Real Analysis and geometrical structures.

Mini-Course 2017.

... il fatto è che un gran numero di incantatori è sempre tra noi,
e tramuta le cose conferendo loro un aspetto ingannevole,
dirigendole come conviene alla loro fantasia,
secondo che vogliono distruggerci o favorirci.

Don Chisciotte, Miguel de Cervantes Saavedra

Chapter 1

Lipschitz Analysis

Let (X, d_X) and (Y, d_Y) be two metric spaces and $L \geq 0$. A function $f : X \longrightarrow Y$ is said to be L-Lipschitz if

$$d_Y(f(x), f(y)) \leq L \cdot d_X(x, y) \tag{1.1}$$

for every pair of points $x, y \in X$.

For a function $f : X \to Y$ the Lipschitz constant is defined by

$$\|f\|_{Lip} = \sup\{\frac{d_Y(f(x) - f(y))}{d_X(x, y)} : \ x, y \in X, \ x \neq y\}. \tag{1.2}$$

Indeed, f is Lipschitz if $\|f\|_{Lip} < \infty$.

Let us notice that Condition (1.1) is purely metric and appears nearly everywhere in mathematics. Throughout these notes will shall use the following notation: n-dimensional Euclidean space stands for ℓ_2^n; i.e., the space $\mathbb{R}^n$ with the usual distance

$$\|x - y\|_2 = \left[\sum_{i=1}^{n} |x_i - y_i|^2\right]^{\frac{1}{2}}.$$

For a general metric space, $B(x, r)$ will stand for open ball centered at $x \in X$ with radius $r > 0$ while $\overline{B}(x, r)$ for closed balls. Usually, we use B_X instead of $B(\theta_X, 1)$, as unit ball, in case we distinguish a point θ_X as origin, and by S_X we consider the related unite sphere associate with. In case $X = \mathbb{R}^n$ we simply use S^{n-1}.

1.1 Extension

For every $X \subseteq Y$ we denote by $e(X, Y, Z)$ the infimuum over all constants K such that every Lipschitz function $f : X \to Z$ can be extended to a function

$\widetilde{f} : Y \to Z$ satisfying $\|\widetilde{f}\|_{Lip} \leq K\|f\|_{Lip}$. We also define

$$e(Y, Z) = \sup\{e(X, Y, Z) : \ X \subseteq Y\}$$

and

$$e(X) = \sup\{e(X, Y, Z) : \ X \subseteq Y, \ Z \text{ a Banach space}\}.$$

In this section, we would like to review the question whether or not, given a L-Lipschitz $f : Z \longrightarrow Y$, $Z \subseteq X$, it can be extended by a Lipschitz function on the whole space X.

Remark 1.1. In [9], Lindenstrauss provided an example in which Banach-space valued Lipschitz functions do not admit extension.

1.1.1 Euclidian case

In case $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$ this can be done always. Let us first start to recall an easy exercise.

Lemma 1.2. *Let $\{f_i : i \in I\}$ be a collection of L-Lipschitz functions, $f_i : A \longrightarrow \mathbb{R}$, $A \subseteq \mathbb{R}^n$. Then the functions*

$$x \longmapsto \inf_{i \in I} f_i(x), \ x \in A,$$

and

$$x \longmapsto \sup_{i \in I} f_i(x), \ x \in A,$$

are L-Lipschitz on A (if finite at one point).

Note that a L-Lipschitz function $f : A \longrightarrow \mathbb{R}^m$ can be trivially extended on $\overline{A}$, simply by uniformly continuity.

Theorem 1.3 (McShane-Whitney extension theorem)**.** *Let $f : A \longrightarrow \mathbb{R}$, $A \subseteq \mathbb{R}^n$, be a L-Lipschitz function. Then there exist a largest and smallest L-Lipschitz functions*

$$f_{min} : \mathbb{R}^n \longrightarrow \mathbb{R}$$

and

$$f_{max} : \mathbb{R}^n \longrightarrow \mathbb{R},$$

respectively, which extend f; i.e. for any L-Lipschitz function $G : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $G|_A = f$ then

$$f_{min} \leq G \leq f_{max}.$$

Proof. It is enough to use the previous lemma and define

$$f_{\min}(x) = \sup_{a\in A}\{f(a) + L\|x-a\|_2\}$$

and

$$f_{\max}(x) = \inf_{a\in A}\{f(a) - L\|x-a\|_2\}$$

$\square$

If we apply McShane-Whitney extension theorem on each coordinate, we obtain

Corollary 1.4. *Let $f : A \longrightarrow \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$, be a L-Lipschitz function. Then there exists $\sqrt{m}L$-Lipschiz function $\widetilde{f} : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $\widetilde{f}|_A = f$.*

A natural question is whenever we can find a better Lipschitz extension, with constant independent of the dimension of the target arrive space. The answer came out in 1934 by M.D. Kirszbraun [7].

Theorem 1.5. *Let $f : A \longrightarrow \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$, be a L-Lipschitz function. Then there exists L-Lipschiz function $\widetilde{f} : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $\widetilde{f}|_A = f$.*

Lemma 1.6. *Let H be a Hilbert space, $C \subseteq H$ be a non empty closed convex set and $b \in H$. Then there is $b_1 \in C$ such that*

$$\langle z-b, b_1-b\rangle \geq \|b_1-b\|^2, \quad \textit{for every } z \in C.$$

Proof. Set $\alpha = \inf\{\|z-b\| : z \in C\}$. For each $n \in \mathbb{N}$, let $C_n = \{z \in C : \|z-b\| \leq \alpha + 4^{-n}\}$. It easy to show that

$$\|z_1 - z_2\| \leq 2^{-n}\sqrt{8\alpha+4}, \text{ for all } z_1, z_2 \in C_n.$$

Indeed, since $\frac{1}{2}(z_1+z_2) \in C$ then $\|\frac{1}{2}(z_1+z_2) - b\| \geq \alpha$. Now

$$\begin{aligned} 4\alpha^2 &\leq \|2(\frac{1}{2}(z_1+z_2)) - 2b\|^2 \\ &= \|(z_1-b)+(z_2-b)\|^2 \\ \text{(parallelogram law)} \quad &= 2\|z_1-b\|^2 + 2\|z_2-b\|^2 - \|(z_1-b)-(z_2-b)\|^2 \\ &\leq 4(\alpha+4^{-n})^2 - \|z_1-z_2\|^2. \end{aligned}$$

Therefore,

$$\|z_1-z_2\|^2 \leq 4((\alpha+4^{-n})^2 - \alpha^2) \leq 4^{-n}(8\alpha+4).$$

For each $n \in \mathbb{N}$, let $z_n \in C_n$. For what we said before, $\|z_m - z_n\| \leq 2^{-m}\sqrt{8\alpha+4}$, $n \geq m$. Therefore the sequence $(z_n)_n$ is Cauchy and then admits limit point $b_1 \in C$ (C is closed!). Of course,

$$\alpha \leq \|b_1 - b\| = \lim_n \|z_n - b\| \leq \alpha,$$

thus $\|b_1 - b\| = \alpha$.

Now, let $z \in C$, since C is convex then $\lambda z + (1-\lambda)b_1 \in C$ for all $\lambda \in]0,1[$. But then,

$$\begin{aligned}\|b_1 - b\|^2 &\leq \|(\lambda z + (1-\lambda)b_1) - b\|^2 \\ &= \lambda^2\|z-b\|^2 + 2\lambda(1-\lambda)\langle z-b, b_1-b\rangle + (1-\lambda)^2\|b_1-b\|^2;\end{aligned}$$

subtracting $\|b_1 - b\|^2$ from both sides,

$$0 \leq \lambda^2\|z-b\|^2 + 2\lambda(1-\lambda)\langle z-b, b_1-b\rangle - 2\lambda\|b_1-b\|^2 + \lambda^2\|b_1-b\|^2;$$

dividing by $2\lambda > 0$, we get

$$0 \leq \frac{1}{2}\lambda\|z-b\|^2 + (1-\lambda)\langle z-b, b_1-b\rangle - \|b_1-b\|^2 + \frac{1}{2}\lambda\|b_1-b\|^2;$$

letting $\lambda \to 0$ we get the thesis. □

Lemma 1.7. *Let H_1, H_2 two Hilbert spaces and $J \subseteq H_1$ be a non empty finite set. Let $g : J \longrightarrow H_2$ be a 1-Lipschitz function such that $\|g(x)\| > \|x\|$ for all $x \in J$. Then $0 \notin co(g(J))$.*

Proof. Note first that

$$\begin{aligned}\langle x, y\rangle &= \frac{1}{2}(\|x\|^2 + \|y\|^2 - \|x-y\|^2) \\ &< \frac{1}{2}(\|g(x)\|^2 + \|g(y)\|^2 - \|g(x)-g(y)\|^2) \\ &= \langle g(x), g(y)\rangle\end{aligned}$$

Now, if $w = \sum_{x\in J}\lambda_x g(x)$, with $\lambda_x \in [0,1]$ and $\sum_{x\in J}\lambda_x = 1$,

$$\begin{aligned}\|w\|^2 &= \langle w, w\rangle \\ &= \langle \sum_{x\in J}\lambda_x g(x), \sum_{y\in J}\lambda_y g(y)\rangle \\ &= \sum_{x,y\in J}\lambda_x\lambda_y\langle g(x), g(y)\rangle\end{aligned}$$

$$
\begin{aligned}
&> \sum_{x,y\in J} \lambda_x \lambda_y \langle x, y\rangle \\
&= \langle \sum_{x\in J} \lambda_x x, \sum_{y\in J} \lambda_y y\rangle \\
&= \| \sum_{x\in J} \lambda_x x\|^2 \\
&\geq 0.
\end{aligned}
$$

Therefore, w cannot be the origin. □

Lemma 1.8. *Let H_1, H_2 be two Hilbert spaces, $I \subseteq H_1$ finite, $f : I \longrightarrow H_2$ a 1-Lipschitz function and $a \in H_1$. Then there exists $b \in H_2$ such that $\|b - f(x)\| \leq \|a - x\|$ for every $x \in I$.*

Proof. Trivial cases: $I = \emptyset$ choose $b = 0$ and $a \in I$ choose $b = f(a)$. In the other cases, let $C = \mathrm{co}(f(I)) \subseteq H_2$ which is nonempty convex set. Let us define $G : C \longrightarrow [0, +\infty[$ as

$$G(z) = \max_{x\in I} \frac{\|z - f(x)\|}{\|a - x\|}.$$

Since I is finite, G is continuous on C (see C as bounded convex set of a finite dimensional space, then compactness holds) then G attains its minimum at a point $b \in C$. Set

$$\gamma = G(b) \text{ and } J = \{x \in I : \ \frac{\|b - f(x)\|}{\|a - x\|} = \gamma\}.$$

Of course J is non empty. First, let us prove that $b \in \mathrm{co}(f(J))$.

Suppose that $b \notin \mathrm{co}(f(J))$, by Lemma 1.6 there is $b_1 \in \mathrm{co}(f(J))$ such that

$$\langle z - b, b_1 - b\rangle \geq \|b_1 - b\|^2, \quad \text{for every } z \in \mathrm{co}(f(J)).$$

In particular, $\langle f(x) - b, b_1 - b\rangle \geq \|b_1 - b\|^2$, for every $x \in J$. For a small $\delta > 0$, let us consider $b_\delta = (1 - \delta)b + \delta b_1 \in C$. If $x \in J$, then

$$\langle f(x) - b, b_\delta - b\rangle = \delta\langle f(x) - b, b_1 - b\rangle \geq \delta\|b_1 - b\|^2$$

thus

$$
\begin{aligned}
\|f(x) - b_\delta\|^2 &= \|(f(x) - b) - (b_\delta - b)\|^2 \\
&= \|f(x) - b\|^2 - 2\langle f(x) - b, b_\delta - b\rangle + \|b_\delta - b\|^2 \\
&\leq \|f(x) - b\|^2 - 2\delta\|b_1 - b\|^2 + \delta^2\|b_1 - b\|^2
\end{aligned}
$$

$$(\text{since } 0 < \delta \le 1) \ < \|f(x) - b\|^2$$

On the other hand, if $x \in I \setminus J$,

$$\lim_{\delta \to 0} \frac{\|f(x) - b_\delta\|}{\|x - a\|} = \frac{\|f(x) - b\|}{\|x - a\|} < \gamma,$$

so there is a $\delta_x > 0$ such that $\frac{\|f(x)-b_\delta\|}{\|x-a\|} < \gamma$ whenever $0 < \delta \le \delta_x$. Since $I \setminus J$ is finite, it is evidente that we can choose $\delta \le \delta_x$ for all $x \in I \setminus J$. For this $\delta > 0$ we have

$$\frac{\|f(x) - b_\delta\|}{\|x - a\|} < \gamma, \quad \forall x \in I.$$

This would implies $G(b_\delta) < \gamma = G(b)$, against the fact that b was minimum for G. Thus $b \in \mathrm{co}(f(J))$ and then it can be written as $b = \sum_{x \in J} \lambda_x f(x)$ with $\lambda_x \in [0,1]$ and $\sum_{x \in J} \lambda_x = 1$.

Set $J_1 = \{x - a : \ x \in J\}$ and $h : J_1 \longrightarrow H_2$ be $h(x) = f(x + a) - b$. One has,

$$\begin{aligned}\|h(x) - h(y)\| &= \|f(x + a) - f(y + a)\| \\ &\le \|(x + a) - (y + a)\| \\ &= \|x - y\|,\end{aligned}$$

and (since $x + a \in J$)

$$\|h(x)\| = \|f(x + a) - b\| = \gamma\|x\|.$$

Therefore,

$$\begin{aligned}\sum_{x \in J_1} \lambda_{x+a} h(x) &= \sum_{x \in J_1} \lambda_{x+a}(f(x + a) - b) \\ &= \sum_{x \in J} \lambda_x(f(x) - b) \\ &= \sum_{x \in J} \lambda_x f(x) - \sum_{x \in J} \lambda_x b \\ &= b - b \\ &= 0.\end{aligned}$$

Of course, $\lambda_{x+a} \in [0,1]$ with $\sum_{x \in J_1} \lambda_{x+a} = 1$. Thus, $0 \in \mathrm{co}(h(J_1))$. By Lemma 1.7 there must exists $x \in J_1$ such that $\|x\| \ge \|h(x)\| = \gamma\|x\|$. Since $a \notin I$ we have that $x = x' - a \ne 0$. Thus $\gamma \le 1$. Finally

$$G(b) \le 1 \Rightarrow \|f(x) - b\| \le \|x - a\|, \quad \forall x \in I.$$

$\square$

Lemma 1.9. *Let $\{x_1, \ldots, x_k\}$ be a finite collection of points in $\mathbb{R}^n$, and let $\{y_1, \ldots, y_k\}$ be a finite collection of points in $\mathbb{R}^m$ such that*

$$\|y_i - y_j\|_2 \leq \|x_i - x_j\|_2, \quad \forall i, j \in \{1, \ldots, k\}.$$

If $r_1, \ldots, r_k$ are positive numbers such that

$$\bigcap_{i=1}^{k} \overline{B}(x_i, r_i) \neq \emptyset,$$

then

$$\bigcap_{i=1}^{k} \overline{B}(y_i, r_i) \neq \emptyset, .$$

Proof. It is enough to define $f(x_i) = y_i$, choose $a \in \bigcap_{i=1}^{k} \overline{B}(x_i, r_i)$ and apply the previous Lemma. □

Remark 1.10. M. Gromov in [6] has proved a stronger version of the previous Lemma. Under the same assumption, actually the following holds

$$\lambda\left(\bigcap_{i=1}^{k} \overline{B}(x_i, r_i)\right) \leq \lambda\left(\bigcap_{i=1}^{k} \overline{B}(y_i, r_i)\right),$$

where λ denotes the Lebesgue measure on the Euclidean spaces.

Proposition 1.11. *Let $f : F \longrightarrow \mathbb{R}^m$ a 1-Lipschitz function, with $F \subseteq \mathbb{R}^n$ a finite set. If $x \in \mathbb{R}^n$, then there exists an extension of f to $F \cup \{x\}$ as 1-Lipschitz function.*

Proof. Suppose $F = \{x_1, \ldots, x_k\}$. Set $r_i = \|x - x_i\|_2$ and $y_i = f(x_i)$. By the previous Lemma, there is a point $y \in \mathbb{R}^m$ such that

$$\|y - f(x_i)\|_2 \leq \|x - x_i\|_2, \quad \text{for all } i \in \{1, \ldots, k\}.$$

The desired extension is accomplished by setting $f(x) = y$. □

We are ready to prove the main theorem.

Proof. (of Theorem 1.5)

By dividing the function f by L, we may assume that $f : A \longrightarrow \mathbb{R}^m$ is 1-Lipschitz. We use a standard Ascoli-Arzelá type argument. Choose a countable dense $\{a_n,\ n \in \mathbb{N}\}$ of A and $\{b_n,\ n \in \mathbb{N}\}$ dense of $\mathbb{R}^n \setminus A$. By Proposition 1.11, for each $k \in \mathbb{N}$ there exists a 1-Lipschitz function

$$f_k : \{a_1, \ldots, a_k\} \cup \{b_1, \ldots, b_k\} \longrightarrow \mathbb{R}^m,$$

such that $f_k(a_i) = f(a_i)$ for every $i = 1, \dots, k$. Since the sequence $(f_k(b_1))_k$ is bounded in $\mathbb{R}^m$, it has a convergent subsequence, say $(f_{k_j^1}(b_1))_j$. Similarly, from the mappings corresponding to this subsequence, we can subtract another subsequence, say $(f_{k_j^2})_j$ such that $(f_{k_j^2}(b_2))_j$ converges. Continuing this way and passing to the diagonal sequence $g_j = f_{k_j^j}$ we find that the limit

$$\widetilde{f}(c) = \lim_{j\to\infty} g_j(c) \in \mathbb{R}^m$$

exists for every $c \in \{a_1, a_2, \dots\} \cup \{b_1, b_2, \dots\}$. Moreover,

$$\widetilde{f} : \{a_1, a_2, \dots\} \cup \{b_1, b_2, \dots\} \longrightarrow \mathbb{R}^m$$

is 1-Lipschitz and $\widetilde{f}(a_i) = f(a_i)$ for all $i \in \mathbb{N}$. Since $\{a_1, a_2, \dots\} \cup \{b_1, b_2, \dots\}$ is dense in $\mathbb{R}^n$ we can extend easily $\widetilde{f}$ on the whole space $\mathbb{R}^n$. $\square$

1.1.2 Doubling metric space case

Recall that the *doubling constant* of a metric space (X, d), denoted by $\lambda(X)$, is the infimuum over all natural numbers λ such that every ball in X can be covered by λ balls of half the radius. When $\lambda(X) < \infty$ one says that (X, d) is *doubling*. A measure μ on a metric space (X, d) is said *doubling*, in such case one says that the *metric measure space* (X, d, μ) *is doubling*, if there exists a constant $C \geq 1$ such that whenever we pick a ball $B(x, r)$ in X

$$\mu(B(x, 2r)) \leq C\mu(B(x, r)).$$

Let us note that every Euclidean space is doubling. In [3], it is proved that if (X, d) is a doubling metric, then for any $0 < \alpha < 1$, (X, d^α) embeds into ℓ_2^k with distortion D, where k and D depend only on the doubling constant of X. That means, there exists a injective Lipschitz function,

$$f : (X, d^\alpha) \longrightarrow \ell_2^k,$$

with inverse also Lipschitz $f^{-1} : f(X) \longrightarrow X$, such that $D = \|f\|_{Lip} \cdot \|f^{-1}\|_{Lip} < \infty$. The infimum of such constants D is usually called *Lipschitz distortion.*

Here, (X, d^α) is the metric with all distances raised to the power α (this is called a *snowflaked* version of X). Unfortunately, the dependence of k and D on the doubling contant is exponential. Assouad also conjectured that the above result holds even when $\alpha = 1$. This question was solved by Semmes [12] which disproved this conjecture.

It is still unknown the following

Question 1.12. Does every doubling subset of ℓ_2 admit a bi-Lipschitz embedding into some Euclidean space?

Throughout this section, we would like to extend in a much stronger version the theorem we have seen in the last subsection.

Before to go on the main result of this section, we would like to treat a particular case where in spirit will help to understand the idea behind.

A metric space (X, d) is called *uniformly discrete* if $d(x, y) \geq \varepsilon > 0$, for all $x \neq y$, $x, y \in X$. The following observation is due by W.B. Johnson, J. Lindenstrauss and G. Schechtman (1986).

Proposition 1.13. *Let (X, d) be a metric space, $Y \subseteq X$ be a uniformly ε-discrete subspace with finite diameter D and Z be a Banach space. Then every Lipschitz function $f : Y \longrightarrow Z$ admits a Lipschitz extension $\widetilde{f} : Y \longrightarrow Z$ such that*

$$\|\widetilde{f}\|_{Lip} \leq \frac{2D}{\varepsilon}\|f\|_{Lip}.$$

Proof. Fix any $x_0 \in Y$ and define

$$\widetilde{f}(x) = \begin{cases} \frac{2}{\varepsilon}\left[d(x,t)f(x_0) + (\frac{\varepsilon}{2} - d(x,t))f(t)\right], & \text{if } x \in B(t, \frac{\varepsilon}{2}), \text{ some } t \in Y \\ \\ f(x_0), \quad \text{if } x \notin \bigcup_{t \in Y} B(t, \frac{\varepsilon}{2}) \end{cases}$$

Of course $\widetilde{f}$ extends f. Let us suppose $x \in B(t_x, \frac{\varepsilon}{2})$ and $y \in B(t_y, \frac{\varepsilon}{2})$, some $t_x, t_y \in Y$. If $t_x \neq t_y$ then $(\frac{\varepsilon}{2} - d(x, t_x)) + (\frac{\varepsilon}{2} - d(y, t_y)) \leq d(x, y)$. Then

$$\begin{aligned}
&\|\widetilde{f}(x) - \widetilde{f}(y)\| \\
&= \frac{2}{\varepsilon}\|(d(x,t_x)f(x_0) + (\frac{\varepsilon}{2} - d(x,t_x))f(t_x)) - (d(y,t_y)f(x_0) + (\frac{\varepsilon}{2} - d(y,t_y))f(t_y))\| \\
&\leq \frac{2}{\varepsilon}\|(\frac{\varepsilon}{2} - d(x,t_x))(f(t_x) - f(x_0)) + \frac{\varepsilon}{2}f(x_0) - (\frac{\varepsilon}{2} - d(y,t_y))(f(t_y) - f(x_0)) - \frac{\varepsilon}{2}f(x_0)\| \\
&\leq \frac{2}{\varepsilon}[(\frac{\varepsilon}{2} - d(x,t_x))\|f\|_{Lip}D + (\frac{\varepsilon}{2} - d(y,t_y))\|f\|_{Lip}D] \\
&\leq \frac{2D}{\varepsilon}\|f\|_{Lip}d(x,y)
\end{aligned}$$

On the other hand, if $t_x = t_y$ then

$$\begin{aligned}
\|\widetilde{f}(x) - \widetilde{f}(y)\| &= \frac{2}{\varepsilon}\|d(x,t_x)f(x_0) + (\frac{\varepsilon}{2} - d(x,t_x))f(t_x) - d(y,t_x)f(x_0) - (\frac{\varepsilon}{2} - d(y,t_x))f(t_x)\| \\
&= \frac{2}{\varepsilon}|d(x,t_x) - d(y,t_x)|\|f(x_0) - f(t_x)\|
\end{aligned}$$

$$\leq \frac{2D}{\varepsilon}\|f\|_{Lip}d(x,y)$$

All the other cases are easer. $\square$

We shall focus on the following

Theorem 1.14 (Lee-Naor [8]). *Let (X,d) be a doubling metric space, $Y \subseteq X$ and Z be a Banach space. Then every L-Lipschitz function $f : X \longrightarrow Z$ can be extended to a function $\widetilde{f} : X \longrightarrow Z$ such that $\|\widetilde{f}\|_{Lip} \leq K\|f\|_{Lip}$. Furthermore, the extension depends linearly and continuously on f.*

Before to go on, let us make some comment. For any pointed metric space $(X, d, \bar{x})$ and Banach space Z, we denote by $\mathrm{Lip}_0(X,Z)$ (omitting for notational simplicity the dependence on d and $\bar{x}$) the Banach space of all Z-valued Lipschitz functions on X which vanish at $\bar{x}$, equipped with the natural norm

$$\|f\|_{Lip_0(X,Z)} = \sup\{\frac{\|f(x)-f(y)\|}{d(x,y)} :\ x \neq y \text{ in } X\}.$$

The Theorem of Lee and Noar precisely states that there exists a bounded linear operator

$$T : \mathrm{Lip}_0(Y,Z) \longrightarrow \mathrm{Lip}_0(X,Z)$$

such that $T(f)|_Y = f$ for every $f \in \mathrm{Lip}_0(Y,Z)$. Such an operator is called extension operator.

For simplicity, we shall treat the case $Z = \mathbb{R}$, real valued Lipschitz functions.

For all $x \in X$, the Dirac measure δ_x defines a continuous linear functional on $\mathrm{Lip}_0(X)$, defined by $\langle f, \delta_x\rangle = f(x)$, with $\|\delta_x\| \leq d(x,\bar{x})$

Proposition 1.15. *Let $(X,d,\bar{x})$ be any pointed metric space and let M be a closed subset with $\bar{x} \in M$. Then the following properties are equivalent:*

(a) there exists a bounded linear extension operator

$$E : Lip_0(M) \longrightarrow Lip_0(X);$$

(b) there exists a Lipschitz map $\hat{\delta} : X \longrightarrow Lip_0(M)^*$ such that the following diagram commutes

$$\begin{tikzcd} M \arrow[r, hook, "\delta"] \arrow[d] & Lip_0(M)^* \\ X \arrow[ur, "\hat{\delta}"] & \end{tikzcd}$$

In addition, for any $\hat{\delta}$ as in (b) one has $\|E\| = \|\hat{\delta}\|$.

Proof. $(a) \Rightarrow (b)$ It is enough to define $\hat{\delta} : X \longrightarrow \mathrm{Lip}_0(M)^*$ by

$$\hat{\delta}(x) = E^*(\delta_x) \quad \forall x \in X.$$

Let us first observe that $\hat{\delta}(x) \in \mathrm{Lip}_0(M)^*$ for every $x \in X$. Linearity follows directly by the definition of $\hat{\delta}$. Moreover,

$$\begin{aligned} \|\hat{\delta}(x)\|_{Lip_0(M)^*} &= \sup_{\|g\|_{Lip_0(M)} \leq 1} \langle g, \hat{\delta}(x) \rangle \\ &= \sup_{\|g\|_{Lip_0(M)} \leq 1} \langle E(g), \delta_x \rangle \\ &\leq \|E\| d(x, \bar{x}). \end{aligned}$$

Of course, $\hat{\delta}(x) = \delta_x$ for every $x \in M$. Finally, since $\|Eg\|_{Lip_0(X)} \leq \|E\| \|g\|_{Lip_0(M)}$ for all $g \in \mathrm{Lip}_0(M)$, we get

$$\begin{aligned} \|\hat{\delta}(x) - \hat{\delta}(y)\|_{Lip_0(M)^*} &= \sup_{\|g\|_{Lip_0(M)} \leq 1} |\langle g, \hat{\delta}(x) - \hat{\delta}(y) \rangle| \\ &= \sup_{\|g\|_{Lip_0(M)} \leq 1} |E(g)(x) - E(g)(y)| \\ &\leq \|E\| d(x, y). \end{aligned}$$

It follows that

$$\|\hat{\delta}\| \leq \|E\|. \tag{1.3}$$

$(b) \Rightarrow (a)$ Let us define $E : \mathrm{Lip}_0(M) \longrightarrow \mathrm{Lip}_0(X)$ by

$$E(g)(x) = \langle \hat{\delta}(x), g \rangle \quad \forall x \in X, \forall g \in \mathrm{Lip}_0(M).$$

Of course E is a bounded linear operator. Let us estimate its norm: for $g \in \mathrm{Lip}_0(M)$ with $\|g\|_{Lip_0} \leq 1$ we have

$$\begin{aligned} |E(g)(x) - E(g)(y)| &= |\langle \hat{\delta}(x), g \rangle - \langle \hat{\delta}(y), g \rangle| \\ &= |\langle \hat{\delta}(x) - \hat{\delta}(y), g \rangle| \\ &\leq \|\hat{\delta}(x) - \hat{\delta}(y)\|_{Lip_0(M)^*} \\ &\leq \|\hat{\delta}\| d(x, y). \end{aligned}$$

Thus

$$\|E\| \leq \|\hat{\delta}\|. \tag{1.4}$$

Finally, $E(g)(x) = \langle \hat{\delta}(x), g \rangle = \langle \delta(x), g \rangle = g(x)$ for all $x \in M$. Therefore, E is an extension operator and (1.3) and (1.4) give $\|E\| = \|\hat{\delta}\|$. $\square$

An immediate consequence of the previous Proposition is the finite extension property for Lipschitz maps.

Theorem 1.16. *Let $(X, d, \bar{x})$ be a pointed metric space and let $M \subseteq X$ be a closed subspace with $\bar{x} \in M$ and with the following* finite extension Lipschitz property*:*

(F) For every $F \subseteq M$ finite with $\bar{x} \in F$ there exists a linear extension operator
$$E_F : Lip_0(F) \longrightarrow Lip_0(X)$$
with $\|E_F\| \leq C$, for some constant independent of F.

Then, there exists a linear extension operator

$$E : Lip_0(M) \longrightarrow Lip_0(X)$$

with $\|E\| \leq C$.

Proof. Firstly, let us notice that if $R_F : \mathrm{Lip}_0(M) \longrightarrow \mathrm{Lip}_0(F)$ denotes the restriction operator, since R_F is continuous and surjective the dual operator

$$R_F^* : \mathrm{Lip}_0(F)^* \hookrightarrow \mathrm{Lip}_0(M)^*$$

is continuous and injective, hence its range is a closed subspace of $\mathrm{Lip}_0(M)^*$.

By the previous proposition, we can translate the hypothesis by the following: for every $F \subseteq M$ finite with $\bar{x} \in F$ there exists a Lipschitz map $f_F : X \longrightarrow \mathrm{Lip}_0(F)^*$ such that the following diagram commutes

$$\begin{tikzcd} F \arrow[r, hook, "\delta"] \arrow[d] & \mathrm{Lip}_0(F)^* \\ X \arrow[ur, "f_F"] & \end{tikzcd}$$

This tells us that

(*i*) $f_F(x) = \delta_x$ for every $x \in F$;

(*ii*) $\|f_F(x) - f_F(y)\|_{Lip_0(F)^*} \leq Cd(x, y)$ for every x, $y \in F$.

Still by the previous proposition, we need to build a Lipschitz map $f : X \longrightarrow \mathrm{Lip}_0(M)^*$ such that the diagram

$$\begin{tikzcd} M \arrow[r, hook, "\delta"] \arrow[d] & \mathrm{Lip}_0(M)^* \\ X \arrow[ur, "f"] & \end{tikzcd}$$

commutes.

Let us denote by $B_F(r) = \{x^* \in \mathrm{Lip}_0(F)^* : \ \|x^*\|_{Lip_0(F)^*} \leq r\}$ be the closed ball in $\mathrm{Lip}_0(F)^*$ centered at 0 with radius $r > 0$. Since $\mathrm{Lip}_0(F)^*$ is finite dimensional, each ball $B_F(r)$ is a compact set, and then, by the natural embeddings R_F^*, $B_F(r)$ can be seen as a compact subset of $\mathrm{Lip}_0(M)^*$.

In particular, (ii) implies that

$$f_F \in \prod_{x \in X} B_F(Cd(x, \bar{x})) = B \subseteq (\mathrm{Lip}_0(M)^*, \mathrm{weak}^*)^X.$$

When we partially order the collection $\mathcal{F}$ of finite subsets of M by inclusion we have a net; hence, by the compactness of B in $(\mathrm{Lip}_0(M)^*, \mathrm{weak}^*)^X$, there exist a cofinal subnet $\mathcal{G}$ and $f : X \longrightarrow \mathrm{Lip}_0(M)^*$ such that

$$\lim_{F \in \mathcal{G}} f_F = f \qquad \text{in } (\mathrm{Lip}_0(M)^*, \mathrm{weak}^*)^X.$$

Now, by cofinality, for each $x \in M$ there exists $F \in \mathcal{G}$ such that $x \in F$. Since the convergence is in weak* topology, in particular we have pointwise convergence. Thus (i) implies that

$$\delta_x = \lim_{F \in \mathcal{G}} f_F(x) = f(x).$$

Similarly, for every x, $y \in X$, by (ii),

$$\begin{aligned}\|f(x) - f(y)\|_{Lip_0(M)^*} &= \lim_F \|f_F(x) - f_F(y)\|_{Lip_0(M)^*} \\ &\leq Cd(x, y).\end{aligned}$$

Thus, f is Lipschitz and $\|f\|_{Lip} \leq C$. □

The previous theorem permit us to assume that Y in Theorem 1.14 is finite; i.e., $Y = \{\bar{x}, x_1, \ldots x_n\}$. For each $R > r > 0$, let us denote by $C(X, R, r)$ the largest cardinality of a set $N \subseteq X$ such that for every distinct $x, y \in N$, $r \leq d(x, y) \leq R$. Let us note that if λ denotes the doubling constant of X, then for every $\Delta > 0$

$$C(X, 2\Delta, \frac{\Delta}{4}) \leq \lambda^4.$$

Indeed, let N such that $\frac{\Delta}{4} \leq d(x, y) \leq 2\Delta$, for every distinct $x, y \in N$. Then N is contained in a ball of radius 2Δ which can be covered by λ^4 balls of radius $\frac{\Delta}{8}$. Since every such ball contains at most one point of N, we see that $|N| \leq \lambda^4$.

Let us recall that $N \subseteq Y$ is said to be a Δ-net if for every $x, y \in N$, $d(x, y) \geq \Delta$ and $Y \subseteq \cup_{y \in N} B(y, \Delta)$.

Lemma 1.17. *Let $\Delta > 0$ and M be a Δ-net of Y. Then for each σ random permutation on M there exists a partition $\{C_y(\sigma) : \ y \in M\}$ of Y and a probability measure μ over the set of all random permutation on M, such that*

(i) $diam C_y(\sigma) \leq 4\Delta$;

(ii) $\mu\left(\bigcup_{y\in M}\{\sigma : \ d(x, Y \setminus C_y(\sigma)) \geq \varepsilon\Delta\}\right) \geq \frac{1}{2}$, *for all* $x \in Y$,

where $\varepsilon = \frac{1}{4\cdot 64\cdot \ln(\lambda)}$.

Proof. Let σ be a random permutation on M and choose $R \in]\Delta, 2\Delta]$ uniformly at random. For each $y \in M$

$$C_y(\sigma) := \{x \in Y : \ x \in B(y, R) \text{ and } \sigma(y) < \sigma(z) \text{ for all } z \in M \text{ with } x \in B(z, R)\}.$$

Clearly, $\text{diam}(C_y(\sigma)) \leq 4\Delta$. Finally, $P = \{C_y(\sigma)\}_{y\in M}$ is a partition of Y because M is a Δ-net and $R \geq \Delta$. Now, fix a value $t \in [0, \Delta]$ and some $x \in Y$. Let $W = B(x, 2\Delta + t) \cap M$, and note that $m = |W| \leq C(Y, 6\Delta, \Delta)$. Arrange the points $w_1, \ldots, w_m \in W$ in order of decreasing distance from x, and let I_k be the interval

$$[d(x, w_k) - t, d(x, w_k) + t].$$

We say that $B(x,t)$ is *cut* by a cluster C_{w_k}, if $C_{w_k} \cap B(x,t) \neq \emptyset$ but $B(x,t) \not\subseteq C_{w_k}$. Finally, write $\mathcal{E}_k$ for the event that w_k is the minimal element (according to σ) in W for which C_{w_k} cuts $B(x,t)$. Observe that for every $1 \leq k \leq m$, $\Pr[\mathcal{E}_k | R \in I_k] \leq \frac{1}{k}$, since we require that in the uniformly random permutation induced by σ on $\{w_1, \ldots, w_m\}$, w_k appears before $w_1, \ldots, w_{k-1}$. Recall that $\Pr[\mathcal{E}_k | R \in I_k]$ is the probability of event $\mathcal{E}_k$ given that event $R \in I_k$ has already taken place.

Now

$$\begin{aligned}
\Pr[B(x,t) \text{ is cut}] &\leq \sum_{k=1}^{m} \Pr[\mathcal{E}_k] \\
&= \sum_{k=1}^{m} \Pr[R \in I_k] \cdot \Pr[\mathcal{E}_k | R \in I_k] \\
&\leq \sum_{k=1}^{m} \frac{2t}{\Delta} \cdot \frac{1}{k} \\
&\leq \frac{2t}{\Delta}(1 + \ln m)
\end{aligned}$$

$$\leq \frac{8t}{\Delta} \ln(C(Y, 6\Delta, \Delta)).$$

Setting

$$t = \frac{\Delta}{64 \ln(C(Y, 6\Delta, \Delta))}$$

yields the required result; i.e,

$$\Pr[B(x,t) \text{ is cut}] \leq \frac{1}{8}$$

Finally, note that when $B(x,t)$ is not cut, we have $B(x,t) \subseteq C$ for some $C \in P$. Therefore,

$$\begin{aligned}
\Pr\left[\bigcup_{y\in M}\{\sigma:\ d(x, Y\setminus C_y(\sigma)) \geq \varepsilon\Delta\}\right] &\geq \Pr\left[\bigcup_{y\in M}\{\sigma:\ B(x,t) \subseteq C_y(\sigma)\}\right] \\
&\geq \Pr\left[B(x,t) \text{ is not cut}\right] \\
&\geq 1 - \frac{1}{8} \\
&\geq \frac{1}{2}
\end{aligned}$$

□

For each $y \in M$ let $\gamma_y(\sigma)$ be the minimal element of M (with respect to σ) in $C_y(\sigma)$.

Now we would like to extend those partition $C_y(\sigma)$'s to the whole space X. For each $x \in X$ let us denote by $t_x \in Y$ the minimum distance element with respect to Y; i.e., $d(x, t_x) = d(x, Y)$. Let us denote by Ω the set of all permutation on M. Now, for each $\sigma \in \Omega$ and $y \in M$ we define

$$\hat{C}_y(\sigma) := \begin{cases} C_y(\sigma) \bigcup \{x \in X:\ d(t_x, Y\setminus C_y(\sigma)) \geq \frac{\varepsilon\Delta}{2} \text{ and } d(x, t_x) \leq \frac{\varepsilon\Delta}{4}\}, \text{ if } y \in M \\ \{y\}, \text{ if } y \notin \bigcup_{z\in M} \hat{C}_z(\sigma) \end{cases}$$

and $\hat{\gamma}_y(\sigma) : \Omega \longrightarrow Y$ by

$$\hat{\gamma}_y(\sigma) = \begin{cases} \gamma_y(\sigma), \text{ if } y \in M \\ t_y,\ y \notin \bigcup_{z\in M} \hat{C}_z(\sigma). \end{cases}$$

Let us note that $\{\hat{C}_y(\sigma)\}_y$ is a partition of whole space X and $\hat{\gamma}_y$ is measurable (in a trivial way!). Let us note that, whenever $x, z \in \hat{C}_y(\sigma)$ we have

$$d(x,z) \leq d(x, t_x) + d(z, t_z) + d(t_x, t_z)$$

$$\leq \frac{\varepsilon\Delta}{4} + \frac{\varepsilon\Delta}{4} + \mathrm{diam}(C_y(\sigma))$$
$$\leq (4 + \frac{\varepsilon}{2})\Delta.$$

Let us also observe that, if $d(x, Y) \leq \frac{\varepsilon\Delta}{16}$ then

$$\mu\left(\bigcup_{y\in M}\{\sigma:\ d(x, X\setminus \hat{C}_y(\sigma)) \geq \frac{\varepsilon\Delta}{16}\}\right) \geq \frac{1}{2}.$$

Since $d(t_x, Y\setminus C_y(\sigma)) \geq \varepsilon\Delta$, for some $y \in M$, our goal will be to show that in this case $d(x, X\setminus \hat{C}_y(\sigma)) \geq \frac{\varepsilon\Delta}{16}$. Assume to the contrary that there is some $z \in X \setminus \hat{C}_y(\sigma)$ with $d(x, z) \leq \frac{\varepsilon\Delta}{16}$. Observe that

$$\begin{aligned} d(t_x, t_z) &\leq d(x, t_x) + d(z, t_z) + d(x, z) \\ &\leq d(x, Y) + (d(z, x) + d(x, Y)) + \frac{\varepsilon\Delta}{16} \\ &< \frac{\varepsilon\Delta}{16} + 2\frac{\varepsilon\Delta}{16} + \frac{\varepsilon\Delta}{16} \\ &< \frac{\varepsilon\Delta}{2} \end{aligned}$$

Hence,

$$d(t_z, Y\setminus C_y(\sigma)) \geq d(t_x, Y\setminus C_y(\sigma)) - d(t_x, t_z) > \frac{\varepsilon\Delta}{2}.$$

Since we also have that

$$d(z, t_z) = d(z, Y) \leq d(x, Y) + d(x, z) \leq \frac{\varepsilon\Delta}{4},$$

which would implies that $z \in \hat{C}_y(\sigma)$, against the choice of z.

Since all the construction above depends on Δ, we can reformulate it for $\Delta = 2^n$, $n \in \mathbb{Z}$.

Proposition 1.18. *For every $n \in \mathbb{Z}$ there exists (Ω_n, μ_n) a probability measure space, a set of index I,*

$$\gamma_i^n : \Omega_n \longrightarrow Y$$

a measurable function, and for all $\omega \in \Omega_n$ a partition $\{\Gamma_n^i(\omega)\}_{i\in I}$ such that diam$(\Gamma_n^i(\omega)) \leq 2^n$ and if $x \in X$ with $d(x, Y) \leq \varepsilon 2^n$ then

$$\mu\left(\bigcup_{i\in I}\{\omega:\ d(x, X\setminus \Gamma_n^i(\omega)) \geq \varepsilon 2^n\}\right) \geq \frac{1}{2}.$$

Now, we are in position to prove the main result of this section

Proof. (of Theorem 1.14).

Let $\varphi : \mathbb{R} \longrightarrow [0, +\infty[$ be 2-Lipschitz function with $\mathrm{supp}\varphi \subseteq [\frac{1}{2}, 4]$ and $\varphi = 1$ on $[1, 2]$ and define

$$\varphi_n(x) := \varphi\left(\frac{d(x, Y)}{\varepsilon 2^{n-3}}\right).$$

If on I we consider the counting measure, let (Ω, μ) be the disjoint union of $(I \times \Omega_n)_{n\in\mathbb{Z}}$. Let us define

$$\gamma : \Omega \longrightarrow Y$$

by

$$\gamma(i, \omega) = \gamma_n^i(\omega), \text{ if } i \in I \text{ and } \omega \in \Omega_n.$$

Let $g : [0, +\infty[\longrightarrow [0, +\infty[$ given by

$$g(x) = \begin{cases} 1, \text{ if } x \geq 2 \\ x - 1, \text{ if } 1 \leq x \leq 2 \\ 0, \text{ if } 0 \leq x \leq 1. \end{cases}$$

Finally, let

$$\theta_\omega^n(x) := g\left(\frac{1}{\varepsilon 2^{n-1}} \cdot \sum_{i\in I} \min\{d\left(x, X \setminus \Gamma_n^i(\omega)\right),\ 2^n\}\right) = g(\pi_\omega^n(x)).$$

Observe that since $\{\Gamma_n^i(\omega)\}_{i\in I}$ is a partition of X, the above sum consist of only one element.

We are ready to define the main tool of the construction: let

$$\Phi : \Omega \times X \longrightarrow [0, +\infty[$$

$$\Phi(i, \omega, x) := \frac{1}{S(x)} \theta_\omega^n(x)\varphi_n(x) \cdot \chi_{\Gamma_n^i(\omega)}(x),$$

where

$$\begin{aligned} S(x) &= \sum_{n\in\mathbb{Z}} \sum_{i\in I} \int_{\Omega_n} \theta_\omega^n(x)\varphi_n(x) \cdot \chi_{\Gamma_n^i(\omega)}(x)\ d\mu_n(\omega) \\ &= \sum_{n\in\mathbb{Z}} \varphi_n(x) \int_{\Omega_n} \theta_\omega^n(x) d\mu_n(\omega). \end{aligned}$$

Important: observe that $\mathrm{supp}\varphi_n \subseteq \{x \in X :\ \varepsilon 2^{n-4} \leq d(x, Y) \leq \varepsilon 2^{n-4}\}$ so the sum above consists of at most 5 terms.

Of course Φ has the following properties (*i*) $\Phi(\cdot,\cdot,x) = 0$ if $x \in Y$;

(*ii*) $\|\Phi(\cdot,\cdot,x)\|_{L^1(\Omega,\mu)} = 1$ if $x \in X \setminus Y$.

Key: for all $x, y \in X$,

$$\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n} d(\gamma(i,\omega),x)\,|\Phi(i,\omega,x)-\Phi(i,\omega,y)|\ d\mu_n(\omega) \le \frac{2C}{\varepsilon}\cdot d(x,y) \quad (1.5)$$

Fix $\omega \in \Omega_n$ and assume that $\Phi(i,\omega,x) \neq \Phi(i,\omega,y)$. Then either $\Phi(i,\omega,x) > 0$ or $\Phi(i,\omega,y) > 0$. In the first case $x \in \Gamma_n^i(\omega)$. Notice that either $d(\gamma(i,\omega),\Gamma_n^i(\omega)) = 0$ or $\Gamma_n^i(\omega) = \{x\}$ and then $\gamma(i,\omega) = t_x$. In any case,

$$\begin{aligned}
d(\gamma(i,\omega),x) &\le d(\gamma(i,\omega),\Gamma_n^i(\omega)) + \mathrm{diam}(\Gamma_n^i(\omega))\\
&\le d(Y,\Gamma_n^i(\omega)) + 2^n\\
&= d(x,Y) + 2^4\cdot 2^{n-4}\\
(\text{since } \varphi_n(x) > 0) \quad &\le d(x,Y) + \frac{2^4}{\varepsilon}d(x,Y)\\
&\le \frac{18}{\varepsilon}d(x,Y)\\
&\le d(x,y) + \frac{18}{\varepsilon}\max\{d(x,Y),\ d(y,Y)\}.
\end{aligned}$$

In the second case, $\Phi(i,\omega,y) > 0$, so

$$\begin{aligned}
d(\gamma(i,\omega),x) &\le d(\gamma(i,\omega),y) + d(x,y)\\
&\le d(y,Y) + \mathrm{diam}(\Gamma_n^i(\omega))\ + d(x,y)\\
&\le d(y,Y) + 2^n + d(x,y)\\
&\le d(x,y) + \frac{18}{\varepsilon}\max\{d(x,Y),\ d(y,Y)\}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n} d(\gamma(i,\omega),x)\,|\Phi(i,\omega,x)-\Phi(i,\omega,y)|\ d\mu_n(\omega)\\
&\le d(x,y)\cdot\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n}[\Phi(i,\omega,x)+\Phi(i,\omega,y)]\ d\mu_n(\omega)\\
&+\frac{18}{\varepsilon}\max\{d(x,Y),\ d(y,Y)\}\cdot\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n}|\Phi(i,\omega,x)-\Phi(i,\omega,y)|\ d\mu_n(\omega)\\
&= 2d(x,y) + \frac{18}{\varepsilon}\max\{d(x,Y),\ d(y,Y)\}\cdot\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n}|\Phi(i,\omega,x)-\Phi(i,\omega,y)|\ d\mu_n(\omega)
\end{aligned}$$

Thus, to reach (1.5), it is enough to show that

$$\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n}|\Phi(i,\omega,x)-\Phi(i,\omega,y)|\ d\mu_n(\omega)\le C'\cdot\frac{d(x,y)}{\max\{d(x,Y),\ d(y,Y)\}} \tag{1.6}$$

Notice that, if $d(x,y)\ge d(\{x,y\},Y)$ then it is enough to choose any $C'>4$. Indeed, $d(x,Y)\le d(x,y)+d(y,Y)\le 2d(x,y)$ and analogously $d(y,Y)\le 2d(x,y)$. Hence the right-hand side of (1.6) is greater than 2 while the left-hand side of (1.6) is at most 2.

Then we can assume that $d(x,y)<d(\{x,y\},Y)$. Moreover, we assume as well that $d(x,Y)\ge d(y,Y)$. Now,

$$\begin{aligned}
&\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n}|\Phi(i,\omega,x)-\Phi(i,\omega,y)|\ d\mu_n(\omega)\\
&=\sum_{n\in\mathbb{Z}}\int_{\Omega_n}\sum_{i\in I}\left|\frac{\theta^n_\omega(x)\varphi_n(x)\chi_{\Gamma^i_n(\omega)}(x)S(y)-\theta^n_\omega(y)\varphi_n(y)\chi_{\Gamma^i_n(\omega)}(y)S(x)}{S(x)S(y)}\right|\ d\mu_n(\omega)\\
&\le\sum_{n\in\mathbb{Z}}\int_{\Omega_n}\sum_{i\in I}\frac{\left|\theta^n_\omega(x)\varphi_n(x)\chi_{\Gamma^i_n(\omega)}(x)-\theta^n_\omega(y)\varphi_n(y)\chi_{\Gamma^i_n(\omega)}(y)\right|}{S(x)}\ d\mu_n(\omega)\\
&+\left(\sum_{n\in\mathbb{Z}}\int_{\Omega_n}\theta^n_\omega(y)\varphi_n(y)\ d\mu_n(\omega)\right)\frac{|S(x)-S(y)|}{S(x)S(y)}\\
&\le\sum_{n\in\mathbb{Z}}\int_{\Omega_n}\sum_{i\in I}\frac{\left|\theta^n_\omega(x)\varphi_n(x)\chi_{\Gamma^i_n(\omega)}(x)-\theta^n_\omega(y)\varphi_n(y)\chi_{\Gamma^i_n(\omega)}(y)\right|}{S(x)}\ d\mu_n(\omega)\\
&+\left(\sum_{n\in\mathbb{Z}}\int_{\Omega_n}\theta^n_\omega(y)\varphi_n(y)\ d\mu_n(\omega)\right)\cdot\\
&\cdot\sum_{k\in\mathbb{Z}}\int_{\Omega_k}\sum_{i\in I}\frac{\left|\theta^n_\tau(x)\varphi_k(x)\chi_{\Gamma^i_k(\tau)}(x)-\theta^n_\tau(y)\varphi_k(y)\chi_{\Gamma^i_k(\tau)}(y)\right|}{S(x)S(y)}\ d\mu_k(\tau)\\
&=\frac{2}{S(x)}\sum_{n\in\mathbb{Z}}\int_{\Omega_n}\sum_{i\in I}\left|\theta^n_\omega(x)\varphi_n(x)\chi_{\Gamma^i_n(\omega)}(x)-\theta^n_\omega(y)\varphi_n(y)\chi_{\Gamma^i_n(\omega)}(y)\right|\ d\mu_n(\omega).
\end{aligned}$$

Let $n_0\in\mathbb{Z}$ such that $\frac{d(x,Y)}{\varepsilon 2^{n_0-3}}\in[1,2]$. Let us denote by $\pi^n_\omega(x)=\sum_{i\in I}\min\{d\left(x,X\setminus\Gamma^i_n(\omega)\right),\ 2^n\}$.

$$\begin{aligned}
S(x)&\ge\sum_{n:\ d(x,Y)\le\varepsilon 2^n}\varphi_n(x)\int_{\{\omega:\Omega_n:\ \pi^n_\omega(x)\ge\varepsilon 2^n\}}g\left(\frac{\pi^n_\omega(x)}{\varepsilon 2^{n-1}}\right)\ d\mu_n(\omega)\\
&\ge\varphi\left(\frac{d(x,Y)}{\varepsilon 2^{n_0-3}}\right)\mu_{n_0}\left(\bigcup_{i\in I}\{\omega\in\Omega_{n_0}:\ d(x,Y\setminus\Gamma^i_{n_0}(\omega))\ge\varepsilon 2^{n_0}\}\right)
\end{aligned}$$

$$= \mu_{n_0}\left(\bigcup_{i\in I}\{\omega\in\Omega_{n_0}:\ d(x, Y\setminus\Gamma^i_{n_0}(\omega))\geq\varepsilon 2^{n_0}\}\right)$$
$$\geq\frac{1}{2}$$

Fix $n\in\mathbb{Z}$ and $\omega\in\Omega_n$. Assume $\varphi_n(x)+\varphi_n(y)>0$. In this case $\{d(x,Y),\ d(y,Y)\}\cap[\varepsilon 2^{n-4},\ \varepsilon 2^{n-1}]\neq\emptyset$, so in particular $d(y,Y)\leq\varepsilon 2^{n-1}$. Since we are assuming $d(x,y)<d(y,Y)$, then

$$d(x,Y)\leq d(x,y)+d(y,Y)\leq 2d(y,Y)\leq\varepsilon 2^n.$$

If $x,y\in\Gamma^j_n(\omega)$ for soem $j\in I$, since g is 1-Lipschitz,

$$\begin{aligned}
&\sum_{i\in I}|\theta^n_\omega(x)\varphi_n(x)\chi_{\Gamma^i_n(\omega)}(x)-\theta^n_\omega(y)\varphi_n(y)\chi_{\Gamma^i_n(\omega)}(y)|\\
&=\left|\theta^n_\omega(x)\varphi\left(\frac{d(x,Y)}{\varepsilon 2^{n-3}}\right)-\theta^n_\omega(y)\varphi\left(\frac{d(y,Y)}{\varepsilon 2^{n-3}}\right)\right|\\
&\leq\varphi\left(\frac{d(x,Y)}{\varepsilon 2^{n-3}}\right)|\theta^n_\omega(x)-\theta^n_\omega(y)|+\theta^n_\omega(y)\left|\varphi\left(\frac{d(x,Y)}{\varepsilon 2^{n-3}}\right)-\varphi\left(\frac{d(y,Y)}{\varepsilon 2^{n-3}}\right)\right|\\
&\leq\varphi\left(\frac{d(x,Y)}{\varepsilon 2^{n-3}}\right)|\pi^n_\omega(x)-\pi^n_\omega(y)|+\theta^n_\omega(y)\left|\varphi\left(\frac{d(x,Y)}{\varepsilon 2^{n-3}}\right)-\varphi\left(\frac{d(y,Y)}{\varepsilon 2^{n-3}}\right)\right|\\
&\leq d(x,y)+\theta^n_\omega(y)\frac{16d(x,y)}{\varepsilon 2^n}\\
&\leq d(x,y)+\frac{16d(x,y)}{\varepsilon 2^n}\\
&\leq\frac{6d(x,y)}{\varepsilon 2^n}\\
&\leq\frac{6d(x,y)}{d(x,Y)}
\end{aligned}$$

On the other hand, if there exist distinct $i,j\in I$ such that $x\in\Gamma^i_n(\omega)$ and $y\in\Gamma^j_n(\omega)$, then using the fact that $\pi^n_\omega(x),\ \pi^n_\omega(y)\leq d(x,y)$, we get

$$\begin{aligned}
&\sum_{i\in I}|\theta^n_\omega(x)\varphi_n(x)\chi_{\Gamma^i_n(\omega)}(x)-\theta^n_\omega(y)\varphi_n(y)\chi_{\Gamma^i_n(\omega)}(y)|\\
&\leq g\left(\frac{\pi^n_\omega(x)}{\varepsilon 2^{n-1}}\right)+g\left(\frac{\pi^n_\omega(y)}{\varepsilon 2^{n-1}}\right)\\
&\leq\frac{\pi^n_\omega(x)}{\varepsilon 2^{n-1}}+\frac{\pi^n_\omega(y)}{\varepsilon 2^{n-1}}
\end{aligned}$$

$$\leq \frac{4d(x,y)}{d(x,Y)}$$

Putting all together,

$$\begin{aligned}
&\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n}|\Phi(i,\omega,x)-\Phi(i,\omega,y)|\ d\mu_n(\omega)\\
&\leq \frac{2}{S(x)}\sum_{n\in\mathbb{Z}}\int_{\Omega_n}\sum_{i\in I}\left|\theta^n_\omega(x)\varphi_n(x)\chi_{\Gamma^i_n(\omega)}(x)-\theta^n_\omega(y)\varphi_n(y)\chi_{\Gamma^i_n(\omega)}(y)\right|\ d\mu_n(\omega)\\
&\leq \sum_{n\in\mathbb{Z}}\int_{\Omega_n}\sum_{i\in I}\left|\theta^n_\omega(x)\varphi_n(x)\chi_{\Gamma^i_n(\omega)}(x)-\theta^n_\omega(y)\varphi_n(y)\chi_{\Gamma^i_n(\omega)}(y)\right|\ d\mu_n(\omega)\\
&\leq \sum_{n:\{x,y\}\cap supp\varphi_n\neq\emptyset}\int_{\Omega_n}\frac{6d(x,y)}{d(x,Y)}\ d\mu_n\\
&\leq \frac{10\cdot 6\cdot d(x,y)}{d(x,Y)}
\end{aligned}$$

where we have used the fact that, for every $z\in X$, $|\{n:\ z\in \operatorname{supp}\varphi_n\}|\leq 5$, which concludes (1.6).

Now, we are ready to define the extension operator. Let

$$T:\mathrm{Lip}_0(Y)\longrightarrow \mathrm{Lip}_0(X)$$

defined by

$$T(f)(x)=\begin{cases}\int_\Omega f(\gamma(\omega))\Phi(\omega,x)\ d\mu(\omega), & \text{if } x\in X\setminus Y\\ \\ f(x), & \text{if } x\in Y\end{cases}\tag{1.7}$$

Let us collect some information. For every $y\in Y$,

$$\begin{aligned}
&\int_\Omega |f(\gamma(\omega))|\Phi(\omega,x)\ d\mu(\omega)\\
&\leq \int_\Omega[|f(\gamma(\omega))-f(x)|+|f(x)|]\Phi(\omega,x)\ d\mu(\omega)\\
\text{(by } (i) \text{ and } (ii) \text{ above)}\quad &\leq \|f\|_{Lip}\int_\Omega d(\gamma(\omega),x)|\Phi(\omega,x)-\Phi(\omega,y)|\ d\mu(\omega)+|f(y)|\\
\text{(by (1.5) above)}\quad &\leq \|f\|_{Lip}Cd(x,y)+|f(y)|<\infty.
\end{aligned}$$

Therefore T is well defined. Of course $T(f)$ extends f. Let us show that $T(f)$ is Lipschitz. Let us consider $x,y\in X$ such that al least one of them not in

Y; i.e., $y \in X \setminus Y$. Take $z \in Y$ such that $d(x,z) = d(x,Y)$ in case $x \in X \setminus Y$ otherwise $z = x$, and observe that

$$T(f)(y) - T(f)(x) = \int_\Omega [f(\gamma(\omega)) - f(z)] \cdot [\Phi(\omega, y) - \Phi(\omega, x)]\; d\mu(\omega)$$

Thus

$$\begin{aligned}
|T(f)(y) - T(f)(x)| &\le \|f\|_{Lip} \int_\Omega d(\gamma(\omega), z)|\Phi(\omega, y) - \Phi(\omega, x)|\; d\mu(\omega) \\
&\le \|f\|_{Lip} \int_\Omega [d(\gamma(\omega), x) + d(x, z)]|\Phi(\omega, y) - \Phi(\omega, x)|\; d\mu(\omega) \\
(\gamma(\omega) \in Y) \quad &\le \|f\|_{Lip} \int_\Omega [d(\gamma(\omega), x) + d(x, \gamma(\omega))]|\Phi(\omega, y) - \Phi(\omega, x)|\; d\mu(\omega) \\
&= 2\|f\|_{Lip} \int_\Omega d(\gamma(\omega), x)|\Phi(\omega, y) - \Phi(\omega, x)|\; d\mu(\omega) \\
(\text{by } (1.5)) \quad &\le 2C\|f\|_{Lip} d(x, y).
\end{aligned}$$

$\square$

Chapter 2

Differentiability

Recall that a function $f : \Omega \longrightarrow \mathbb{R}^m$, where $\Omega \subseteq \mathbb{R}^n$ is open, is *differentiable at* $a \in \Omega$ if there exists a linear map $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that

$$\lim_{x \to a} \frac{\|f(x) - f(a) - L(x-a)\|}{\|x-a\|} = 0. \tag{2.1}$$

If such a linear map L exists, it is unique, called the *derivative* of f at a, and denoted by $Df(a)$. We also note that $f = (f_1, \dots, f_m)$ is differentiable at a if and only if each of the coordinate functions f_i are differentiable at a.

To analyze (2.3) more carefully, suppose $f : \Omega \longrightarrow \mathbb{R}$ is a real-valued function, differentiable at a point $a \in \Omega$. For $t \in \mathbb{R}$, $t \neq 0$, consider the functions

$$f_t(x) := \frac{f(a+tx) - f(a)}{t},$$

which are defined for t small enough. Then

$$\lim_{t \to 0} |f_t(x) - L(x)| = 0$$

uniformly in $x \in \mathbb{B}^n$. This procedure can be reversed and we conclude that a function f is differentiable at a point a if and only if the functions $(f_t)_t$ converges uniformly in $\mathbb{B}^n$ to a linear map as $t \to 0$.

Assume now that $f : \Omega \longrightarrow \mathbb{R}$ is L-Lipschitz, and that $a \in \Omega$. Then the family $(f_t)_t$ consists of uniformly bounded L-Lipschitz functions on $\mathbb{B}^n$ (for small enough t). The Arzelá-Ascoli theorem guarantees that there is a subsequence of the sequence $(f_t)_t$ that converges uniformly to an L-Lipschitz function on $\mathbb{B}^n$. What Rademacher's theorem claims, in effect, is that for almost all points a in Ω this limit is independent of the subsequence, and that the limit function is linear.

Before to go on, we need to recall some classical covering theorem.

Theorem 2.1. *Every family $\mathcal{F}$ of balls of uniformly bounded diameter in a metric space X contains a disjointed subfamily $\mathcal{G}$ such that*

$$\bigcup_{B\in\mathcal{F}} B \subseteq \bigcup_{B\in\mathcal{G}} 5B.$$

Proof. Let Ω denote the set consisting of all disjointed subfamilies $\widetilde{\mathcal{F}}$ of $\mathcal{F}$, partially ordered by inclusion, with the following property: if a ball B from $\mathcal{F}$ meets some ball from $\widetilde{\mathcal{F}}$, then it meets one whose radius is at least half the radius of B. Note that Ω is nonempty because the one ball family $\widetilde{\mathcal{F}} = \{B\}$ is in Ω whenever $B \in \mathcal{F}$ has radius close to the supremum one.

Then, if $\mathcal{C} \subseteq \Omega$ is a chain, it is easy to see that $\widetilde{\mathcal{F}}_0 = \bigcup_{\widetilde{\mathcal{F}}\in\mathcal{C}} \widetilde{\mathcal{F}}$ belongs to Ω, so there is a maximal element $\mathcal{G}$ in Ω. By construction, $\mathcal{G}$ is disjointed. If there is a ball $B \in \mathcal{F}$ that does not meet any ball from $\mathcal{G}$, then pick a ball B_0 from $\mathcal{F}$ such that the radius of B_0 is larger than half of the radius of any other ball that does not meet the balls from $\mathcal{G}$. Then, if a ball B from $\mathcal{F}$ meets a ball from the collection $\mathcal{G}' = \mathcal{G} \cup \{B_0\}$, by construction it meets one whose radius is at least half that of B, showing that $\mathcal{G}'$ belongs to Ω. But this contradicts the maximality of $\mathcal{G}$. Thus, every ball $B = B(x, r)$ from $\mathcal{F}$ meets a ball $B' = B(x', r')$ from $\mathcal{G}$ so that $r \le 2r'$, and the triangle inequality shows that $B \subseteq 5B'$. $\square$

Theorem 2.2 (Vitali covering)**.** *Let A be a subset in a doubling metric measure space (X, d, μ) and let $\mathcal{F}$ be a collection of closed balls centered at A such that*

(a) $A \subseteq \bigcup_{B\in\mathcal{F}} B$*,*

(b) For each point $x \in A$ and $\varepsilon > 0$ there is $B \in \mathcal{F}$ that contains x and $\text{radius}B < \varepsilon$.

Then for each $\varepsilon > 0$ there is a finite disjoint subcollection $\{B_k\}_{k=1}^n$ of $\mathcal{F}$ for which

$$\mu^*\left(A \setminus \bigcup_{k=1}^{n} B_k\right) < \varepsilon,$$

where μ^* denotes the outer measure associate to μ.

Proof. Assume first that A is bounded. Next, we may assume that the balls in $\mathcal{F}$ have uniformly bounded radii; in particular, by the Theorem 2.1, we find a disjointed subcollection $\mathcal{G}$ of $\mathcal{F}$ such that A is contained in $\bigcup_{\mathcal{G}} 5B$. Since $\bigcup_{\mathcal{G}} B$ is contained in some fixed ball and μ is finite and strictly positive on balls, the collection $\mathcal{G}$ is necessarily countable; i.e., $\mathcal{G} = \{B_n,\ n \in \mathbb{N}\}$.

We claim,

$$(*) \quad A \setminus \bigcup_{i=1}^{n} B_i \subseteq \bigcup_{k=n+1}^{\infty} 5B_k, \text{ for each } n \in \mathbb{N}.$$

Indeed, let $a \in A \setminus \bigcup_{i=1}^{N} B_i$. Since the balls in $\mathcal{F}$ are closed we can find $B \in \mathcal{F}$ such that $a \in B$ and $B \cap (\cup_{i=1}^{N} B_i) = \emptyset$. Therefore, by property of $\mathcal{G}$

$$a \in B \subseteq \bigcup_{k=N+1}^{\infty} 5B_k.$$

Since B_i's are disjoints, by σ-additivity of the measure and $(*)$, it is enough to choose $N \in \mathbb{N}$ such that

$$\begin{aligned} \mu^* \left(A \setminus \bigcup_{k=1}^{N} B_k \right) &\leq \mu(\bigcup_{k=N+1}^{\infty} 5B_k) \\ &= \sum_{k=N+1}^{\infty} \mu(B_k) \\ &< \varepsilon. \end{aligned}$$

The case A is unbounded is easy. □

Theorem 2.3 (Lebesgue)**.** *Let $f : (a,b) \longrightarrow \mathbb{R}$ be a monotone function. Then f is differentiable at almost every point in (a,b).*

Proof. Of course, we can assume that the interval (a,b) is bounded and f is increasing. Let

$$D^+ f(x) := \lim_{h \to 0} \left[\sup_{0<|t|<h} \frac{f(x+t) - f(x)}{h} \right]$$

and

$$D^- f(x) := \lim_{h \to 0} \left[\inf_{0<|t|<h} \frac{f(x+t) - f(x)}{h} \right].$$

First, let us note that, for each $\alpha > 0$

$$\lambda^*(\{x \in (a,b) : \ D^+ f(x) \geq \alpha\}) \leq \frac{1}{\alpha}[f(b) - f(a)].$$

Indeed, let $E_\alpha = \{x \in (a,b) : \ D^+ f(x) \geq \alpha\})$ and $0 < \alpha' < \alpha$. Let $\mathcal{F}$ be the collection of all closed, bounded intervals $[c,d]$ contained in (a,b) such that

$f(d) - f(c) \geq \alpha'(d-c)$. Since $D^+f(x) \geq \alpha$ on E_α, we get that $\mathcal{F}$ satisfies the hypothesis of Vitaly's covering theorem above. Then there exists a finite disjoint subcollection $\{[c_k, d_k]\}_{k=1}^N$ of $\mathcal{F}$ for which

$$\lambda^*(E_\alpha \setminus \cup_{k\in\mathbb{N}}[c_k, d_k]) < \varepsilon.$$

Since $E_\alpha \subseteq (\cup_{k\in\mathbb{N}}[c_k, d_k]) \cup (E_\alpha \setminus \cup_{k\in\mathbb{N}}[c_k, d_k])$, by the finite subadditivity of outer measure, we have

$$\begin{aligned}\lambda^*(E_\alpha) &\leq \sum_{i=1}^N (d_i - c_i) + \varepsilon \\ &\leq \frac{1}{\alpha'}\sum_{i=1}^N (f(d_i) - f(c_i)) + \varepsilon.\end{aligned}$$

However, the function f is increasing on (a,b) and $\{[c_k, d_k]\}_{k=1}^N$ are disjoint intervals. Therefore, for each $\varepsilon > 0$ and $0 < \alpha' < \alpha$, we have

$$\lambda^*(E_\alpha) \leq \frac{1}{\alpha'}(f(b) - f(a)) + \varepsilon.$$

Hence,

$$\lambda^*(\{x \in (a,b) : \ D^+f(x) = \infty\}) = 0.$$

For what we said before, it is enough to show that

$$\lambda^*(\{x \in (a,b) : \ D^+f(x) > D^-f(x)\}) = 0,$$

or

$$\lambda^*(\{x \in (a,b) : \ D^+f(x) > \alpha > \beta > D^-f(x)\}) = 0, \quad \text{for all } \alpha, \beta \in \mathbb{Q}^+.$$

Let us call the above set $E_{\alpha,\beta}$. Fix $\varepsilon > 0$ and consider $O \subseteq (a,b)$ open such that $\lambda(O) \leq \lambda^*(E_{\alpha,\beta}) + \varepsilon$. Let $\mathcal{I}$ the collection of all closed intervals

$$\mathcal{I} = \{[c,d] \subseteq O, \ f(d) - f(c) < \beta(d-c)\}.$$

Since $D^-f(x) < \beta$ on $E_{\alpha,\beta}$, it is easy to see that $\mathcal{I}$ verifies the hypothesis of Vitaly's covering theorem above. Then there exists a finite collection of mutually disjoint closed intervals $[c_1, d_1], \ldots, [c_n, d_n] \in \mathcal{I}$, such that

$$\lambda^*\left(E_{\alpha,\beta} \setminus \bigcup_{i=1}^n [c_i, d_i]\right) < \varepsilon.$$

Now,

$$\begin{aligned}\sum_{i=1}^{n}[f(d_i)-f(c_i)] &< \beta\sum_{i=1}^{n}[d_i-c_i]\\ &\le \beta\lambda(O)\\ &\le \beta[\lambda^*(E_{\alpha,\beta})+\varepsilon].\end{aligned}$$

Moreover, $E_{\alpha,\beta}\cap[c_i,d_i]\subseteq\{x\in[c_i,d_i]:\ D^+f(x)>\alpha\}$, for $i=1,\dots,n$. Thus,

$$\begin{aligned}\lambda^*(E_{\alpha,\beta}\cap[c_i,d_i]) &\le \lambda^*(\{x\in(c_i,d_i):\ D^+f(x)>\alpha\})\\ &\le \frac{1}{\alpha}[f(d_i)-f(c_i)]\end{aligned}$$

Finally, write $E_{\alpha,\beta}=(E_{\alpha,\beta}\setminus\bigcup_{i=1}^{n}[c_i,d_i])\cup(\bigcup_{i=1}^{n}E_{\alpha,\beta}\cap[c_i,d_i])$, to have

$$\begin{aligned}\lambda^*(E_{\alpha,\beta}) &\le \varepsilon+\sum_{i=1}^{n}\lambda(E_{\alpha,\beta}\cap[c_i,d_i])\\ &\le \varepsilon+\frac{1}{\alpha}\sum_{i=1}^{n}[f(d_i)-f(c_i)]\\ &\le \varepsilon+\frac{\beta}{\alpha}(\lambda^*(E_{\alpha,\beta})+\varepsilon).\end{aligned}$$

Since $\beta<\alpha$, we get $\lambda^*(E_{\alpha,\beta})=0$. The proof is completed. □

Corollary 2.4. *Let $f:(a,b)\longrightarrow\mathbb{R}$ be a Lipschitz function. Then f is differentiable at almost every point in (a,b).*

Proof. It is enough to observe that f can be written as

$$f(x)=P_f(x)-(N_f(x)-f(a))$$

where

$$P_f(x)=\sup\sum_{i=1}^{N}[f(x_{i+1})-f(x_i)]^+,$$

$$N_f(x)=\sup\sum_{i=1}^{N}[f(x_{i+1})-f(x_i)]^-,$$

and both supremum are taken over all finite sequences $a=x_1<\cdots<x_{N+1}=x$. Moreover the functions P_f and $N_f-f(a)$ are both increasing. Then one concludes by the previous theorem. □

2.0.1 Rademacher's theorem

Now we are in position to state the main result of this section.

Theorem 2.5 (Rademacher). *Let $\Omega \subseteq \mathbb{R}^n$ be open, and let $f : \Omega \longrightarrow \mathbb{R}^m$ be Lipschitz. Then f is differentiable at almost every point in Ω.*

Proof. By using the extension theorem of the previous chapter, we may assume for simplicity and without loss of generality that $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is Lipschitz; $m = 1$ and $\Omega = \mathbb{R}^n$.

The proof splits into three parts. First the one-dimensional result is used to conclude that the partial derivatives $(\frac{\partial f}{\partial x_i})$ of f exists almost everywhere. This gives us a candidate for the total derivative, namely the (almost everywhere defined) formal gradient

$$\nabla f(x) = (\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}) \tag{2.2}$$

Next, it is shown that all directional derivatives exist almost everywhere, and can be given in terms of the gradient. Finally, by using the fact that there are only "countably many directions" in $\mathbb{R}^n$, the total derivative is shown to exist; it is only in this last step that the Lipschitz condition is seriously used.

We will now carry out these steps. For every $x, v \in \mathbb{R}^n$, with $\|v\| = 1$, the function

$$f_{x,v}(t) := f(x + tv), \quad t \in \mathbb{R}$$

is Lipschitz as a one variable function, then by Lebesgue theorem it is differentiable at almost every $t \in \mathbb{R}$. Let

$$D_v f(x) = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t} \tag{2.3}$$

of course when the limit exists. Let $A_v = \{x : \ D_v f(x) \text{ exists}\}$. Since f is continuous, it can be seen easily that A_v is measurable. Indeed,

$$D^+ f(x) = \lim_{h \to 0^+} \sup_{\substack{0<|t|<h \\ t \text{ rational}}} \frac{f(x + tv) - f(x)}{t}$$

is Borel measurable, similarly for $D^- f$.

For each $y \in v^\perp$, let us decompose $\mathbb{R}^n = \mathbb{R}v \oplus v^\perp$. The intersection of A_v^c with the line parallel to $\mathbb{R} \cdot v$ and passing through y has one dimensional measure zero for what we said before (Theorem 2.4). Then, by Fubini's theorem the set A_v^c has measure zero. In other words, each directional derivative (2.3) exists almost everywhere. In particular, when $v = e_i$, the i-th vector of the basis of $\mathbb{R}^n$, $\frac{\partial f}{\partial x_i}$ exists almost everywhere and then $\nabla f(x)$ exists at almost every $x \in \mathbb{R}^n$.

Claim: for every $v \in \mathbb{R}^n$ we have

$$D_v f(x) = v \cdot \nabla f(x), \text{ for almost every } x \in \mathbb{R}^n. \tag{2.4}$$

Let us fix $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ and let $g \in C_0^\infty(\mathbb{R}^n)$ be a fixed test function. Then

$$\begin{aligned}
\int_{\mathbb{R}^n} D_v f(x) g(x) d\lambda(x) &= \int_{\mathbb{R}^n} \lim_{t\to 0} \frac{f(x+tv)-f(x)}{t} g(x) d\lambda(x) \\
\text{(dominate convergence)} &= \lim_{t\to 0} \int_{\mathbb{R}^n} \frac{f(x+tv)-f(x)}{t} g(x) d\lambda(x) \\
\text{(change of variables)} &= \lim_{t\to 0} \int_{\mathbb{R}^n} -f(x) \frac{g(x)-g(x-tv)}{t} d\lambda(x) \\
\text{(dominate convergence)} &= -\int_{\mathbb{R}^n} f(x) \lim_{t\to 0} \frac{g(x)-g(x-tv)}{t} d\lambda(x) \\
&= -\int_{\mathbb{R}^n} f(x) D_v g(x) d\lambda(x) \\
&= -\sum_{i=1}^n v_i \int_{\mathbb{R}^n} f(x) \frac{\partial g}{\partial x_i}(x) d\lambda(x) \\
\text{(integration by parts)} &= \sum_{i=1}^n v_i \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i}(x) g(x) d\lambda(x) \\
&= \int_{\mathbb{R}^n} v \cdot \nabla f(x) g(x) d\lambda(x)
\end{aligned}$$

Notice that the integration by part step is used on almost every line parallel to the coordinate axes, which is possible by the fact that f is Lipschitz and Fubini theorem.

Now, let $(v_k)_k$ be a dense in the unite sphere of $\mathbb{R}^n$ and define

$$A_k = \{x \in \mathbb{R}^n,\ D_{v_k} f(x) \text{ exists},\ \nabla f(x) \text{ exists, and } D_{v_k} f(x) = v_k \cdot \nabla f(x)\}$$

For what we have proved, we already know that A_k^c has measure zero and so does A^c, where $A = \cap_k A_k$. We claim that f is differentiable at all $x \in A$. More precisely, we claim that for all $x \in A$ we have

$$f(x+w) - f(x) = w \cdot \nabla f(x) + o(\|w\|).$$

Indeed, for $w \neq 0$, let

$$u = \frac{w}{\|w\|},$$

so that

$$w = \|w\|u.$$

Hence,

$$\begin{aligned}
\left|\frac{[f(x+w)-f(x)] - w\cdot\nabla f(x)}{\|w\|}\right| &= \left|\frac{[f(x+\|w\|u)-f(x)] - \|w\|u\cdot\nabla f(x)}{\|w\|}\right| \\
&\leq \left|\frac{[f(x+\|w\|v_i)-f(x)] - \|w\|v_i\cdot\nabla f(x)}{\|w\|}\right| \\
&+ \left|\frac{f(x+\|w\|u)-f(x+\|w\|v_i)}{\|w\|}\right| \\
&+ |(u-v_i)\cdot\nabla f(x)| \\
\text{(since } x\in A \text{ for } \|w\| \text{ small)} &\leq \frac{1}{3}\varepsilon \\
\text{(by Lipschitz condition)} &+ L\|u-v_i\| \\
\text{(by linearity)} &+ C\|u-v_i\| \\
&\leq \varepsilon.
\end{aligned}$$

In the last inequality, since u lives in the unite sphere, we have used the density to pick $i \in \mathbb{N}$.

This completes the proof of Rademacher's. □

There is a generalization of Rademacher's theorem. Let us recall the *pointwise Lipschitz constant* of a function $f : A \longrightarrow \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$:

$$\mathrm{Lip} f(x) := \lim_{y\to x} \sup_{y\in A} \frac{\|f(x)-f(y)\|}{\|x-y\|}$$

Theorem 2.6 (Stepanov)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be open, and let $f : \Omega \longrightarrow \mathbb{R}^m$ be a function. Then f is differentiable almost everywhere in the set*

$$L(f) = \{x \in \Omega : \ \mathit{Lip} f(x) < \infty\}.$$

Proof. We may assume that $m = 1$. Let $\{B_1, B_2, \dots\}$ be the countable collection of all balls contained in Ω such that each B_i has rational center and radius, and that $f|_{B_i}$ is bounded. In particular, this collection covers $L(f)$. Define

$$u_i(x) := \inf\{u(x) : \ u \text{ is } i\text{-Lipschitz with } u \geq f \text{ on } B_i\}$$

and

$$v_i(x) := \sup\{v(x) : \quad v \text{ is } i\text{-Lipschitz with } v \le f \text{ on } B_i\}.$$

By Lemma 1.2, u_i and v_i are i-Lipschitz and $v_i \le f|_{B_i} \le u_i$. It is clear that f is differentiable at every point a, where for some i both u_i and v_i are differentiable with $v_i(a) = u_i(a)$. We claim that almost every point in $L(f)$ is such a point. By Rademacher's theorem, the set

$$Z = \bigcup_{i=1}^{\infty} \{x \in B_i : \text{ either } u_i \text{ or } v_i \text{ is not differentiable at } x\}$$

has measure zero. If $a \in L(f) \setminus Z$, then there is a radius $r > 0$ such that

$$|f(a) - f(x)| \le M\|x - a\|, \quad \forall x \in B(a, r),$$

for some M independent of x. Clearly there is an index $i > M$ such that $a \in B_i \subseteq B(a, r)$. Since $f(a) - i\|x - a\|$ is i-Lipschitz on B_i with $f(a) - i\|x - a\| \le f(x)$ and v_i is the supremum of such functions, then we have

$$f(a) - i\|a - x\| \le v_i(x) \le u_i(x) \le f(a) + i\|a - x\|$$

for $x \in B_i$, which gives the claim. $\square$

A more abstract version of Rademacher's theorem has been done by Cheeger [5], where Lipschitz functions in certain measure metric spaces are considered.

As a consequence of Rademacher's theorem we have that, given a Lipschitz function $f : \Omega \longrightarrow \mathbb{R}^m$, $\Omega \subseteq \mathbb{R}^n$ open, the set of all points at which f is not differentiable has measure zero. For many years it was open the following

Question 2.7. Let $A \subseteq \Omega$ be a Lebesgue zero-measure. It is always possible to find $f : \Omega \longrightarrow \mathbb{R}^m$ such that the set of all points at which f is not differentiable is contained in A?

The question was solved negatively by D. Preiss [10], given a counterexample in case $n = 2$ and $m = 1$. Recently it has been realized that this phenomena is strictly connected by the dimensions m, n. In [2] is has been shown that any Lebesgue zero-measure A in $\mathbb{R}^n$ it is contained in the set of all points at which f is not differentiable, for some Lipschitz function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. On the other hand, D. Preiss, G. Speight [11] have proved that for any $m < n$ the exists a Lebegue zero-measure $A \subseteq \mathbb{R}^n$ such that every Lipschitz function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at some point of A.

Chapter 3

An intermission: Γ-convergence

The Γ-convergence was introduced by De Giorgi in the early 1970s, it has gained an undiscussed role as the most flexible and natural notion of convergence for variational problems, and is now been widely used also outside the field of the Calculus of Variations and of Partial Differential Equations. We report just a quick glance, recommending the book [4].

Definition 3.1. Let (X, d) be a metric space and let $F_h : X \longrightarrow [-\infty, +\infty]$. We say that F_h Γ-converge to $F : X \longrightarrow [-\infty, +\infty]$ if:

(i) For every sequence $(u_h)_h \subseteq X$ convergent to $u \in X$ we have

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

(ii) For all $u \in X$ there exists a sequence $(u_h)_h \subseteq X$ converging to u such that

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

Of course, sequence in (ii) satisfies $\lim_{h\to\infty} F_h(u_h) = F(u)$.

Proposition 3.2. *If (X, d) is separable, any sequence of functionals $F_h : X \longrightarrow [-\infty, +\infty]$ admits a Γ-convergent subsequence.*

Proof. The proof use the compactness of $[-\infty, +\infty]$, the Aleksandrov compactification of $\mathbb{R}$. Let $(U_k)_{k\in\mathbb{N}}$ be a countable basis for the topology of X. By compactness of $[-\infty, +\infty]$, there exists an increasing sequence of integers $(\sigma_j^0)_j$ such that

$$\lim_{j\to\infty} \inf_{y\in U_0} F_{\sigma_j^0}(y)$$

exists, and for all $k \geq 1$ we define $(\sigma_j^k)_j$ as any subsequence of $(\sigma_j^{k-1})_j$ along which the limit

$$\lim_{k\to\infty} \inf_{y\in U_k} F_{\sigma_j^k}(y)$$

exists. By diagonal argument, $j_k = \sigma_k^k$, we have that the limit

$$\lim_{k\to\infty} \inf_{y\in U_l} F_{j_k}(y)$$

exists for all $l \in \mathbb{N}$. For each $x \in X$ let

$$F(x) := \sup_{\substack{l\in\mathbb{N}\\ x\in U_l}} \lim_{k\to\infty} \inf_{y\in U_l} F_{j_k}(y).$$

Then F_{j_k} Γ-converges to F. □

Let V be a normed and let $\omega : [0+\infty[\longrightarrow [0+\infty[$ be a continuous, non-decreasing and positive with $\omega > 0$ on $]0+\infty[$. We say that a 1-homogeneous function $\mathcal{N} : V \longrightarrow [0+\infty[$ is *uniformly convex modulus* ω if

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

for all $u, v \in V$.

Proposition 3.3. *Let $\mathcal{N}_h$ be a sequence of 1-homogeneous uniformly convex functionals with uniformly modulus ω. If $\mathcal{N}_h$ Γ-converges to some functional $\mathcal{N}$, then $\mathcal{N}$ is 1-homogeneous uniformly convex modulus ω.*

Proof. That $\mathcal{N}$ is 1-homogeneous is easy. Let $u, v \in V$ such that $\mathcal{N}(u) = \mathcal{N}(v) = 1$. By (ii), let $(u_h)_h$ and $(v_h)_h$ convergint to u and v respectively, such that $\lim_{h\to\infty} \mathcal{N}_h(u_h) = 1$ and $\lim_{h\to\infty} \mathcal{N}_h(v_h) = 1$. Therefore, if we denote by

$$u_h' = u_h/\mathcal{N}_h(u_h) \text{ and } v_h' = v_h/\mathcal{N}_h(v_h),$$

we get $\mathcal{N}_h(u_h') = \mathcal{N}_h(v_h') = 1$ (use 1-homogenuity). By assumption,

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

Since still $(u_h')_h$ and (v_h') converges to u and v respectively, we get

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

$$\leq 1.$$

Where we have used (i), the monotonicity and continuity of ω and the superadditivity of $\liminf$.

$\square$

Chapter 4

Sobolev Spaces

Let$1 \leq p < \infty$, $\Omega \subseteq \mathbb{R}^n$ and let $u \in L^p(\Omega)$. Then u is said to belong to the Sobolev space $W^{1,p}(\Omega)$ if there exists, for each $i = 1, \ldots, n$, a function $v_i \in L^p(\Omega)$ such that the distributional ith partial derivative of u is determined by v_i via integration; that is,

$$\int_\Omega v_i \varphi \, dx = -\int_\Omega u \frac{\partial \varphi}{\partial x_i} \, dx, \quad \text{for each } \varphi \in C_0^\infty(\Omega).$$

It is easy to see that such a function v_i, if it exists, is unique as an L^p-function, and we set

$$\partial_i u := v_i.$$

The functions $\partial_i u$ a priori have has nothing to do with the partial derivative of u, they are usually called *distributional partial derivatives* of u. The space $W^{1,p}(\Omega)$ became a Banach space equipped with the norm

$$\|u\|_{W^{1,p}} = \|u\|_{L^p} + \|\nabla u\|_{L^p},$$

where $\nabla u = (\partial_1 u, \ldots, \partial_n u)$ is the distributional gradient of u.

We recall the standard approximation procedure. If $\eta \in C_0^\infty(\mathbb{R}^n)$ is a function with

$$\int_{\mathbb{R}^n} \eta(x) \, dx = 1,$$

the one defines the *mollifier function*,

$$u_\varepsilon(x) := u * \eta_\varepsilon(x) = \int_{\mathbb{R}^n} u(y) \eta_\varepsilon(x - y) \, dy,$$

where

$$\eta_\varepsilon(x) := \varepsilon^{-n} \eta(\frac{x}{\varepsilon}),$$

is C^∞-smooth function, and $u_\varepsilon \to u$ in $L^p(\Omega)$, if $u \in L^p(\Omega)$ and $1 \le p < \infty$. Notice that to integrate over $\mathbb{R}^n$ in the definition of u_ε, we set u to be zero outside Ω.

We also have that $u_\varepsilon \to u$ locally uniformly, if u is continuous. Moreover,

$$\partial_i u_\varepsilon = u * \partial_i \eta_\varepsilon = \partial_i u * \eta_\varepsilon,$$

is $u \in W^{1,p}(\Omega)$. It follows that smooth functions are dense in the Sobolev space $W^{1,p}(\Omega)$ for $1 \le p < \infty$.

Essentially, $W^{1,\infty}(\Omega)$ consists of Lipschitz functions.

Theorem 4.1. *The space $W^{1,\infty}(\Omega)$ consists of those bounded functions on Ω that are locally L-Lipschitz (for some L depending on the function). In particular, if Ω is convex, then $W^{1,\infty}(\Omega)$ consists of all bounded Lipschitz functions on Ω.*

Proof. Note that the second claim follows from the first and the fact that a locally L-Lipschitz in a convex set is globally L-Lipschitz.

Assume that $u : \Omega \longrightarrow \mathbb{R}$ is locally L-Lipschitz, for some L. Thus u is Lipschitz on each line that is parallel to a coordinate axis. By using integration by parts on such a line, and then Fubini's theorem (see proof of Theorem 2.5), we find that

$$\int_\Omega \frac{\partial u}{\partial x_i}\eta(x)\,dx = -\int_\Omega u(x)\frac{\partial \eta}{\partial x_i}(x)\,dx, \text{ for each } \eta \in C_0^\infty(\Omega), \text{ and } i = 1, \dots, n.$$

This proves that the almost everywhere existing classical gradient of u is the distributional gradient as well. Moreover, $\|\nabla u\|_\infty \le L$.

Next, assume that $u \in W^{1,\infty}(\Omega)$. Fix a ball B with compact closure in Ω. The convolutions u_ε converge to u pointwise almost everywhere in B. Moreover, we have that

$$\|\nabla u_\varepsilon\|_{\infty,B} \le \|\nabla u\|_\infty < \infty, \quad \text{for } \varepsilon \text{ small enough.}$$

On the other hand, the function u_ε are smooth, so that

$$u_\varepsilon(a) - u_\varepsilon(b) = \int_0^1 \nabla u_\varepsilon(b + t(a-b)) \cdot (a-b)\,dt,$$

and consequently,

$$|u_\varepsilon(a) - u_\varepsilon(b)| \le \|\nabla u\|_\infty \|a - b\|,$$

whenever $a, b \in B$. By letting $\varepsilon \to 0$ we find

$$|u(a) - u(b)| \le \|\nabla u\|_\infty \|a - b\|,$$

for a, b outside a set of measure zero in B. It can be proved that the above inequality holds everywhere in B. $\square$

Although Sobolev functions can exhibit rather singular behavior, there is some regularity beneath the rough surface.

Theorem 4.2. *Let $u \in W^{1,p}(\Omega)$, $1 \le p \le \infty$. Then*

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

where E_i are measurable sets such that $u|_{Ei}$ is i-Lipschitz, and Z has measure zero.

4.0.1 Cheeger-Sobolev space

The fundamental theorem of calculus gives global information about a function after integrating its derivative. There are important several variable versions of this phenomenon; these are the various Sobolev-Poincaré inequalities.

Let u be a smooth real-valued function on $\mathbb{R}^n$. Fix two points $x, y \in \mathbb{R}^n$. We can apply the fundamental theorem of calculus on the line segment $[x, y]$ connecting x and y, and obtain from the chain rule that (as in the previous proof)

$$u(y) - u(x) = \int_0^1 \nabla u(ty + (1-t)x) \cdot (y - x)\ dt,$$

Inserting absolute value, it yields the estimate

$$|u(y) - u(x)| \le \int_{[x,y]} |\nabla u|\ ds \tag{4.1}$$

for the oscillation of the function at the end points in terms of a scalar line integral. The above reasoning holds not only for the straight line segment from x to y but for every rectifiable path γ joining the two points; we have the estimate

$$|u(y) - u(x)| \le \int_\gamma |\nabla u|\ ds \tag{4.2}$$

for each such γ.

Let $\Gamma = \{\gamma\}$ be a family of rectifiable curves in $\mathbb{R}^n$, where a curve in $\mathbb{R}^n$ is a continuous map $\gamma : [a,b] \longrightarrow \mathbb{R}^n$, and a curve is rectifiable if it is (componentwise) of bounded variation. A Borel measurable function $\rho : \mathbb{R}^n \longrightarrow [0,+\infty]$ is said to be an *admissible function*, or density, for Γ if

$$\int_\gamma \rho \, ds \geq 1, \quad \text{for all } \gamma \in \Gamma.$$

The p-modulus of Γ is defined as

$$\operatorname{mod}_p(\Gamma) = \inf \int_{\mathbb{R}^n} \rho^p \, dx \tag{4.3}$$

where the infimum is taken over all admissible functions ρ. A family Γ is said to be *p-exceptional* if $\operatorname{mod}_p(\Gamma) = 0$. If a property of curves holds outside a p-exceptional family, then it is said to hold on *p-almost every curve.*

Proposition 4.3. *A family Γ of curves in $\mathbb{R}^n$ is p-exceptional if and only if there exists a Borel function $\rho : \mathbb{R}^n \longrightarrow [0,\infty]$ such that $\rho \in L^p(\mathbb{R}^n)$ and*

$$\int_\gamma \rho \, ds = \infty, \ \textit{for each locally rectifiable } \gamma \in \Gamma.$$

Proof. Easy. □

Lemma 4.4 (Fuglede's lemma)**.** *If a sequence of Borel functions $(g_k)_k$ converges in $L^p(\mathbb{R}^n)$ to a Borel function g, then there is a subsequence $(g_{k_j})_j$ such that*

$$\int_\gamma g_{k_j} \, ds \to \int_\gamma g \, ds,$$

for p-almost every curve γ in $\mathbb{R}^n$.

Proof. We may assume that $g = 0$ and pick a subsequence $(g_{k_j})_j$ such that $\|g_{k_j}\|_{L^p} \leq 2^{-(p+1)j}$. Let Γ be the family of locally rectifiable curves γ in $\mathbb{R}^n$ for which the statement

$$\lim_{j\to\infty} \int_\gamma g_{k_j} \, ds = 0$$

fails to hold, and let Γ_j be the family of all locally rectifiable curves γ for which

$$\int_\gamma |g_{k_j}| \, ds \geq 2^{-j}.$$

Then on the one hand

$$\Gamma \subseteq \bigcup_{j=k}^{\infty} \Gamma_j$$

for all $k \geq 1$. On the other hand, $2^j |g_{k_j}|$ is admissible for Γ_j, for each j, so that

$$\begin{aligned}\mathrm{Mod}_p(\Gamma_j) &\leq 2^{jp} \int_{\mathbb{R}^n} |g_{k_j}|^p \, dx \\ &\leq 2^{-j}\end{aligned}$$

Consequently, we have that

$$\mathrm{Mod}_p(\Gamma) \leq \sum_{j=k}^{\infty} \mathrm{Mod}_p(\Gamma_j) \leq 2^{-k+1}$$

for each $k \geq 1$, whence $\mathrm{Mod}_p(\Gamma) = 0$ and the lemma is proved. $\square$

The following gives an alternative description of Sobolev functions in $\mathbb{R}^n$.

Theorem 4.5. *A function $u \in L^p(\mathbb{R}^n)$ has a representative in $W^{1,p}(\mathbb{R}^n)$ if and only if there exists a Borel function $\rho \in L^p(\mathbb{R}^n)$ such that the inequality*

$$|u(\gamma(a)) - u(\gamma(b))| \leq \int_\gamma \rho \, ds \tag{4.4}$$

holds for p-almost every curve $\gamma : [a, b] \longrightarrow \mathbb{R}^n$.

Actually, one can realize that distributional gradient $|\nabla u|$ is the (almost everywhere) smallest function ρ that satisfies (4.4) for p-almost every curve γ. It follows that the norm in $W^{1,p}(\mathbb{R}^n)$ can be defined as

$$\|u\|_{W^{1,p}} = \|u\|_{L^p} + \inf \|\rho\|_{L^p}$$

where the infimum is taken over all Borel functions ρ satisfying (4.4) for p-almost every curve γ.

This characterization of Sobolev functions requires no smooth structure of the underlying space. One uses the metric structure for the line integration and measure for the modulus. This lead is followed when Sobolev spaces in arbitrary metric measure spaces.

Let (X, d) be a metric space. Every rectifiable curve $\gamma : [a, b] \longrightarrow X$ has an arc length parametrization

$$\gamma_0 : [0, \mathrm{length}(\gamma)] \longrightarrow X$$

and we define, for any Borel function $\rho : X \longrightarrow [0, +\infty]$,

$$\int_\gamma \rho \, ds := \int_0^{length(\gamma)} \rho(\gamma_0(t)) \, dt.$$

In particular, the definition and basic properties of modulus as before can be transferred over to general metric measure spaces (X, d).

Definition 4.6. A Borel function $\rho : X \longrightarrow [0, +\infty]$ is said to be an *upper gradient* of a function $u : X \longrightarrow \mathbb{R}$ if

$$|u(a) - u(b)| \leq \int_\gamma \rho \, ds \tag{4.5}$$

whenever $a, b \in X$ and γ is a rectifiable curve in X with end points a and b.

As trivial examples, we note that $\rho = \infty$ is an upper gradient of every function.

Definition 4.7. Let $X = (X, d, \mu)$ be a metric measure space and let $1 \leq p < \infty$. A Borel function $\rho : X \longrightarrow [0, \infty]$ is said to be a *p-weak upper gradient* of a function $u : X \longrightarrow \mathbb{R}$ if the inequality in (4.5) holds for p-almost every curve γ in X.

Let us observe that despite that upper gradient can be define on every metric space, for the p-weak upper gradient we need a Borel measure μ on X such that the definition (4.3) makes sense on (X, d). However, it turns out that that every p-integrable p-weak upper gradient of a function can be approximated in $L^p(X, \mu)$ by upper gradients.

Proposition 4.8. *Suppose that a function $u : X \longrightarrow \mathbb{R}$ has a p-integrable p-weak upper gradient. Then there exists a* **minimal p-weak upper gradient** *ρ_u characterized by the following two properties:*

(i) ρ_u is a p-integrable p-weak upper gradient of u;

(ii) if ρ is another p-integrable p-weak upper gradient of u, then $\rho_u \leq \rho$ almost everywhere.

In particular, $\|\rho_u\|_{L^p} = \inf \|\rho\|_{L^p}$ where the infimum is taken over all p-weak upper gradient ρ of u in $L^p(X, \mu)$.

Proof. First one proves (we omit the details) that p-integrable p-weak upper gradients form a lattice in the sense that if τ and σ are two p-integrable p-weak upper gradients of a given function, then so is $\min\{\tau,\ \sigma\}$. Consequently, any sequence $(\rho_i)_i$ of p-weak upper gradients of a given function u satisfying

$$\lim_i \|\rho_i\|_{L^p} = \inf \|\rho\|_{L^p}$$

where the infimum is taken over all p-weak upper gradients ρ of u, can be chosen to be pointwise decreasing:

$$\rho_1 \geq \rho_2 \geq \dots .$$

Clearly, then, $(\rho_i)_i$ converges in $L^p(X,\mu)$ to a Borel function ρ_u whose L^p-norm assumes the above infimum. By Fuglede's Lemma 4.4, ρ_u is a p-weak upper gradient of u. It is also clear from the lattice property of upper gradients that ρ_u is minimal as asserted. □

Definition 4.9. Consider the vector space $\widetilde{W}^{1,p}(X)$ consisting of all functions $u : X \longrightarrow \mathbb{R}$ such that u is in $L^p(X)$ and there exists an upper gradientρ of u in $L^p(X)$. Then we can define a seminorm in $\widetilde{W}^{1,p}(X)$ by

$$\|u\|_{W^{1,p}} := \|u\|_{L^p} + \|\rho_u\|_{L^p}$$

where ρ_u is the minimal p-weak upper gradient of u. The *Cheeger-Sobolev space* $W^{1,p}(X)$ is the space

$$W^{1,p}(X) := \widetilde{W}^{1,p}(X)/\sim$$

where

$$u \sim v \text{ if and only if } \|u - v\|_{W^{1,p}} = 0.$$

Theorem 4.10. *$W^{1,p}(X)$ is a Banach space.*

Proof. See [13, Theorem 3.7]. □

For a function $u \in L^p(X,\mu)$ let us define the extended real valued Borel function,

$$\mathrm{Lip}(u)(x) := \lim_{r\to 0} \sup_{y\in B(x,r)} \frac{|u(y) - u(x)|}{d(x,y)},$$

where we put $\mathrm{Lip}(u)(x) := 0$ if x is isolated point.

Proposition 4.11. *If $u : X \longrightarrow \mathbb{R}$ is Lipschitz, then Lip(u) is an upper gradient for u.*

Proof. Since the restriction of u to any rectifiable curve, γ, is Lipschitz, by Rademacher's theorem 2.5, it follows that $u(\gamma(s))$ is differentiable for almost every s. Thus $|u'(\gamma(s))|$ is an upper gradient (see the argument in (4.2)). Therefore, it is enough to show that

$$|u'(\gamma(s))| \leq \mathrm{Lip}(u)(\gamma(s))$$

for those s such that $|u'(\gamma(s))|$ exists.

Fix a value $\bar{s}$. We can assume that $d(\gamma(s),\gamma(\bar{s})) > 0$ for s sufficiently close to $\bar{s}$; since otherwise, $u'(\gamma(\bar{s})) = 0$. Then, by continuity, for all sufficiently small r, there exists some smallest $s(r)$, with $d(\gamma(s(r)),\gamma(\bar{s}))=r\le|s(r)-\bar{s}|$. In addition, $s(r)\to\bar{s}$ as $r\to 0$. We have

$$\frac{|u(\gamma(s(r)))-u(\gamma(\bar{s}))|}{|s(r)-\bar{s}|}\le\sup_{y\in\partial B(\gamma(\bar{s}),r)}\frac{|u(y)-u(\gamma(\bar{s}))|}{r}.$$

Since $u'(\gamma(\bar{s}))$ exists, we get

$$\begin{aligned}|u'(\gamma(\bar{s}))|&=\liminf_{r\to 0}\frac{|u(\gamma(s(r)))-u(\gamma(\bar{s}))|}{|s(r)-\bar{s}|}\\&\le\liminf_{r\to 0}\sup_{y\in\partial B(\gamma(\bar{s}),r)}\frac{|u(y)-u(\gamma(\bar{s}))|}{r}\\&=\mathrm{Lip}(u)(\gamma(\bar{s})).\end{aligned}$$

□

Definition 4.12. We say that $g\in L^p(X,\mu)^+$ is a p-relaxed slope of $u\in L^p(X,\mu)$ if there exists $\widetilde{g}\in L^p(X,\mu)^+$ and a sequence of Lipschitz functions $(u_n)_n$ such that

(*i*) $u_n\to u$ in $L^p(X,\mu)$ and $\mathrm{Lip}(u_n)$ weakly converges to $\widetilde{g}$ in $L^p(X,\mu)$;

(*ii*) $\widetilde{g}\le g$ μ-a.e. in X.

It is easy to see that the set of all p-relaxed slope of u is a closed convex set in $L^p(X,\mu)$. By uniformly convexity of $L^p(X,\mu)$, among all p-relaxed slope of u there is one with minimal L^p-norm. Such an element will be denoted by $|\nabla u|_{*,p}$.

By Proposition 4.11, since $\mathrm{Lip}(u_n)$ are in particular p-weak upper gradient, and weak convergence is stable under p-weak upper gradient, it follows that

$$\rho_u\le|\nabla u|_{*,p},\quad \mu\text{-a.e. in } X,$$

for all $u\in L^p(X,\mu)$.

It is an exercise to prove that the vector space $W^{1,p}(X,\mu)$ endowed with the new norm $\|u\|_{L^p}+\||\nabla u|_{*,p}\|_{L^p}$ became a Banach space, and the identity operator

$$\mathrm{id}:(W^{1,p}(X,\mu),\|\cdot\|_{L^p}+\|\,|\nabla\cdot|_{*,p}\|_{L^p})\longrightarrow(W^{1,p}(X,\mu),\|\cdot\|_{L^p}+\|\rho_{\cdot}\|_{L^p})$$

is a surjective bounded linear operator. As a consequence of the open mapping theorem, it follows that the two norms above in $W^{1,p}(X,\mu)$ are equivalents.

Remark 4.13. Actually, in [1, Theorem 6.1] the authors proved that the two norms above are equals.

4.0.2 Reflexivity of Cheeger-Sobolev space

Though the Cheeger-Sobolev space is always a Banach space, the deep difference with the classical one is that the space is not in general reflexive as the following shows.

Example 4.14. Let $(a_n)_n$ be a sequence in $]0,1[$ which converges to zero, and let

$$X = \prod_{n\in\mathbb{N}} [0,\ a_n],$$

endowed with the norm $\|(x_n)_n\|_\infty = \max_n x_n$. There is a natural product (probability) measure μ on X, that is the product of normalized Lebesgue measures on each factor $[0, a_n]$. Now consider a sequence $(f_n)_n$ of functions $f_n : X \longrightarrow \mathbb{R}$ that are the projections on the factors; i.e., $f_n(x) = x_n$ for $x = (x_n)_n \in X$. Because each function f_n is 1-Lipschitz on X, and bounded by 1, the sequence $(f_n)_n$ is a bounded sequence in $W^{1,p}(X)$. Assuming that the Cheeger-Sobolev space in question is reflexive, a weakly convergent subsequence could be found with weak limit f. By passing to Mazur's lemma we would further find a sequence $(g_m)_m$ consisting of convex combinations of the functions f_n,

$$g_m = \lambda_{m,1} f_{m_1} + \cdots + \lambda_{m,k_m} f_{m_{k_m}}$$

where $0 \leq \lambda_{m,j} \leq 1$ and $\lambda_{m,1} + \cdots + \lambda_{m,k_m} = 1$, with $m_{k_i} \to \infty$ as $m \to \infty$, such that $g_m \to f$ strongly in $W^{1,p}(X)$. It follows that $f = 0$. Moreover, from the norm structure it is easy to check that the minimal upper gradient of each g_m is identically $\lambda_{m,1} + \cdots + \lambda_{m,k_m} = 1$, which contradicts the fact that $g_m \to 0$ in $W^{1,p}(X)$.

Now, we would like to prove the marvelous result of Ambrosio, Colombo, Di Marino [1], regarding the reflexivity of the Cheeger-Sobolev space $W^{1,p}(X)$ in case X is a doubling separable metric and μ is a Borel measure on X finite on bounded sets. To do so, we take advance on the doubling property to create nice partition on X.

Lemma 4.15. *For every $\delta > 0$ there exist $\ell_\delta \in \mathbb{N} \cup \{\infty\}$ and pairs set-point (A_i^δ, z_i^δ), $0 \leq i < \ell_\delta$, where $A_i^\delta \subseteq X$ are Borel sets and $z_i^\delta \in X$, satisfying:*

(i) the sets $\{A_i^\delta :\ 0 \leq i < \ell_\delta\}$ forms a partition of X;

(ii) $d(z_i^\delta, z_j^\delta) > \delta$, for $i \neq j$;

(iii) $B(z_i^\delta, \frac{\delta}{3}) \subseteq A_i^\delta \subseteq B(z_i^\delta, \frac{5}{4}\delta)$;

(iv) for each $0 \leq i < \ell_\delta$ *then the cardinality of*

$$s(i) = \{j \in \mathbb{N} : \; d(A_i^\delta, A_j^\delta) < \delta\}$$

is less than λ^3 *(doubling condition);*

(v) $x \in A_k^\delta$ *if and only if* $k = \min I_x^\delta$ *where*

$$I_x^\delta = \{i \in \mathbb{N} : \; d(x, z_i^\delta) \leq d(x, z_j^\delta) + \frac{\delta}{8}, \; \forall j \in \mathbb{N}\};$$

we are minimizing the quantity $d(x, z_i^\delta)$ *and among those indeces* i *who are minimizing up to* $\frac{\delta}{8}$ *we take the least one* i_x*.*

Proof. Let $\{x_k\}_{k\in\mathbb{N}}$ be a dense in X. Let us choose $z_0^\delta = x_0$. Recursively consider $B_i = X \setminus \bigcup_{j<i} \overline{B}(z_j^\delta, \delta)$ and choose $k_i = \min\{k \in \mathbb{N} : \; x_k \in B_i\}$. Then choose $z_i^\delta = x_{k_i}$. Using density and by construction, it easily follows that

$$X = \bigcup_i B(z_i^\delta, \delta + \frac{\delta}{8}).$$

Then define

$$A_0^\delta := \{x \in X : \; d(x, z_0^\delta) \leq d(x, z_j^\delta) + \frac{\delta}{8}, \; \forall j > 0\}$$
$$A_i^\delta := \{x \in X : \; d(x, z_i^\delta) \leq d(x, z_j^\delta) + \frac{\delta}{8}, \; \forall j > i\}$$

It is easy to see that A_i^δ and the points z_i^δ satisfy all the properties required. Let us note that (iv) follows form the fact that we can cover $B(z_i^\delta, 4\delta)$ with λ^3 balls with radius $\frac{\delta}{2}$ but each of them can contain only one of the z_j^δ's (by (ii)). □

In order to define our discrete gradients we give more terminology. We say that A_i^δ is a neighbor of A_j^δ , and we denote by $A_i^\delta \sim A_j^\delta$, if their distance is less than δ. In particular $A_i^\delta \sim A_j^\delta$ implies that $d(z_i^\delta, z_j^\delta) \leq 4\delta$: indeed, if $\widetilde{z}_i^\delta \in A_i^\delta$ and $\widetilde{z}_j^\delta \in A_j^\delta$ satisfy $d(\widetilde{z}_i^\delta, \widetilde{z}_j^\delta) \leq \delta'$, we have

$$\begin{aligned} d(z_i^\delta, z_j^\delta) &\leq d(z_i^\delta, \widetilde{z}_i^\delta) + d(\widetilde{z}_i^\delta, \widetilde{z}_j^\delta) + d(\widetilde{z}_j^\delta, z_j^\delta) \\ &\leq \frac{10}{4}\delta + \delta' \\ \text{(letting } \delta' \to \delta) \;\; &\leq \frac{14}{4}\delta \end{aligned}$$

$$\leq 4\delta.$$

Now we fix $\delta \in]0,1[$ and we consider a partition $\{A_i^\delta : \ 0 \leq i < \ell_\delta\}$ of X on scale δ. For every $u \in L^p(X,\mu)$ we define the average $u_{\delta,i}$ of u in each cell of the partition by $⨍_{A_i^\delta} u \, d\mu$. We denote by $P\mathcal{C}_\delta(X)$, which depends on the chosen decomposition as well, the set of functions $u \in L^p(X,\mu)$ constant on each cell of the partition at scale δ, namely

$$u(x) = \text{constant}, \quad \text{for } \mu\text{-a.e. } x \in A_i^\delta.$$

Then we can define the linear projection

$$P_\delta : L^p(X,\mu) \longrightarrow P\mathcal{C}_\delta(X)$$

by

$$P_\delta(u) = u_{\delta,i}, \quad \text{for every } x \in A_i^\delta.$$

Lemma 4.16. *P_δ are contractions in $L^p(X,\mu)$ and $P_\delta(u) \stackrel{\delta\to 0}{\longrightarrow} u$ in $L^p(X,\mu)$ for all $u \in L^p(X,\mu)$.*

Proof. The contractivity of P_δ is a simple consequence of Jensen's inequality.

To show that $P_\delta(u) \stackrel{\delta\to 0}{\longrightarrow} u$ in $L^p(X,\mu)$, it is enough to get that convergence in a dense subset of $L^p(X,\mu)$, in the set of bounded continuous functions with bounded support. Since X is doubling, if u is any of such a function, then u is uniformly continuous and then $P_\delta(u) \stackrel{\delta\to 0}{\longrightarrow} u$ pointwise. The conclusion follows by the dominated convergence theorem. □

For each $\delta \in]0,1[$ and $i \in \mathbb{N}$, let

$$|\mathcal{D}_\delta u|^p(x) := \frac{1}{\delta^p} \sum_{j:\ A_i^\delta \sim A_j^\delta} |u_{\delta,i} - u_{\delta,j}|^p, \quad \forall x \in A_i^\delta.$$

Let us also define $\mathcal{F}_{\delta,p} : L^p(X,\mu) \longrightarrow [0,\infty]$ by

$$\mathcal{F}_{\delta,p}(u) := \int_X |\mathcal{D}_\delta u|^p(x) \ d\mu(x).$$

Lemma 4.17. *For each curve $\gamma : [a,b] \longrightarrow X$, with length$(\gamma) \geq \delta/2$, we have*

$$|P_\delta(u)(\gamma(b)) - P_\delta(u)(\gamma(a))| \leq 4 \int_\gamma |\mathcal{D}_\delta u|(t) \ dt.$$

Proof. WLOG we can assume that $\delta/2 \leq$ length$(\gamma) \leq \delta$; otherwise slice $[a,b]$ in subintervals in which we have that estimates and then use the triangle inequality. Since, for each $t \in [a,b]$, $d(\gamma(t),\gamma(a)) < \delta$ and $d(\gamma(t),\gamma(b)) < \delta$, so that the cell relative to $\gamma(a)$ and $\gamma(b)$ are both neighbors of the one relative to $\gamma(t)$. So, by definition

$$\begin{aligned}|\mathcal{D}_\delta u|^p(\gamma(t)) &\geq \frac{1}{\delta^p}[|P_\delta(u)(\gamma(b)) - P_\delta(u)(\gamma(t))|^p + |P_\delta(u)(\gamma(t)) - P_\delta(u)(\gamma(a))|^p] \\ &\geq \frac{1}{2^{p-1}\delta^p}|P_\delta(u)(\gamma(b)) - P_\delta(u)(\gamma(a))|^p.\end{aligned}$$

Then

$$\begin{aligned}\int_\gamma |\mathcal{D}_\delta u|(t)\ dt &\geq \frac{1}{2^{1-\frac{1}{p}}\delta}|P_\delta(u)(\gamma(b)) - P_\delta(u)(\gamma(a))| \cdot \text{length}(\gamma) \\ &\geq \frac{1}{4}|P_\delta(u)(\gamma(b)) - P_\delta(u)(\gamma(a))|\end{aligned}$$

□

Lemma 4.18. *Assume $(\varepsilon_n)_n$ be a null sequence of positive reals, $(f_n)_n \subseteq L^p(X,\mu)$ and $(g_n)_n \subseteq L^p(X,\mu)^+$ such that*

$$|f_n(a) - f_n(b)| \leq \int_\gamma g_n\ ds$$

for all $a,b \in X$ and for all p-almost every curve γ in X with end points a and b such that length$(\gamma) \geq \varepsilon_n$.

Assume furthermore that $f_n(x) \to f(x)$ for μ-a.e. $x \in X$ and $(g_n)_n$ converges weakly in $L^p(X,\mu)$ to some $g \in L^p(X,\mu)$. Then g is a p-weak upper gradient of f.

Proof. Combine Mazur's theorem and the argument seen in Fuglede's Lemma 4.4. □

Since $L^p(X,\mu)$ is separable, by Proposition 3.2, unless to pass through a subsequence, we can assume that $\mathcal{F}_{\delta,q}$ have Γ-limit points as $\delta \to 0$; i.e., let

$$\mathcal{F}_p := \lim_{k\to\infty} \mathcal{F}_{\delta_k,p}$$

for some infinitesimal sequence $(\delta_k)_k$, where the limit is computed with respect to L^p distance.

Theorem 4.19. *(a) There exists $\eta = \eta(p,\lambda)$ (keep in mind λ denotes the doubling constant), such that*

$$\frac{1}{\eta}\|\rho_u\|_{L^p} \leq \mathcal{F}_p(u) \leq \|\rho_u\|_{L^p}.$$

(b) The norm on $W^{1,p}(X,\mu)$

$$[\|u\|^p_{L^p} + \mathcal{F}_p(u)]^{\frac{1}{p}}$$

is uniformly convex.

Proof. For what we said in the paragraph preceding Remark 4.13, it is enough to prove the inequality involving $|\nabla u|_{*,p}$.

Let $u : X \longrightarrow \mathbb{R}$ be a Lipschitz function with bounded support. Let $i, j \in \{1, \dots, \ell_\delta\}$ such that A_i^δ and A_j^δ are neighbors. For every $x \in A_i^\delta$ and $y \in A_j^\delta$ we have $d(x,y) \leq (10/4+10/4+1)\delta = 6\delta$ and that $y \in B(z_i^\delta, 19/4\delta) \subseteq B(z_i^\delta, 5\delta)$. Hence

$$\begin{aligned} \frac{|u_{\delta,i} - u_{\delta,j}|}{\delta} &\leq \frac{1}{\delta\mu(A_i^\delta)\mu(A_j^\delta)} \int_{A_i^\delta \times A_j^\delta} |u(x) - u(y)| \, d\mu(x) d\mu(y) \\ &\leq 6\mathrm{Lip}(u|_{B(z_i^\delta, 5\delta)}) \end{aligned}$$

Thanks to the fact that the number of neighbors of A_i^δ does not exceed λ^3, we obtain

$$|\mathcal{D}_\delta u|^p(x) \leq 6^p \lambda^3 \left(\mathrm{Lip}(u|_{B(z_i^\delta, 5\delta)})\right)^p.$$

Integrating on X we obtain that

$$\mathcal{F}_p(u) \leq \liminf_{k\to\infty} \mathcal{F}_{p,\delta_k}(u) \leq 6^p\lambda^3 \int_X |\nabla u|^p_{*,p} \, d\mu.$$

Hence we have the right hand side of (a). For the left hand side, we can assume that $\mathcal{F}_p(u)$ is finite. Therefore, the sequence $(|\mathcal{D}_{\delta_k} u_k|)_k$ is bounded in $L^p(X,\mu)$. By weak compactness, unless to pass through a subsequence,

$$|\mathcal{D}_{\delta_k} u_k| \overset{k\to\infty}{\longrightarrow} g \text{ weakly in } L^p(X,\mu).$$

By the lower semicontinuity of the p-norm with respect to the weak convergence, we have that

$$\int_X g^p \, d\mu \leq \liminf_{k\to\infty} \int_X |\mathcal{D}_{\delta_k} u_k|^p \, d\mu = \lim_{k\to\infty} \mathcal{F}_{\delta_k,p}(u_k).$$

Now, apply Lemma 4.18 to $\mathcal{P}_{\delta_k}(u_k)$ which converges to u in $L^p(X,\mu)$ thanks Lemma 4.16, and to the functions $4|\mathcal{D}_{\delta_k} u_k|$ which are p-weak upper gradient of $\mathcal{P}_{\delta_k}(u_k)$ (up to scale $\delta_k/2$) thanks Lemma 4.17.

We obtain that $4g$ is a p-weak upper gradient of u, hence $g \geq \rho_u/4$ μ-a.e. in X. Therefore,

$$\frac{1}{4^p} \int_X \rho_u^p \, d\mu \leq \int_X g^p \, d\mu \leq \mathcal{F}_p(u).$$

We leave (b) as esercize. □

As consequence of the previous Theorem and Proposition 3.3, we get the main result of this section.

Corollary 4.20. *Let* (X, d) *be a separable doubling metric space supporting a Borel measure* μ *finite on bounded sets. Then the Cheeger-Sobolev space* $W^{1,p}(X, \mu)$ *is reflexive.*

Bibliography

[1] L. Ambrosio, M. Colombo, S. Di Marino; Sobolev spaces in metric measure spaces: reflexivity and lower semicontinuity of slope. Variational methods for evolving objects, *Adv. Stud. Pure Math.* **67** (2015), 1–58.

[2] G. Alberti, M. Csornyei, D.Preiss; Structure of null sets, differentiability of Lipschitz functions, and other problems. Preprint (2016).

[3] P. Assouad; Plongements lipschitziens dans $\mathbb{R}^n$. *Bull. Soc. Math. France* **111** (1983), no. 4, 429–448.

[4] A. Braides; Γ-convergence for beginners. *Oxford Lecture Series in Mathematics and its Applications*, **22** (2002).

[5] J. Cheeger; Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.* **9** (1999), no. 3, 428–517.

[6] M. Gromov; Monotonicity of the volume of intersection of balls. *Geometrical aspects of functional analysis (1986/86)* vol. **1267**, *Lecture Notes in Math.*, Springer, Berlin (1987), 1–4.

[7] M. D. Kirszbraun; Über die zusammenziehende und Lipschitzsche Transformationen. *Fund. Math.* **22** (1934), 77–108.

[8] J.R. Lee, A. Naor; Extending Lipschitz functions via random metric partitions. *Inventiones Mathematicae*, **160** (2005), 59–95.

[9] J. Lindenstrauss; On nonlinear projections in Banach spaces. *Mich. Math. J.* **11** (1964), 263–287.

[10] D. Preiss; Differentiability of Lipschitz functions on Banach spaces. *J. Funct. Anal.*, (2)**91** (1990), 312–345.

[11] D. Preiss, G. Speight ; Differentiability of Lipschitz functions in Lebesgue null sets. Invent. Math., (2)**199** (2015), 517–559.

[12] S. Semmes; On the nonexistence of bi-Lipschitz parameterizations and geometric problems about A^∞-weights. *Rev. Mat. Iberoamericana*, **12**(2) (1997), 337–410.

[13] N. Shanmugalingam; Newtonian spaces: an extension of Sobolev spaces to metric measure spaces. *Rev. Mat. Iberoamericana* **16** (2000), no. 2, 243–279.

Contents