) - \psi_i(y)\big) |f|_{B(x_i,3r)} \Big| \\ &\le C\frac{L}{r} d(x, y) M^*_\varepsilon f(x) \le C\frac{L}{\varepsilon} d(x, y)\big(M^*_\varepsilon f(x) + M^*_\varepsilon f(y)\big).\end{aligned}$$

By taking the supremum over all rationals $r > \varepsilon$, we obtain (5.2).

Now it is enough to notice that the restricted maximal function is μ-measurable, hence by Luzin's Theorem 2.4, it is approximately continuous μ-a.e., and by (5.2)

$$\operatorname{ap}\limsup_{y \to x} \frac{|M^*_\varepsilon f(x) - M^*_\varepsilon f(y)|}{d(x, y)} \le \frac{\widetilde{Q}}{\varepsilon} 2M^*_\varepsilon f(x) < \infty \quad \mu\text{-a.e. in } X.$$

Using a Stepanov-type characterization (see Corollary 3.4), we obtain that $M^*_\varepsilon f$ is approximately differentiable μ-a.e. in X. ■

It would be interesting to know, whether Theorem 5.1 holds for the standard Hardy–Littlewood maximal function as well. In the metric space setting, this would require a totally different proof, since estimates like (5.2) do not hold in spaces where the measure of balls does not behave nicely. Even with the annular decay property, only Hölder type estimates are available [11].

Now we are ready to prove Theorem 5.1.

Proof of Theorem 5.1 First, we can split the space into three parts

$$X = \left\{ x : M^* f(x) > |f(x)| \right\} \cup \left\{ x : M^* f(x) = |f(x)| \right\} \cup N.$$

By [1, Lemma 4.5], $|f(x)| \leq M^* f(x)$ a.e. in X. Thus $\mu(N) = 0$.

Observe that if $f \in L^1(X)$ is approximately differentiable function at μ-almost every point in X, then $|f|$ is approximately differentiable μ-a.e. in X as well. This fact easily follows, for example, from Theorem 3.1 on Whitney-type characterization of approximate differentiability.

Thus, the maximal function $M^* f$ is approximate differentiable μ-a.e. on the second set. Note also that since μ-almost every point of X is a Lebesgue point of f (see *e.g.*, [21, Theorem 1.8]), it is enough to show that $M^* f$ is approximately differentiable almost everywhere on the set

$$A := \{x : M^* f(x) > |f(x)| \text{ and } x \text{ is a Lebesgue point of } f\}.$$

If $x \in A$, there exists a sequence $\{r_n\}_{n=1}^{\infty}$ such that

$$\lim_{n\to\infty} |f|_{r_n}(x) = M^* f(x).$$

The sequence r_n is bounded (since $M^* f(x) > 0$ and $f \in L^1(X)$), and we can find a convergent subsequence. Let us denote its limit by r. Note that $r > 0$, otherwise $M^* f(x) = |f|(x)$. Thus for each $x \in A$ there exists $k \in \mathbb{N}$ such that $M^* f(x) = M^*_{1/k} f(x)$ and

$$A \subset \bigcup_{k=1}^{\infty} \left\{ x : M^* f(x) = M^*_{1/k} f(x) \right\}.$$

By Lemma 5.2, each maximal function $M^*_{1/k} f(x)$, $k \in \mathbb{N}$, is approximately differentiable μ-a.e. in X, and, since the sets $\{x : M^* f(x) = M^*_{1/k} f(x)\}$ are measurable, $M^* f$ is approximately differentiable μ-a.e. in A. ∎

Acknowledgments Part of this research was conducted while the first and fourth authors visited Aalto University; they wish to thank the institution for its kind hospitality. The authors would also like to thank Prof. Juha Kinnunen for valuable comments and suggestions, which have been helpful in improving the manuscript. The authors would like to thank the referee for a very careful reading of the manuscript and many useful comments.

References

[1] D. Aalto and J. Kinnunen, *The discrete maximal operator in metric spaces.* J. Anal. Math. **111**(2010), 369–390. http://dx.doi.org/10.1007/s11854-010-0022-3

[2] ———, *Maximal functions in Sobolev spaces.* In: Sobolev spaces in mathematics. I, Int. Math. Ser. (N. Y.), 8, Springer, New York, 2009, pp. 25–67.

[3] L. Ambrosio and P. Tilli, *Topics on analysis in metric spaces.* Oxford Lecture Series in Mathematics and its Applications, 25, Oxford University Press, Oxford, 2004.

[4] Z. M. Balogh, K. Rogovin, and T. Zürcher, *The Stepanov differentiability theorem in metric measure spaces.* J. Geom. Anal. **14**(2004), no. 3, 405–422. http://dx.doi.org/10.1007/BF02922098

[5] S. G. Basalaev and S. K. Vodopyanov, *Approximate differentiability of mappings of Carnot-Carathéodory spaces.* 2012. arxiv:1206.5197

[6] D. Bate and G. Speight, *Differentiability, porosity and doubling in metric measure spaces.* Proc. Amer. Math. Soc. **141**(2013), no. 3, 971–985. http://dx.doi.org/10.1090/S0002-9939-2012-11457-1

[7] A. Björn and J. Björn, *Nonlinear potential theory on metric spaces.* EMS Tracts in Mathematics, 17, European Mathematical Society, Zürich, 2011.

[8] A. Björn, J. Björn, and N. Shanmugalingam, *The Dirichlet problem for p-harmonic functions on metric spaces.* J. Reine Angew. Math. **556**(2003), 173–203.

[9] J. Björn, *L^q-differentials for weighted Sobolev spaces.* Michigan Math. J. **47**(2000), no. 1, 151–161. http://dx.doi.org/10.1307/mmj/1030374674

[10] B. Bojarski, *Differentiation of measurable functions and Whitney-Luzin type structure theorems.* Helsinki University of Technology Institute of Mathematics Research Reports, 2009.

[11] S. M. Buckley, *Is the maximal function of a Lipschitz function continuous?* Ann. Acad. Sci. Fenn. Math. **24**(1999), no. 2, 519–528.

[12] J. Cheeger, *Differentiability of Lipschitz Functions on metric measure spaces.* Geom. Funct. Anal. **9**(1999), no. 3, 428–517. http://dx.doi.org/10.1007/s000390050094

[13] A. Denjoy, *Sur les fonctions dérivées sommables.* Bull. Soc. Math. France **43**(1915), 161–248.

[14] E. Durand-Cartagena, J. A. Jaramillo, and N. Shanmugalingam, *The ∞-Poincaré inequality in metric measure spaces.* Michigan Math. J. **61**(2012), no. 1, 63–85. http://dx.doi.org/10.1307/mmj/1331222847

[15] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions.* Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992.

[16] H. Federer, *Geometric measure theory.* Die Grundlehren der mathematischen Wissenschaften, 153, Springer-Verlag New York Inc., New York, 1969.

[17] P. Hajłasz, *Sobolev spaces on metric-measure spaces.* In: Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), Contemp. Math., 338, Amer. Math. Soc., Providence, RI, 2003, pp. 173–218.

[18] P. Hajłasz and P. Koskela, *Sobolev met Poincaré.* Mem. Amer. Math. Soc. **145**(2000), no. 688.

[19] P. Hajłasz and J. Malý, *On approximate differentiability of the maximal function.* Proc. Amer. Math. Soc. **138**(2010), no. 1, 165–174. http://dx.doi.org/10.1090/S0002-9939-09-09971-7

[20] J. Heinonen and P. Koskela, *Quasiconformal maps in metric spaces with controlled geometry.* Acta Math. **181**(1998), no. 1, 1–61. http://dx.doi.org/10.1007/BF02392747

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

[22] ———, *Nonsmooth calculus.* Bull. Amer. Math. Soc. (N.S.) **44**(2007), no. 2, 163–232. http://dx.doi.org/10.1090/S0273-0979-07-01140-8

[23] S. Keith, *A differentiable structure for metric measure spaces.* Adv. Math. **183**(2004), no. 2, 271–315. http://dx.doi.org/10.1016/S0001-8708(03)00089-6

[24] ———, *Measurable differentiable structures and the Poincaré inequality.* Indiana Univ. Math. J. **53**(2004), no. 4, 1127–1150. http://dx.doi.org/10.1512/iumj.2004.53.2417

[25] J. Kinnunen, *The Hardy-Littlewood operator of a Sobolev function.* Israel J. Math. **100**(1997), 117–124. http://dx.doi.org/10.1007/BF02773636

[26] J. Kinnunen and V. Latvala, *Lebesgue points for Sobolev functions on metric spaces.* Rev. Mat. Iberoamericana **18**(2002), no. 3, 685–700. http://dx.doi.org/10.4171/RMI/332

[27] B. Kleiner and J. Mackay, *Differentiable structures on metric measure spaces: A Primer.* 2011. arxiv:1108.1324.

[28] P. Lahti and H. Tuominen, *A pointwise characterization of functions of bounded variation on metric spaces.* 2012. arxiv:1301.6897

[29] F. C. Liu, *Luzin type property of Sobolev functions.* Indiana Univ. Math. J. **26**(1977), no. 4, 645–651. http://dx.doi.org/10.1512/iumj.1977.26.26051

[30] F. C. Liu and W. S. Tai, *Approximate Taylor polynomials and differentiation of functions.* Topol. Methods Nonlinear Anal. **3**(1994), no. 1, 189–196.

[31] ———, *Lusin properties and interpolation of Sobolev spaces.* Topol. Methods Nonlinear Anal. **9**(1997), no. 1, 163–177.
[32] H. Luiro, *On the size of the set of non-differentiability points of maximal function.* 2012. arxiv:1208.3971
[33] N. Luzin, *Sur les propriétés des fonctions mesurables.* Comptes Rendus Acad. Sci. Paris **154**(1912), 1688–1690.
[34] J. Malý, *A simple proof of the Stepanov theorem on differentiability almost everywhere.* Exposition. Math. **17**(1999), no. 1, 59–61.
[35] J. Malý and L. Zajíček, *Approximate differentiation: Jarník points.* Fund. Math. **140**(1991), no. 1, 87–97.
[36] M. Miranda Jr., *Functions of bounded variation on "good" metric spaces.* J. Math. Pures Appl. (9) **82**(2003), no. 8, 975–1004.
[37] A. Ranjbar-Motlagh, *Generalized Stepanov type theorem with applications over metric-measure spaces.* Houston J. Math. **34**(2008), no. 2, 623–635.
[38] S. Semmes, *Some novel types of fractal geometry.* Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2001.
[39] N. Shanmugalingam, *Newtonian spaces: an extension of Sobolev spaces to metric spaces.* Rev. Mat. Iberoamericana **16**(2000), no. 2, 243–279. http://dx.doi.org/10.4171/RMI/275
[40] W. Stepanoff, *Sur les conditions de l'existence de la différentielle totale.* Rec. Math. Soc. Math. Moscou **32**(1925), no. 3, 511–526.
[41] H. Whitney, *On totally differentiable and smooth functions.* Pacific J. Math. **1**(1951), 143–159. http://dx.doi.org/10.2140/pjm.1951.1.143

Departamento de Matemática Aplicada. ETSI Industriales, UNED c/Juan del Rosal 12 Ciudad Universitaria, 28040 Madrid, Spain
e-mail: edurand@ind.uned.es

Department of Mathematics and Statistics, P.O. Box 68, FI-00014 University of Helsinki, Finland
e-mail: lizaveta.ihnatsyeva@helsinki.fi riikka.korte@helsinki.fi

Faculty of Mathematics, Informatics, and Mechanics University of Warsaw, Banacha 2, 02-097 Warszawa, Poland
e-mail: m.szumanska@mimuw.edu.pl