UNIVERSITY OF PALERMO

PHD JOINT PROGRAM:

UNIVERSITY OF CATANIA - UNIVERSITY OF MESSINA

XXXIV CYCLE

DOCTORAL THESIS

Topics in calculus and geometry on metric spaces

Author:
VINCENZO PALMISANO

Supervisor:
Prof. BIAGIO RICCERI

Co-Supervisor:
Prof. ANDREA O. CARUSO

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy
in
Mathematics and Computational Sciences

May 9, 2022

Signed:

Date:

"Un bel problema, anche se non lo risolvi, ti fa compagnia se ci pensi ogni tanto."

Ennio De Giorgi

UNIVERSITY OF PALERMO

Abstract

Department of Mathematics and Computer Sciences

Doctor of Philosophy

Topics in calculus and geometry on metric spaces

by VINCENZO PALMISANO

In this thesis we present an overview of some important known facts related to topology, geometry and calculus on metric spaces. We discuss the well known problem of the existence of a lipschitz equivalent metric to a given quasi-ultrametric, revisiting known results and counterexamples and providing some new theorems, in an unified approach. Also, in the general setting of a quasi-metric doubling space, suitable partition of unity lemmas allows us to obtain, in step two Carnot groups, the well known Whitney's extension theorem for a given real function of class C^m defined on a closed subset of the whole space: this result relies on relevant properties of the symmetrized Taylor's polynomial recently introduced in this setting. Finally, some first interesting investigations on Menger convexity in the setting of a general metric spaces concludes this work.

Acknowledgements

I would like to thank all people who made this thesis possible. I'm grateful to Prof. Biagio Ricceri for his generous availability, and for constant willingness to talk about mathematical questions. Also, my thanks goes to Prof. Andrea Caruso for a few nice moments through mathematics. And, last but not least, a special thanks goes to my wife Cristina, who has always supported my choices.

Contents

Abstract **v**

Acknowledgements **vii**

Introduction **1**

1 Distance spaces and the metrization problem **5**
- 1.1 Distances and triangular type inequalities 5
- 1.2 The metrization problem for quasi-ultrametric spaces 19
- 1.3 Equivalent metrics on quasi-metric spaces 36

2 A brief glimpse into doubling spaces and applications **43**
- 2.1 Homogeneity and geometric doubling properties 43
- 2.2 Partition of unity 53
- 2.3 Lipschitz and hölder continuous functions 57
- 2.4 Lipschitz extension theorems 61
- 2.5 Extension of lipschitz functions on space of homogeneous type 72

3 Selected topics on calculus on Carnot groups **81**
- 3.1 Preliminaries and notations 81
- 3.2 Whitney's extension theorem 96

A First insights into Menger convexity **101**
- A.1 Betweenness in metric spaces 101
  - A.1.1 Metric segments and convex sets 103
- A.2 Menger and Takahashi's convexity 107
- A.3 Convex function 114
  - A.3.1 Metric convex hull 120
  - A.3.2 Metric convex function and Jensen's inequality 122

Bibliography **127**

Dedicated to my father

Introduction

Chapter 1 is dedicated to the problem of the existence of a metrically equivalent metric with respect to a given quasi-ultrametric. The idea of measuring distance between points on abstract sets arises naturally as a generalization of the concept of geometric distance; clearly every distance must satisfy some characterizing requirements. The notion of space with a distance, or *distance space*, appears in 1906, appropriately exposed in relation to some sets of functions [1], and in the first half of twentieth century has gradually found full acceptance. We stress immediately the in our study, the terms distance and metric have a different meaning, due to the level of generality and completeness of the results. The problem of the existence of a suitable metric with a prescribed topology is well known, and the techniques employed are typical of the topological setting; instead, rather geometric are the techniques known in literature for the problem of the existence o a geometrically equivalent metric. Indeed, in 1937, generalizing the proof of E. W. Chittenden, A. H. Frink exhibits for the first time an equivalent metric to a given – nowadays called – quasi-ultrametric (with constant $K = 2$); more precisely in [2] the existence of a topological equivalent metric is proved, then in [3], by the introduction of the *chain approach techniques* jointly with a candidate metric d_ρ , the existence of a bi-lipschitz equivalent metric is proved, arguing by contradiction; after that, several proof of the existence of such a metric, and related facts, do appear in literature (see for instance [4, 5]): in this chapter we organize main key points and provide simple and new proofs in a unified setting by using standard techniques based both on the the validity of a relaxed polygonal inequality (see, for instance, [6])

$$\rho(x, y) \leq c \sum_{i=1}^{n} \rho(x_{i-1}, x_i), \quad x_0 = x, x_1, ..., x_n = y \in X,$$

and topological techniques based on the existence of suitable uniformities. For the sake of completeness, known results related to existence of a bi-lipschitz metric for a suitable power of a given quasi-metric conclude the chapter.

Chapter 2 is a technical one. Space of homogeneous type and doubling spaces are compared: the equivalence of these facts is known, we present several equivalent facts in the setting of a general quasi-ultrametric spaces. Note that spaces of homogeneous type appear in [7], more recent is the notion of geometrically doubling spaces; spaces of homogeneous type are metric spaces endowed with doubling measure, as, for instance, Euclidean spaces and the group $\mathbb{H}^1$. Moreover, in the setting of doubling quasi-metric spaces we provide autonomous proofs of global and local partition of unity lemmas, employed later in Chapter 3. A standard application of these lemmas provide extension theorem for normed valued Lipschitz mappings.

Chapter 3 is dedicated to the generalization of Whitney extension theorem in the setting of step two Carnot groups: actually, for the sake of clarity we provide a quite detailed proof in the setting of the Heisenberg group $\mathbb{H}^1$, avoiding more technicalities. The history of Carnot groups goes back to E. M. Stein, in '82, when jointly with G. B. Folland began calculus in some connected and simply connected nilpotent Lie groups whose algebra is endowed with a family of dilations δ_λ; after more than twenty years a particular class of these groups, nowadays called *Carnot groups*, are of relevant interest as a natural setting for real analysis, calculus and geometric measure theory. These algebraic structures arise as the natural tangent structures at regular point of a sub-Riemannian manifold, in analogy with the Euclidean vector spaces and the Riemannian manifolds. A sub-Riemannian manifold (M, D, g) is a Riemannian manifold (M, g) where a given distribution D of m-dimensional vector spaces of the tangent bundle has been fixed. An absolutely continuous curve $\gamma : \mathbb{R} \supset I \to M$ joining two points p and q in M such that $\dot{\gamma}(t) \in D_{\gamma(t)}$ for a.e. any $t \in I$ is called horizontal; through these curves we can define a metric $d(p, q)$, the *C-C* or *Carnot-Carathéodory metric*, as the infimum of all the lengths of all such a curves joining p e q. The induced topology is the original one but d and d_g, the Riemannian metric, are non equivalent in general. Hörmander condition requires that, given a finite family $X_1, X_2, \dots, X_m$ of smooth vector fields on a n-dimensional smooth manifold M, there exists a positive integer r such that we can find at least n linearly independent vectors at each point p of M among all the vector fields of the kind X_i, $[X_{i_1}, X_{i_2}]$, $[X_{i_1}, [X_{i_2}, X_{i_3}]], \dots, [X_{i_1}, [X_{i_2}, [X_{i_3}, \cdots [X_{i_{r-1}}, X_{i_r}] \cdots]]]$, calculated at this point. So, in the setting of step two Carnot groups we exhibit a quite direct proof of the well known Whitney's extension theorem. Let us briefly recall some historical facts about. In 1925 P. Urysohn proved how a given continuous function defined on a closed subset F of a normal topological space can be extended to a continuous function defined on the whole space [8]; indeed this theorem, whose proof is quite simple when F is a closed subset of $\mathbb{R}$, was extended first by H. Lebesgue in 1907 to the plane and after by H. Tietze in 1915 to a metric space [9, 10]). Later on H. Whitney raised and solved in 1934 the problem of finding a differentiable or even analytic extension of a given function of the same regularity defined on a closed subset $F \subseteq \mathbb{R}^n$ [11]; finally, in 2005, C. Fefferman has given a deep generalization and characterization of this extension theorem for an extension of class $C^{m-1,1}(\mathbb{R}^n)$ [12]. Whitney's proof relies on two key tools: a particular open covering for open sets in $\mathbb{R}^n$ and an extension operator built up with the help of a suitable partition of unity subordinated to such a covering. Relatively to the first tool, in the original paper, rather than a type covering argument we should speak about a decomposition into cubes of open sets in $\mathbb{R}^n$ [13]; a simplified argument avoiding this dyadic decomposition appears in [14], where the same extension operator is constructed through a partition of unity subordinated to a Vitali type covering of a given open set. We provide a proof of Whitney's extension theorem for functions of class $C^m_{\mathbb{H}^1}(F)$, with $F \subseteq \mathbb{H}^1$ closed. The proof is based on Whitney's original one, following the scheme of H. Federer. From an analysis of the original proof, it is evident that various steps are made possible merely thanks to the properties of the Taylor polynomial in the ordinary Euclidean case. Unfortunately, the well-posedness of the original definition of the Taylor polynomial given by Folland and Stein, relies on the existence of a linear isomorphism between the vector space of polynomial and and the vector space of left invariant vector fields, but considerable difficulties of computational kinds occurs if we try to write it explicitly. The Taylor polynomial introduced in [15] and recalled in the first section helps with this problem, thanks to its properties that make it similar to analogous one in the Euclidean case; furthermore, the use of the polynomial is also necessary in the proof of the extension theorem because the condition for the extension is characterized in terms of it, jointly with a suitable family of functions, i.e. a jet associated to the set F, which assigns the coefficients of the polynomial to each point of space. It comes out that the proof of the Whitney extension theorem, in spite of the highly non linear structure of Carnot groups, reduce, in spirit, to a Euclidean one.

Finally an appendix is dedicated to the study of the notion of convexity. An accurate bibliographic research has shown some properties of sets and candidate functions to be *convex in a metric sense*, although these convex set and functions do not seem to have been the subject of particular study in the general setting of a metric space, possibly with supplementary hypothesis (geometric properties of convex sets, differentiability properties of functions, Caratheodory theorem (which is false in $\mathbb{H}^1$) or Helly theorems); actually such theorems require suitable preliminary definition and properties that, in a general context of a metric space, do not appear obvious at all. So, in this appendix, first basic properties of convex sets and functions (characterization of convex hull, Jensen inequality, basic extensions theorems, etc), according to K. Menger [16] are studied, together with possibly supplementary hypotheses on the metric space.

Chapter 1

Distance spaces and the metrization problem

1.1 Distances and triangular type inequalities

In this section we recall some fundamental statements about distances and distance spaces, in particular distinguishing the different conditions that may be required for a distance in place of the usual triangular inequality. Most of the things contained here can be found on every book about the argument [17, 18, 19].

Definition 1.1.1 (Semi-distance). *A semi-distance ρ on a non-empty set X is a symmetric application $\rho : X \times X \to [0, +\infty[$ such that*

$$x = y \Rightarrow \rho(x, y) = 0.$$

A semi-distance space (X, ρ) is a non-empty set X equipped with a semi-distance ρ.

Example 1.1.1. *Let X be a no empty set. Then the application $\rho : X \times X \to [0, +\infty[$ given by $\rho(x, y) = 0$ for all $x, y \in X$ is trivially a semi-distance on X.*

Example 1.1.2. *Let V be a linear space and p a seminorm on V. Then by definition of seminorm the application $\rho : V \times V \to [0, +\infty[$ given by $\rho(\mathbf{v}, \mathbf{w}) = p(\mathbf{v} - \mathbf{w})$ for all $\mathbf{v}, \mathbf{w} \in V$ is trivially a semi-distance on V.*

Example 1.1.3. *Let X be a no empty set, $x_0 \in X$ a fixed point and let $\mathcal{F}(X)$ be the set of functions $f : X \to \mathbb{R}$. Then considering the application $\rho : \mathcal{F}(X) \times \mathcal{F}(X) \to [0, +\infty[$ given by $\rho(f, g) = |f(x_0) - g(x_0)|$ for all $f, g \in \mathcal{F}(X)$ we have that ρ is symmetric and that $\rho(f, f) = |f(x_0) - f(x_0)| = 0$ for all $f \in \mathcal{F}(x)$, so ρ is a semi-distance on $\mathcal{F}(X)$.*

Definition 1.1.2 (Distance)**.** *A distance ρ on a non-empty set X is a semi-distance on X such that*

$$\rho(x,y) = 0 \Rightarrow x = y.$$

A distance space (X,ρ) is a non-empty set X equipped with a distance ρ.

Example 1.1.4. *Let X be a set with at least two points and $k \in \mathbb{R}$ fixed such that $K > 0$. Then the application $\rho : X \times X \to [0,+\infty[$ given by $\rho(x,y) = 0$ if $x = y$ and $\rho(x,y) = k$ if $x \neq y$ for all $x,y \in X$ is trivially a distance on X.*

Example 1.1.5. *Let V be a linear space and p a norm on V. Then by definition of norm the application $\rho : V \times V \to [0,+\infty[$ given by $\rho(\mathbf{v},\mathbf{w}) = p(\mathbf{v}-\mathbf{w})$ for all $\mathbf{v},\mathbf{w} \in V$ is trivially a distance on V.*

Example 1.1.6. *Considering the application $\rho : \mathbb{R} \times \mathbb{R} \to [0,+\infty[$ given by $\rho(x,y) = (x-y)^2$ for all $x,y \in X$, then ρ is trivially a distance on $\mathbb{R}$.*

Example 1.1.7. *Considering the application $\rho : \mathbb{R}^2 \times \mathbb{R}^2 \to [0,+\infty[$ given by $\rho(x,y) = \left(\sqrt{|x_1-y_1|} + \sqrt{|x_2-y_2|}\right)$ for all $x = (x_1,x_2), y = (y_1,y_2) \in \mathbb{R}^2$, then ρ is trivially a distance on $\mathbb{R}^2$.*

Example 1.1.8. *Fixed $n \in \mathbb{N}$ with $n \geq 1$ and considering the application $\rho : \mathbb{R}^n \times \mathbb{R}^n \to [0,+\infty[$ given by $\rho(x,y) = \sum_{i=1}^{n} |x_i - y_i|^2$ for all $x = (x_1,\ldots,x_n), y = (y_1,\ldots,y_n) \in \mathbb{R}^n$, then ρ is trivially a distance on $\mathbb{R}^n$.*

It is clear that a distance space is also a semi-distance space, but in general the opposite can not be true.

Example 1.1.9. *Let us consider the application $\rho : \mathbb{R}^2 \times \mathbb{R}^2 \to [0,+\infty[$ given by $\rho(x,y) = |x_1 - y_1|$ for all $x = (x_1,x_2), y = (y_1,y_2) \in \mathbb{R}^2$. Then $\rho(x,x) = |x_1 - x_1| = 0$ for all $x \in \mathbb{R}^2$, so ρ is a semi-distance on $\mathbb{R}^2$. However this is not an example of a distance, since we can have two distinct points that are distance 0 from each other, indeed for all $x_1, x_2, y_2 \in \mathbb{R}$ such that $x_1 \neq y_2$ taking $x = (x_1,x_2), y = (x_1,y_2) \in \mathbb{R}^2$ we have $\rho(x,y) = |x_1 - x_1| = 0$.*

Definition 1.1.3 (Semi-metric and metric)**.** *A semi-metric (metric) d on a non-empty set X is a semi-distance (distance) on X such that the so called triangular inequality*

$$d(x,y) \leq d(x,z) + d(z,y) \tag{1.1}$$

holds for all $x,y,z \in X$. A semi-metric (metric) space (X,d) is a non-empty set X equipped with a semi-metric (metric) d.

Example 1.1.10. *Semi-distance spaces in the Examples 1.1.1 and 1.1.2, 1.1.3 have a semi-distance trivially satisfying the triangular inequality* (1.1)*, so they are also semi-metric spaces.*

Example 1.1.11. *Distance spaces in the Examples 1.1.4 and 1.1.5 have a distance trivially satisfying the triangular inequality* (1.1)*, so they are also metric spaces.*

It is clear that a metric space is also a distance space, but in general the opposite can not be true.

Example 1.1.12. *Distance space in the Example 1.1.6 is not a metric space. Indeed taking particular points* $x = 1$*,* $y = -1$ *and* $z = 0$ *we have* $\rho(x, y) = 4 > 1 + 1 = \rho(x, z) + \rho(z, y)$*, i.e. the distance* ρ *does not satisfy the triangular inequality* (1.1).

Example 1.1.13. *Distance space in the Example 1.1.7 is not a metric space. Indeed taking particular points* $x = (1, 0)$*,* $y = (0, 1)$ *and* $z = (0, 0)$ *we have* $\rho(x, y) = 4 > 1 + 1 = \rho(x, z) + \rho(z, y)$*, i.e. the distance* ρ *does not satisfy the triangular inequality* (1.1).

Example 1.1.14. *Distance space in the Example 1.1.8 is not a metric space. Indeed taking particular points* $x = (1, 0, \ldots, 1)$*,* $y = (-1, 0, \ldots, 0)$ *and* $z = (0, 0, \ldots, 1)$ *we have* $\rho(x, y) = 5 > 1 + 2 = \rho(x, z) + \rho(z, y)$*, i.e. the distance* ρ *does not satisfy the triangular inequality* (1.1).

It is clear that a metric space is also a semi-metric space, but in general the opposite can not be true.

Example 1.1.15. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a not injective function. Then the application* $\rho : \mathbb{R} \times \mathbb{R} \to [0, +\infty[$ *given by* $\rho(x, y) = |f(x) - f(y)|$ *for all* $x, y \in \mathbb{R}$ *is trivially a semi-metric on* $\mathbb{R}$*. By injectivity of function* f*, there exist two points* $x, y \in \mathbb{R}$ *with* $x \neq y$ *such that* $f(x) = f(y)$*, i.e.* $\rho(x, y) = |f(x) - f(y)| = 0$*, so* ρ *is not a metric on* X*.*

Definition 1.1.4 (Quasi-metric)**.** *A quasi-metric* ρ *with constant* A *on a non-empty set* X *is a distance on* X *such that there exists a real constant* A *for which the so called relaxed triangular inequality*

$$\rho(x, y) \leq A(\rho(x, z) + \rho(z, y)) \tag{1.2}$$

holds for all $x, y, z \in X$*. A quasi-metric space* (X, ρ) *is a non-empty set* X *equipped with a quasi-metric* ρ*.*

Definition 1.1.5 (Quasi-ultrametric)**.** *A quasi-ultrametric* ρ *with constant* K *on a non-empty set* X *is a distance on* X *such that there exists a real constant* K *for which* *the so called relaxed ultra-triangular inequality*

$$\rho(x,y) \le K \max\{\rho(x,z), \rho(z,y)\} \tag{1.3}$$

holds for all $x, y, z \in X$. A quasi-ultrametric space (X, ρ) is a non-empty set X equipped with a quasi-ultrametric ρ.

Excluding the trivial case by assuming that X be a set with at least two points we have both $A, K \in [1, +\infty[$. We can also note that a quasi-metric with constant $A = 1$ is a metric. Then it is usually to call ultrametric a quasi-ultrametric with constant $K = 1$. Moreover, by inequality $\frac{a+b}{2} \le \max\{a, b\} \le a + b$ true for all real numbers $a, b \ge 0$, it is easy to check that every quasi-metric with constant A is a quasi-ultrametric with constant $K = 2A$, and that every quasi-ultrametric with constant K is a quasi-metric with the same constant $A = K$ (so in particular an ultrametric is also a metric).

Example 1.1.16. *Distance space in the Example 1.1.4 have a distance trivially satisfying the relaxed ultra-triangular inequality* (1.3) *with constant $K = 1$, so it is also an ultrametric space.*

Example 1.1.17. *Let $D = \{0, 1\}$ and let $X = D^{\mathbb{N}}$. We note that elements of X are the sequences $x = \{x_i\}_{i \in \mathbb{N}}$ such that each x_i is either 0 or 1. For all $x, y \in X$ with $x \ne y$ let $\delta(x, y) = \min\{i \in \mathbb{N} : x_i \ne y_i\}$, i.e. $\delta(x, y)$ is the first index at which x and y disagree, and consider the application $\rho : X \times X \to [0, +\infty[$ given by*

$$\rho(x,y) = \begin{cases} 0 & if \quad x = y \\ \\ 2^{-\delta(x,y)} & if \quad x \ne y. \end{cases}$$

It is clear that ρ is symmetric and that $\rho(x, y) = 0$ if and only if $x = y$. If $x = z$, $y = z$, or $x = y$ it is clear that $\rho(x, y) \le \max\{\rho(x, z), \rho(z, y)\}$ holds. Taking three distinct points $x, y, z \in X$, let $h = \delta(x, z)$ and $l = \delta(z, y)$ assuming without loss of generality $h \le l$. Then $x_i = z_i = y_i$ for all $i \le k$, so $\delta(x, y) \ge k$, and therefore

$$\rho(x,y) = 2^{-\delta(x,y)} \le 2^{-k} = \max\left\{2^{-h}, 2^{-l}\right\} = \max\{\rho(x,z), \rho(z,y)\},$$

i.e. ρ satisfies the relaxed ultra-triangular inequality (1.3) *with constant $K = 1$, so X is an ultrametric space.*

Example 1.1.18. *Distance spaces in the Example 1.1.6 is not a metric space but is a quasi-metric space. Indeed for all fixed $x, y, z \in X$, let $a = \rho(x, y)$, $b = \rho(x, z)$*

and $c = \rho(z, y)$*, so we have* $\sqrt{a} = |x - y| \le |x - z| \le + |z - y| = \sqrt{b} + \sqrt{c}$*; by squaring both sides of* $\sqrt{a} \le= \sqrt{b} + \sqrt{c}$ *we have* $a \le b + c + 2\sqrt{bc}$ *and by the well-known fact that the geometric mean is bounded above by the arithmetic mean, i.e.* $\sqrt{bc} \le \frac{b+c}{2}$*, we obtain* $a \le (b + c)$ *so* ρ *satisfying the relaxed triangular inequality* (1.2) *with constant* $A = 2$*.*

Example 1.1.19. *Distance spaces in the Example 1.1.8 is not a metric space but is a quasi-metric space. Indeed for all fixed* $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n), z = (z_1, \ldots, z_n) \in \mathbb{R}^n$ *we have*

$$\begin{aligned}
2(\rho(x,z) + \rho(z,y)) =& 2\left(\sum_{i=1}^{n}(x_i - z_i)^2 + \sum_{i=1}^{n}(z_i - y_i)^2\right) \\
=& \sum_{i=1}^{n} 2\left((x_i - z_i)^2 + (z_i - y_i)^2\right) \\
\ge& \sum_{i=1}^{n}(x_i - z_i + z_i - y_i)^2 \\
=& \sum_{i=1}^{n}(x_i - y_i)^2 \\
=& \rho(x,y),
\end{aligned}$$

so ρ *satisfying the relaxed triangular inequality* (1.2) *with constant* $A = 2$*.*

About quasi-metric spaces we have the following obvious but useful result below.

Proposition 1.1.1. *Let* (X, ρ) *be a quasi-metric space with constant* A*. Then for any* $n \in \mathbb{N}$ *with* $n \ge 1$ *and for any points* $x_0, x_1, ..., x_n \in X$ *we have that*

$$\rho(x_0, x_n) \leqslant \sum_{i=1}^{n-1} A^i \rho(x_{i-1}, x_i) + A^{n-1}\rho(x_{n-1}, x_n). \tag{1.4}$$

Proof. Iterating the relaxed triangular inequality (1.2) the thesis holds. □

Furthermore, a better basic inequality than (1.4) can be found as illustrated in the following theorem [20].

Theorem 1.1.1. *Let* (X, ρ) *be a quasi-metric space with constant* A*. Then for any* $n \in \mathbb{N}$ *with* $n \ge 1$ *and for any points* $x_0, x_1, ..., x_n \in X$ *we have that*

$$\rho(x_0, x_n) \leqslant A^{f(n)} \sum_{i=1}^{n} \rho(x_{i-1}, x_i), \tag{1.5}$$

where the function $f : \mathbb{N} \setminus \{0\} \to \mathbb{N}$ is given by

$$f(n) = -\lceil -\log_2 n \rceil \tag{1.6}$$

Proof. We prove the thesis by induction on n. We first note that $2^{f(n)-1} < n \leqslant 2^{f(n)}$ holds for any $n \in \mathbb{N}$ and that obviously (1.5) holds for $n = 1$. We assume that (1.5) is true for n with $n \geq 2^k$ for same $k \in \mathbb{N}$. Fixed $n \in \mathbb{N}$ with $n \geq 1$ and $2^k < n \leqslant 2^{k+1}$ let us note that $f(n) = k+1$, $f(2^k) = k$ and $f(n-2^k) \leq k$, so it follows that

$$\begin{aligned}
\rho(x_0, x_n) \leqslant & A\rho(x_0, x_{2^k}) + A\rho(x_{2^k}, x_n) \\
& \leqslant AA^{f(2^k)} \sum_{i=1}^{2^k-1} \rho(x_{i-1}, x_i) + AA^{f(n-2^k)} \sum_{i=2^k}^{n} \rho(x_{i-1}, x_i) \\
& \leqslant A^{k+1} \sum_{i=1}^{n} \rho(x_{i-1}, x_i) \\
& = A^{f(n)} \sum_{i=1}^{n} \rho(x_{i-1}, x_i),
\end{aligned}$$

i.e. (1.5) is true for n proving giving the thesis. □

Remark 1.1.1. By proposition 1.1.1 because it is $A \geq 1$, considering the function $g : \mathbb{N} \setminus \{0\} \to \mathbb{N}$ given by

$$g(n) = n - 1 \tag{1.7}$$

clearly one has that

$$\rho(x_0, x_n) \leqslant A^{g(n)} \sum_{i=1}^{n} \rho(x_{i-1}, x_i), \tag{1.8}$$

for any $n \in \mathbb{N}$ with $n \geq 1$ and for any points $x_0, x_1, ..., x_n \in X$. By comparing the function f given by (1.6) with the function g given by (1.7) it is evident that the first allows a better estimate for a condition as the (1.8), being in general g much smaller than f. For instance by explicit computation one can find

$$f(1) = 0 = g(1),$$

$$f(2) = 1 = g(2),$$

$$f(3) = f(4) = 2 = g(3) < 3 = g(4),$$

$$f(5) = \cdots = f(8) = 3 < 4 = g(5) < 5 = g(6) < 6 = g(7) < 7 = g(8),$$

and so on.

About the special case of ultrametric spaces the relative theory is closely connected with various directions of investigations in mathematics, physics, linguistics, psychology and computer science, so different properties of ultrametric spaces have been studied [21, 22]. For completeness we recall some peculiar facts about these distance spaces. In the following for a given distance space (X, ρ), fixed $c \in X$ and $r > 0$, we call open ball of center c and radius r with respect to distance ρ the set $B(c,r) = \{x \in X : \rho(x,c) < r\}$, while $\overline{B}(c,r)$ will denote the closed one.

Proposition 1.1.2. *Let (X, ρ) be an ultrametric space. Then every triangle is isosceles with possibly a small edge.*

Proof. If there exist points $x, y, z \in X$ such that $\rho(x,y) < \rho(y,z) < \rho(x,z)$, then the ultrametric triangular inequality gives $\rho(x,z) \leq \max\{\rho(x,y), \rho(y,z)\} = \rho(xy,z)$, i.e. a contradiction. $\square$

Proposition 1.1.3. *Let (X, ρ) be an ultrametric space. For all $x, y \in X$ if $r = \rho(x,y)$ then $B(x,r) \cap B(y,r) = \emptyset$.*

Proof. Suppose $B(x,r) \cap B(y,r) \neq \emptyset$ and let $z \in B(x,r) \cap B(y,r)$. Considering the triangle formed by points x, y and z, we have $\rho(x,z) < r$ and since $\rho(x,y) = r$, by proposition 1.1.2 then it follows that $\rho(x,y) = r = \rho(z,y)$. But $\rho(y,z) < r$ as $z \in B(x,r) \cap B(y,r)$, i.e. a contradiction. $\square$

Proposition 1.1.4. *Let (X, ρ) be an ultrametric space. Then every point inside a ball is its center, i.e. if $\rho(x,y) < r$ it follows that $B(x,r) = B(y,r)$.*

Proof. Let $z \in B(x,r)$, then $\rho(x,z) < r$. Since $\rho(x,y) < r$, by proposition 1.1.2 we have $\rho(y,z) < r$ so $z \in B(y,r)$ and therefore $B(x,r) \subseteq B(y,r)$. Similarly we can show $B(x,r) \supseteq B(y,r)$ hence $B(x,r) = B(y,r)$. $\square$

Proposition 1.1.5. *Let (X, ρ) be an ultrametric space. Then intersecting balls are contained in each other, i.e. if $B(x,r) \cap B(y,s) \neq \emptyset$ it follows that either $B(x,r) \subseteq B(y,s)$ or $B(x,r) \supseteq B(y,s)$.*

Proof. Supposing $r < s$ let $z \in B(x,r) \cap B(y,s)$ so then $\rho(x,z) < r$ and $\rho(z,y) < s$. Therefore by proposition 1.1.4 we have $B(x,r) = B(z,r)$ and

$B(y,s) = B(z,s)$ and noting that $B(z,s) \supseteq B(z,r)$ as $r < s$ we have $B(x,r) \subseteq B(y,s)$. Similarly, if $s < r$ then $B(x,r) \supseteq B(y,s)$. $\square$

An important class of quasi-metric spaces then are the so called quasi-normed spaces according to next definition [23, 24].

Definition 1.1.6 (Quasi-norm). *A quasi-norm $\|\cdot\|$ with constant B on a linear space V over $\mathbb{K}$ (equal to $\mathbb{R}$ or $\mathbb{C}$) is a map $\|\cdot\| : V \to [0,+\infty[$ such that there exists a real constant B so that*

1) $\|\mathbf{v}\| = 0 \iff \mathbf{v} = \mathbf{0}$;
2) $\|\alpha\mathbf{v}\| = |\alpha|\,\|\mathbf{v}\|$;
3) $\|\mathbf{v}+\mathbf{w}\| \leqslant B(\|\mathbf{v}\| + \|\mathbf{w}\|)$,

for all $\mathbf{v},\mathbf{w} \in V$ and $\alpha \in \mathbb{K}$. A quasi-normed space $(V,\|\cdot\|)$ is a vector space V equipped with a quasi-norm $\|\cdot\|$. A complete quasi-normed space is called a quasi-Banach space.

It is therefore evident that if $(V,\|\cdot\|)$ is a quasi-normed space with constant B, then by definition 1.1.6 the map $\rho : V \times V :\to [0,+\infty[$ defined as $\rho(\mathbf{v},\mathbf{w}) = \|\mathbf{v}-\mathbf{w}\|$ for all $\mathbf{v},\mathbf{w} \in V$ is a quasi-metric, with the same constant B, on V.

Example 1.1.20. *Let $L^p(X,\mu)$ be the linear space of all measurable functions from X to $\mathbb{R}$ or $\mathbb{C}$ whose absolute value raised to the p-th power has a finite integral. Then if $0 < p < 1$ the map $\|\cdot\|_{L^p} : L^p(X,\mu) \to [0,+\infty[$ given by*

$$\|f\|_{L^p} = \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}},$$

for any $f \in L^p(X,\mu)$ is a quasi-norm on $L^p(X,\mu)$. Indeed it is easy to check that $\|\cdot\|_{L^p}$ satisfies conditions 1) and 2) of definition 1.1.6. Then also the inequality

$$\|f+g\|_{L^p} \leqslant 2^{\frac{1-p}{p}} \left(\|f\|_{L^p} + \|g\|_{L^p}\right) \tag{1.9}$$

for all $f,g \in L^p(X,\mu)$ holds, i.e., the condition 3) of definition 1.1.6 with $B = 2^{\frac{1-p}{p}}$. To prove (1.9) *first let us note that for fixed $\theta \in [1,+\infty[$, $n \in \mathbb{N}$ with $n \geq 1$ and $a_1,\dots,a_n \in [0,+\infty[$, the inequality*

$$\left(\sum_{i=1}^n a_i\right)^{\vartheta} \leqslant n^{\vartheta-1}\sum_{i=1}^n a_i^{\vartheta} \tag{1.10}$$

is true: in fact by $\theta \geq 1$ we have that the function $h : [0, +\infty \to [0, +\infty$ given by $h(x) = x^{\theta}$ for all $x \in \mathbb{R}$ is convex, hence

$$\begin{aligned}\left(\sum_{i=1}^{n} a_i\right)^{\theta} &= h\left(\frac{\sum\limits_{i=1}^{n} na_i}{n}\right)\\ &\leqslant \frac{1}{n}\sum_{i=1}^{n} h(na_i)\\ &= n^{\theta-1}\sum_{i=1}^{n} a_i^{\vartheta}.\end{aligned}$$

Then since $\frac{1}{p} > 1$ we obtain

$$\begin{aligned}\left\|\sum_{i=1}^{n} f_i\right\|_{L^p} &= \left(\int_X \left|\sum_{i=1}^{n} f_i\right|^p d\mu\right)^{\frac{1}{p}}\\ &\leq \left(\sum_{i=1}^{n}\int\limits_X |f_i|^p d\mu\right)^{\frac{1}{p}}\\ &\leq n^{\frac{1}{p}-1}\sum_{i=1}^{n}\left(\int_X |f_i|^p d\mu\right)^{\frac{1}{p}}\\ &= n^{\frac{1-p}{p}}\sum_{i=1}^{n}\|f_i\|_{L^p},\end{aligned}$$

i.e. the (1.9) for $n = 2$. We can also note that $n^{\frac{1-p}{p}}$ is the best possible constant in (1.10), so $2^{\frac{1-p}{p}}$ it is in (1.9). Indeed if E is a measurable set such that $\mu(E) = \alpha$, taking $E_i = E$ and the characteristic functions $f_i = \mathbb{1}_E$ for $1 \leq i \leq n$, we have

$$\begin{aligned}\left\|\sum_{i=1}^{n} f_i\right\|_{L^p} &= \left(\sum_{i=1}^{n}\mu(E_i)\right)^{\frac{1}{p}}\\ &= (n\alpha)^{\frac{1}{p}}\\ &= n^{\frac{1-p}{p}}\left(n\alpha^{\frac{1}{p}}\right)\\ &= n^{\frac{1-p}{p}}\sum_{i=1}^{n}\mu(E_i)^{\frac{1}{p}}\\ &= n^{\frac{1-p}{p}}\sum_{i=1}^{n}\|f_i\|_{L^p}.\end{aligned}$$

Example 1.1.21. *Let l^p be the linear space consisting of all real or complex sequences $\{x_k\}_{k\in\mathbb{N}}$ satisfying $\sum_{k\in\mathbb{N}} |x_k|^p < +\infty$. Then if $0 < p < 1$ the map $\|\cdot\|_{l^p} :\to [0,+\infty[$ given by*

$$\|x\|_{l^p} = \left(\sum_{k=1}^{+\infty} |x_k|^p\right)^{\frac{1}{p}},$$

for $x = \{x_k\}_{k\in\mathbb{N}} \in l^p$, is a quasi-norm on l_p. Indeed it is easy to check that $\|\cdot\|_p$ satisfies conditions 1) and 2) of definition 1.1.6 and the inequality

$$\|x+y\|_{l^p} \leq B\left(\|x\|_{l^p} + \|y\|_{l^p}\right)$$

for all $x, y \in l^p$ with the optimal constant $B = 2^{\frac{1-p}{p}}$.

Furthermore we note that quasi-ultrametrics and quasi-metrics can be generated from others using the following general proposition which will be also very useful later on.

Proposition 1.1.6. *Let (X,ρ) be a distance space and $\alpha \in]0,+\infty[$. The following facts hold:*

1) *if ρ is a quasi-ultrametric then ρ^α is a quasi-ultrametric;*
2) *if ρ is a quasi-metric then ρ^α is a quasi-metric;*
3) *if d is a metric then d^α is a metric when $0 < \alpha \leq 1$, a quasi-metric when $\alpha > 1$.*

Proof. 1) is trivial. In both two other cases thesis follows by observing that, for any $c > 0$, the function $[0,+\infty[\ni t \to (t+c)^\alpha - t^\alpha - c^\alpha$ is non-positive for any $0 < \alpha \leq 1$, and the maximum of the function $[0,+\infty[\ni t \to \frac{(t+c)^\alpha}{t^\alpha + c^\alpha}$ is $2^{\alpha-1}$ for any $\alpha > 1$. □

Example 1.1.22. *Let $d(x,y) = |x-y|$ for all $x, y \in X = \mathbb{R}$ the usual euclidean metric. Then $\rho(x,y) = d^2(x,y)$ for all $x, y \in X$ is a quasi-metric on X with constant $A = 2$, as already shown in Example 1.1.18.*

Remark 1.1.2. Note that the inequality $(t+c)^\alpha \leq 2^\alpha (t^\alpha + c^\alpha)$ for any $\alpha > 0$ also holds, and will be useful in the sequel.

Definition 1.1.7 (Relaxed polygonal inequality)**.** *A distance (semi-distance) ρ on a non-empty set X satisfies a relaxed polygonal inequality if there exists a real constant c such that*

$$\rho(x,y) \leq c \sum_{i=1}^{n} \rho(x_{i-1}, x_i) \tag{1.11}$$

for any $n \in \mathbb{N}$ and for any points $x_0 = x, x_1, ..., x_n = y \in X$.

Usually we call the set of points $x_0 = x, x_1, ..., x_n = y \in X$ a chain of length n with extremes x and y and the number $\sum\limits_{i=1}^{n} \rho(x_{i-1}, x_i)$ the sum on this chain. According to definition 1.1.7 immediately we have that an usual metric d satisfies a relaxed polygonal inequality with constant $c = 1$ and that a if a distance ρ satisfies a relaxed polygonal inequality with constant c then it is a quasi-metric with constant $A = c$. In any case, there are quasi-metrics and also quasi-ultrametric that do not satisfy a relaxed polygonal inequality as shown in the next example which will be very significant also for further considerations.

Example 1.1.23. *Let d be the Euclidean metric on $\mathbb{R}$, $\alpha \in]1, +\infty[$ and $\rho = d^\alpha$. Then ρ is both a quasi-metric with constant $A = 2^{\alpha-1}$ and a quasi-ultrametric with constant $K = 2^\alpha$, but the space $(\mathbb{R}, \rho)$ does not satisfy the relaxed polygonal inequality. Indeed, for any given $x, y \in \mathbb{R}$ with $x < y$, for any $n \in \mathbb{N}$, with $n \geq 1$, let us consider the points $x_i = x + (y - x) \cdot \frac{i}{n}$, $i = 0, 1, \ldots, n$. We have $\rho(x_{i-1}, x_i) = \left(\frac{|y-x|}{n}\right)^\alpha$ for any $i = 1, \ldots, n$ and, for some real number $c \geq 1$, the relaxed polygonal inequality would imply $|y - x|^\alpha = \rho(x,y) \leq c \cdot n \cdot \left(\frac{|y-x|}{n}\right)^\alpha$, i.e. $c \geq n^{\alpha-1}$ for all $n \in \mathbb{N}$, $n \geq 1$.*

Now we show a restrictive condition for a distance which guarantees that a relaxed polygonal inequality is verified. For this purpose we say that a distance ρ on X is bounded from below if there exists a constant $m > 0$ such that $m \leq \rho(x,y)$ for all $x, y \in X$ with $x \neq y$ and we say that ρ is bounded from above if a constant $M > 0$ there exists such that $\rho(x,y) \leq M$ for all $x, y \in X$. We also say that m and M are bounds for ρ.

Theorem 1.1.2. *Let (X, ρ) be a distance space. If ρ is bounded from below and from above with bounds m and M, then ρ satisfies a relaxed polygonal inequality with constant $c = \frac{M}{m}$.*

Proof. Fixed a chain $x_0 = x, x_1, ..., x_n = y \in X$, let $a = \sum_{i=1}^{n} \rho(x_{i-1}, x_i)$. If $a = 0$ we have $x = x_i = y$ for all i because ρ is a distance and so $\rho(x, y) = 0$ holds. If $a > 0$ then by hypothesis we have $m \leqslant \sum_{i=1}^{n} \rho(x_{i-1}, x_i)$ and $\rho(x, y) \leq M$ so $\rho(x, y) \leqslant \frac{M}{m} \sum_{i=1}^{n} \rho(x_{i-1}, x_i)$ holds. Setting $c = \frac{M}{m}$ then the inequality $\rho(x, y) \leqslant c \sum_{i=1}^{n} \rho(x_{i-1}, x_i)$ is true for all $x_0 = x, x_1, ..., x_n = y \in X$, i.e. the thesis. □

Remark 1.1.3. Theorem 1.1.2 is valid for all finite spaces and more in general, spaces that have infinite points for which theorem works can be constructed as follows. If d is a metric on X, then d^α is a quasi-metric on X for any $\alpha > 1$. Fixed $m > 0$ then for all $x, y \in X$ we set $\bar{\rho}(x, y) = 0$ if $x = y$ and $\bar{\rho}(x, y) = m + d^\alpha(x, y)$ if $x \neq y$, so $\bar{\rho}$ is a quasi-metric on X bounded below. Finally fixed $M \geq m$ and setting $\rho(x, y) = \min\{m + \bar{\rho}(x, y), M\}$ for all $x, y \in X$, it follows that ρ is a quasi-metric on X bounded below and above.

Example 1.1.24. *Let us take the space* $(\mathbb{R}, \rho)$ *with*

$$\rho(x,y) = \begin{cases} 0 \quad if \quad x = y \\ \\ \min\left\{1 + (x-y)^2, 8\right\} \quad if \quad x \neq y \end{cases}.$$

$(\mathbb{R}, \rho)$ *is a distance space and* ρ *is bounded. If* ρ *is a metric on* $\mathbb{R}$ *then trivially it satisfies a relaxed polygonal inequality with constant* $c = 1$*, but it is easy to check that* ρ *is a quasi-metric* $\mathbb{R}$ *with constant* 2 *but not a metric, in fact choosing points* $x = 3$*,* $y = 0$*,* $z = 1$ *we have* $\rho(x, y) = 8$*,* $\rho(x, z) = 5$*,* $\rho(z, y) = 2$*, i.e.* $\rho(x, y) > \rho(x, z) + \rho(z, y)$*. Anyway by theorem 1.1.2* ρ *satisfies a relaxed polygonal inequality with constant* $c = 8$ *because it is* $1 \leqslant \rho(x, y) \leqslant 8$ *for all* $x, y \in \mathbb{R}$ *with* $x \neq y$*.*

Starting from a semi-distance ρ on X it is possible to generate a semi-metric on X through the so called chain approach technique essentially due to Frink [3]. This semi-metric is defined as the infimum over the sum of all chains of points in the space that have fixed extremes and it proves to be very useful for various arguments involving a relaxed polygonal inequality. Then, under suitable conditions, this can be a bi-lipschitz equivalent metric with respect to the given ρ according to a definition that will be given in the next section [6, 25]. More precisely for a given semi-distance space (X, ρ) let $d_\rho : X \times X \to [0, +\infty[$ be defined as follows:

$$d_\rho(x,y) = \inf_{\substack{n\in\mathbb{N}\\ n\geq 1}} \left\{ \inf_{\substack{z_0,z_1,\dots,z_{n-1},z_n\in X\\ z_0=x,z_n=y}} \left\{ \sum_{i=1}^{n} \rho(z_{i-1},z_i) \right\} \right\} \tag{1.12}$$

$$= \inf \left\{ \sum_{i=1}^{n} \rho(z_{i-1},z_i) : z_0, z_1, \dots, z_{n-1}, z_n \in X, z_0 = x, z_n = y, n \in \mathbb{N}, n \geqslant 1 \right\}.$$

Note that all infima are actually minima when X is a finite set. We call d_ρ the Frink's map induced by ρ and show the next fundamental lemma [26].

Lemma 1.1.1. *Let (X,ρ) be semi-distance space. Then the Frink's map $d_\rho : X \times X \to [0,+\infty[$ induced by ρ is a semi-metric on X.*

Proof. Obviously by 1.12 we have $d_\rho(x,y) = 0$ if $x = y$ so only need to check the triangle inequality for d_ρ. Let x, y and z be three given points in X and let $\epsilon > 0$. Then there exist two chains of points $x_0 = x, x_1, ..., x_{n-1} = z$ and $x_n = z, x_{n+1}, ..., x_m = y \in X$ such that

$$\sum_{i=0}^{n-1} \rho(x_{i-1},x_i) < d_\rho(x,z) + \frac{\varepsilon}{2} \quad \text{and} \quad \sum_{i=n}^{m} \rho(x_{i-1},x_i) < d_\rho(z,y) + \frac{\varepsilon}{2},$$

so adding the above inequalities we see that

$$d_\rho(x,y) \leqslant \sum_{i=0}^{m} \rho(x_{i-1},x_i) < d_\rho(x,z) + d_\rho(z,y) + \varepsilon$$

for every $\epsilon > 0$, getting the thesis. □

The next facts concern results connecting Frink's map induced by a given distance ρ and the relaxed polygonal inequality for ρ.

Theorem 1.1.3. *Let (X,ρ) be a distance space. Then if ρ satisfies a relaxed polygonal inequality with constant $c \geq 1$ then d_ρ is a metric on X.*

Proof. We prove that if d_ρ is not a metric on X, then ρ does not satisfy the relaxed polygonal inequality. Assuming that d_ρ is not a metric on X there exist two points $x, y \in X$ such that $x \neq y$ and $d_\rho(x,y) = 0$. If ρ satisfies the relaxed polygonal inequality for some constant $c \in [1,+\infty[$, then for any points $x_1, ..., x_n \in X$ it follows

$$\rho(x,y) \leq c(\rho(x,x_1) + ... + \rho(x_n,y)), \tag{1.13}$$

and by definition of d_ρ it follows that for any $\varepsilon \in]0,+\infty[$ there exist the points $\bar{x}_1, ..., \bar{x}_{\bar{n}} \in X$, depending on ε and x, y, such that

$$\rho(x,\bar{x}_1) + ... + \rho(\bar{x}_{\bar{n}},y) < \varepsilon. \tag{1.14}$$

By (1.13) choosing $x_1 = \bar{x}_1, ..., x_n = \bar{x}_{\bar{n}}$ and by (1.14) it follows $\frac{1}{c}\rho(x,y) < \varepsilon$, and tending with ε to 0^+ we have $\rho(x,y) \leq 0$, i.e. a contradiction because ρ is a distance but $x \neq y$, thus the thesis is proved. $\square$

Theorem 1.1.3 gives a necessary condition such that ρ satisfies a relaxed polygonal inequality, and in general this theorem cannot be inverted as shown by the following example.

Example 1.1.25. *If ρ is a distance on X bounded from below then also d_ρ is bounded from below, so it is a metric on X. Let us take the space $(\mathbb{R}, \rho)$ with*

$$\rho(x,y) = \begin{cases} 0 \quad if \quad x = y \\ \\ 1 + (x-y)^2 \quad if \quad x \neq y \end{cases}.$$

$(\mathbb{R}, \rho)$ is a distance space bounded from below so d_ρ is a metric on $\mathbb{R}$ but ρ does not satisfy a relaxed polygonal inequality. Indeed if there exists $c \geq 1$ such that $\rho(x,y) \leq c \sum\limits_{i=1}^{n} \rho(x_{i-1}, x_i)$ for any $n \in \mathbb{N}$ and for any $x_i \in \mathbb{R}$ with $x_0 = x$ and $x_n = y$, let us consider the points $x_i = 2\lceil c \rceil \frac{i}{n}$ for any $i = 0, \dots, n$ with n arbitrarily fixed, so $\sum\limits_{i=1}^{n} \rho(x_{i-1}, x_i) = n + \frac{4\lceil c \rceil^2}{n}$ and $\rho(x,y) = 1 + 4\lceil c \rceil^2$. It follows that $1 + 4\lceil c \rceil^2 \leqslant c(n + \frac{4\lceil c \rceil^2}{n})$ and assuming $n = 2\lceil c \rceil$ then we have $1 \leqslant 4\lceil c \rceil(c - \lceil c \rceil) \leqslant 0$, a contradiction.

Corollary 1.1.1. *Let (X, ρ) be a distance space such that ρ does not satisfy a relaxed polygonal inequality. Then ρ is unbounded from below if ρ is bounded from above and ρ is unbounded from above if ρ is bounded from below.*

Remark 1.1.4. Corollary 1.1.1 works if ρ does not satisfy a relaxed polygonal inequality. By theorem 1.1.3 we can assume that d_ρ is not a metric on X to ensure this condition. However we note that if ρ is bounded from below then also d_ρ is bounded from below, so d_ρ is a metric on X. Hence in this case ρ cannot be bounded from below.

The last proposition of this section shows when a distance ρ satisfies a more general relaxed polygonal inequality for the which the constant c is not absolute but depends on the pair of extreme points of the chain (x, y).

Proposition 1.1.7. *Let (X, ρ) be a distance space. If d_ρ is a metric on X then ρ satisfies an extended relaxed polygonal inequality, in the sense that for any points $x_0 = x, x_1, ..., x_n = y \in X$ a constant $c(x, y) \geq 1$ depending only on x and y exists such that $\rho(x, y) \leq c(x, y) \sum\limits_{i=1}^{n} \rho(x_{i-1}, x_i)$.*

Proof. For any fixed $x, y \in X$ such that $x \neq y$ we consider the set

$$I_{x,y} = \left\{ \frac{\sum\limits_{i=1}^{n} \rho(x_{i-1}, x_i)}{\rho(x, y)} : x_0 = x, x_1, ..., x_n = y \in X \right\}.$$

We note that $\inf I_{x,y} > 0$. In fact it is trivially $\inf I_{x,y} \geqslant 0$ and if $\inf I_{x,y} = 0$ the definition of $I_{x,y}$ gives that for any $\varepsilon > 0$ there exist the points $x_0, x_1, ..., x_n \in X$ depending on ε such that $\frac{\sum_{i=1}^{n} \rho(x_{i-1}, x_i)}{\rho(x,y)} < \varepsilon$, thus by definition of d_ρ we have $\frac{d_\rho(x,y)}{\rho(x,y)} < \varepsilon$: tending with ε to 0^+, $d_\rho(x, y) \leq 0$, i.e. a contradiction because d_ρ is a metric but $x \neq y$. Setting

$$c(x, y) = \begin{cases} 1 \quad if \quad x = y \\ \\ \frac{1}{\inf I_{x,y}} \quad if \quad x \neq y \end{cases},$$

we have the thesis. $\square$

1.2 The metrization problem for quasi-ultrametric spaces

Sufficient conditions for the existence of a metric compatible with a topology on a given distance space (X, ρ) are well known with ρ either quasi-ultrametric or quasi-metric. If on the set X it is not a priori defined a topology, it is possible to declare open a set for which any of its points has a ball contained in the given set; however these balls, despite being a basis of neighborhoods, are not open sets in general. Indeed in both two cases, either quasi-ultrametric with constant K or quasi-metric with constant A, given the ball $B(c,r)$, it is immediately verified that for any $x \in B(c,\frac{r}{K})$ (resp. $x \in B(c,\frac{r}{2A})$) there exists a suitable ball $B(x,\sigma) \subseteq B(c,r)$. So, by means of the quasi-distance ρ, we can define a uniformity $\mathcal{U}_\rho$ [27], such that the balls form a basis of neighborhoods for the topology τ_ρ induced by $\mathcal{U}_\rho$; we can then apply a variant of the classic Nagata and Smirnov metrization theorem [18], thus, there exists a metric d on X such that $\tau_d = \tau_\rho$ where τ_d is the topology induced by d.

Theorem 1.2.1 (Nagata-Smirnov's theorem). *Let X be a T_0 topological space. Then X is metrizable if and only if every $y \in X$ has an open neighbourhood basis $\{U_{y,n}\}_{n\in\mathbb{N}}$ such that*

a) If $x \in U_{y,n}$ then $U_{x,n} \subseteq U_{y,n-1}$;

b) If $x \notin U_{y,n-1}$ then $U_{x,n} \cap U_{y,n} = \emptyset$.

Indeed, for all $y \in X$ the sets $B_\rho\big(y,\frac{1}{(2A)^n}\big)$ satisfy the hypothesis of the theorem 1.2.1, so setting $U_{y,n} = \operatorname{int}\Big(B_\rho\big(y,\frac{1}{(2A)^n}\big)\Big)$ there exists a metric d on X such that the topology τ_d induced by d is the topology τ_ρ induced by ρ.

Remark 1.2.1. The topology τ_ρ generated by a quasi-ultrametric or quasi-metric ρ has some peculiarities: a ball $B(c,r)$ need not be τ_ρ open [28].

Example 1.2.1. *Fixed $\epsilon > 0$ consider the space $(\mathbb{N},\rho)$ with $\rho : \mathbb{N}\times\mathbb{N} \to [0,+\infty[$ given by*

$$\rho(0,1) = 1, \quad \rho(0,m) = 1+\varepsilon \quad for \quad m \geq 2$$

$$\rho(1,m) = \frac{1}{m}, \quad \rho(n,m) = \frac{1}{n}+\frac{1}{m} \quad for \quad n \geq 2$$

and extended to $\mathbb{N}\times\mathbb{N}$ by $\rho(n,n) = 0$ and symmetry. Then by a direct check it follows that

$$\rho(k,n) \leq (1+\varepsilon)(\rho(k,m)+\rho(m,n)) \tag{1.15}$$

for all $k,n,m \in \mathbb{N}$, i.e. ρ is a quasi-metric on $\mathbb{N}$ with constant $A = 1+\varepsilon$. In fact it is clear that it suffices to check (1.15) *for pairwise distinct k,m,n only. Let L and R denote the left and the right hand sides of the inequality* (1.15) *respectively. If one of k,n,m is equal to* 0*, then $L \leq 1+\varepsilon \leq R$. If none of k,m,n are* 0*, then we consider subcases. First, assume* 1 *appears among k,m,n. If $k = 1$, then $L = \frac{1}{n}$ while $R = (1+\varepsilon)\left(\frac{2}{m}+\frac{1}{n}\right)$. Similarly if $n = 1$. If $m = 1$, then $L = \frac{1}{k}+\frac{1}{n}$*

while $R = (1+\varepsilon)\left(\frac{1}{k}+\frac{1}{n}\right)$. *Next, assume* 1 *does not appear among* k, n, m. *Then* $L = \frac{1}{k}+\frac{1}{n}$ *and* $R = (1+\varepsilon)\left(\frac{1}{k}+\frac{2}{m}+\frac{1}{n}\right)$, *finishing our checking so* ρ *satisfies* (1.15). *Now noting that* $B(0, 1+\frac{\varepsilon}{2}) = \{0, 1\}$ *while* $B(1, r)$ *contains infinitely many elements for any* $r > 0$, *none of* $B(1, r)$ *are contained in* $B(0, 1+\frac{\varepsilon}{2})$, *i.e. the ball* $B(0, 1+\frac{\varepsilon}{2})$ *is not open.*

Remark 1.2.2. Another peculiarity in a quasi-ultrametric or quasi-metric space (X, ρ) is that ρ could not be continuous [29, 30].

Example 1.2.2. *Let* $X = \left\{0, 1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots\right\}$ *and consider the map* $\rho : X \times X \to [0, +\infty]$ *given by*

$$\rho(x,y) = \begin{cases} 0 \quad if \quad x = y \\ \\ 1 \quad if \quad x, y \in \{0, 1\} \quad with \quad x \neq y \\ \\ |x-y| \quad if \quad x, y \in \left\{0, \frac{1}{2n}, \frac{1}{2m}\right\} \quad with \quad x \neq y \\ \\ 0 \quad otherwise \end{cases}.$$

By explicit calculation we have that ρ *is a quasi-metric on* X *with constant* $A = \frac{8}{3}$. *Indeed for all* $x, y \in X$ *we have* $\rho(x,y) \geq 0$, $\rho(x,y) = 0$ *if and only if* $x = y$ *and* $\rho(x,y) = \rho(y,x)$. *If* $\rho(x,y) = \rho(0,1) = 1$, *then* $\rho(x,z) + \rho(z,y)$ *is*

$$\begin{cases} \rho(0, \dfrac{1}{2n}) + \rho(\dfrac{1}{2n}, 1) = \dfrac{1}{2n} + 4 \quad if \quad z = \dfrac{1}{2n} \\ \rho(0, \dfrac{1}{2n+1}) + \rho(\dfrac{1}{2n+1}, 1) = 4 + 4 = 8 \quad if \quad z = \dfrac{1}{2n+1} \end{cases};$$

if $\rho(x,y) = \rho(0, \frac{1}{2n}) = \frac{1}{2n}$ *then* $\rho(x,z) + \rho(z,y)$ *is*

$$\begin{cases} \rho(0, \dfrac{1}{2m}) + \rho(\dfrac{1}{2m}, \dfrac{1}{2n}) = \dfrac{1}{2m} + \left|\dfrac{1}{2m} - \dfrac{1}{2n}\right| \quad if \quad z = \dfrac{1}{2m} \\ \rho(0, \dfrac{1}{2m+1}) + \rho(\dfrac{1}{2m+1}, \dfrac{1}{2n}) = 4 + 4 = 8 \quad if \quad z = \dfrac{1}{2m+1} \neq 1 \\ \rho(0,1) + \rho(1, \dfrac{1}{2n}) = 1 + 4 = 5 \quad if \quad z = 1 \end{cases};$$

if $\rho(x,y)=\rho(\frac{1}{2k},\frac{1}{2n})=\left|\frac{1}{2k}-\frac{1}{2n}\right|$ *then then* $\rho(x,z)+\rho(z,y)$ *is*

$$\begin{cases}\rho(\frac{1}{2k},\frac{1}{2m})+\rho(\frac{1}{2m},\frac{1}{2n})=\left|\frac{1}{2k}-\frac{1}{2m}\right|+\left|\frac{1}{2n}-\frac{1}{2n}\right| & if \quad z=\frac{1}{2m}\\ \rho(\frac{1}{2k},\frac{1}{2m+1})+\rho(\frac{1}{2m+1},\frac{1}{2n})=4+4=8 \quad if \quad z=\frac{1}{2m+1} & ;\\ \rho(\frac{1}{2k},0)+\rho(0,\frac{1}{2n})=\frac{1}{2k}+\frac{1}{2n} \quad if \quad z=0 & \end{cases}$$

if $\rho(x,y)=\rho(\frac{1}{2k},\frac{1}{2n+1})=4$ *with* $\frac{1}{2n+1}\neq 1$, *then* $\rho(x,z)+\rho(z,y)$ *is*

$$\begin{cases}\rho(\frac{1}{2k},0)+\rho(0,\frac{1}{2n+1})=\frac{1}{2k}+4 \quad if \quad z=0 & \\ \rho(\frac{1}{2k},\frac{1}{2m})+\rho(\frac{1}{2m},\frac{1}{2n+1})=\left|\frac{1}{2k}-\frac{1}{2m}\right|+4 \quad if \quad z=\frac{1}{2m} & ;\\ \rho(\frac{1}{2k},\frac{1}{2m+1})+\rho(\frac{1}{2m+1},\frac{1}{2n+1})=4+4=8 \quad if \quad z=\frac{1}{2m+1} & \end{cases}$$

if $\rho(x,y)=\rho(\frac{1}{2k+1},\frac{1}{2n+1})=4$ *with* $\frac{1}{2k+1}\neq 1$ *and* $\frac{1}{2n+1}\neq 1$ *then* $\rho(x,z)+\rho(z,y)$ *is*

$$\begin{cases}\rho(\frac{1}{2k},0)+\rho(0,\frac{1}{2n+1})=4+4=8 \quad if \quad z=0 & \\ \rho(\frac{1}{2k+1},\frac{1}{2m})+\rho(\frac{1}{2m},\frac{1}{2n+1})=4+4=8 \quad if \quad z=\frac{1}{2m} & ;\\ \rho(\frac{1}{2k+1},\frac{1}{2m+1})+\rho(\frac{1}{2m+1},\frac{1}{2n+1})=4+4=8 \quad if \quad z=\frac{1}{2m+1} & \end{cases}$$

if $\rho(x,y)=\rho(\frac{1}{2k},1)=4$ *then* $\rho(x,z)+\rho(z,y)$ *is*

$$\begin{cases}\rho(\frac{1}{2k},0)+\rho(0,1)=\frac{1}{2k}+1 \quad if \quad z=0 & \\ \rho(\frac{1}{2k},\frac{1}{2m})+\rho(\frac{1}{2m},1)=\left|\frac{1}{2k}-\frac{1}{2m}\right|+4 \quad if \quad z=\frac{1}{2m} & ;\\ \rho(\frac{1}{2k},\frac{1}{2m+1})+\rho(\frac{1}{2m+1},1)=4+4=8 \quad if \quad z=\frac{1}{2m+1}\neq 1 & \end{cases}$$

if $\rho(x,y)=\rho(\frac{1}{2k+1},1)=4$ *then* $\rho(x,z)+\rho(z,y)$ *is*

$$\begin{cases}\rho(\frac{1}{2k+1},0)+\rho(0,1)=4+1=5 \quad if \quad z=0 & \\ \rho(\frac{1}{2+1k},\frac{1}{2m})+\rho(\frac{1}{2m},1)=4+4=8 \quad if \quad z=\frac{1}{2m} & ;\\ \rho(\frac{1}{2k+1},\frac{1}{2m+1})+\rho(\frac{1}{2m+1},1)=4+4=8 \quad if \quad z=\frac{1}{2m+1}\neq 1 & \end{cases}$$

if $\rho(x,y) = \rho(\frac{1}{2k+1}, 0) = 4$ *then* $\rho(x,z) + \rho(z,y)$ *is*

$$\begin{cases} \rho(\frac{1}{2k+1}, 1) + \rho(1,0) = 4+1 = 5 \quad if \quad z = 1 \\ \rho(\frac{1}{2+1k}, \frac{1}{2m}) + \rho(\frac{1}{2m}, 0) = 4 + \frac{1}{2m} \quad if \quad z = \frac{1}{2m} \\ \rho(\frac{1}{2k+1}, \frac{1}{2m+1}) + \rho(\frac{1}{2m+1}, 0) = 4+4 = 8 \quad if \quad z = \frac{1}{2m+1} \neq 1 \end{cases}.$$

So previous calculations give that $\rho(x,y) \leq \frac{8}{3}(\rho(x,z) + \rho(z,y))$ *for all* $x, y, z \in X$, *i.e.* ρ *is a quasi-metric on* X *with constant* $A = \frac{8}{3}$ *as mentioned. Then we also have*

$$\rho(1, \frac{1}{2}) = 4 > 1 + \frac{1}{2} = \rho(1,0) + \rho(0, \frac{1}{2}),$$

proving that ρ *is not a metric on* X. *Now noting that it is*

$$\lim_{n\to+\infty} \rho(0, \frac{1}{2n}) = \lim_{n\to+\infty} \frac{1}{2n} = 0,$$

and also $\lim_{n\to+\infty} \frac{1}{2n} = 0$ *in* (X, ρ), *however we have*

$$\lim_{n\to+\infty} \rho(1, \frac{1}{2n}) = 4 \neq 1 = \rho(1,0),$$

proving that ρ *is a quasi-metric on* X *not continuous in each variable.*

The following result instead gives a sufficient condition such that open balls and closed balls are open and closed in the topology τ_ρ respectively in any quasi-metric space.

Proposition 1.2.1. *Let* (X, ρ) *be a quasi-metric space with constant* A. *Then if* ρ *is continuous in one variable, then* ρ *is continuous in other variable. Moreover, for any* $c \in X$ *and* $r > 0$ *we have*

i) $B(c,r)$ *is open in* τ_ρ *;*

ii) $\bar{B}(c,r)$ *is open in* τ_ρ.

Proof. Assuming without loss of generality ρ continuous with respect to the first variable, then for any $x \in X$ if $\{y_n\}_{n\in\mathbb{N}} \subseteq X$ is a sequence such that

$\lim_{n\to+\infty} y_n = y$ we have

$$\begin{aligned}\lim_{n\to+\infty} \rho(x, y_n) &= \lim_{n\to+\infty} \rho(y_n, x) \\ &= \rho(y, x) \\ &= \rho(x, y),\end{aligned}$$

proving that ρ is continuous with respect to the second variable.

i) Let $\{x_n\}_{n\in\mathbb{N}} \subseteq X\backslash B(c,r)$ be a sequence such that $\lim_{n\to+\infty} x_n = x$. Then for all $n \in \mathbb{N}$ we have $\rho(c, x_n) \geq r$ and $\rho(c, x) = \lim_{n\to+\infty} \rho(c, x_n) \geqslant r$. This implies that $x \in X\backslash B(c,r)$ and therefore $X\backslash B(c,r)$ is a closed set, i.e. the ball $B(c,r)$ is open.

ii) Let $\{x_n\}_{n\in\mathbb{N}} \subseteq \bar{B}(c,r)$ be a sequence such that $\lim_{n\to+\infty} x_n = x$. Then for all $n \in \mathbb{N}$ we have $\rho(c, x_n) \leq r$ and $\rho(c, x) = \lim_{n\to+\infty} \rho(c, x_n) \leq r$. This implies that $x \in \bar{B}(c,r)$ and therefore $\bar{B}(c,r)$ is a closed set. $\square$

Anyway for special case there exist also direct proofs of the metrizability of the topology of a quasi-ultrametric space based on an approach due to Frink that uses chains [3, 4] and general results of metrizability were obtained by a slight modification of Frink's technique [26, 28]. In the rest of this section we present new proofs of some of these facts and we are interested to the existence of a metric d on X bi-lipschitz equivalent with respect to a given quasi-ultrametric ρ, or to a suitable power of a given quasi-metric ρ. Let us start with a definition.

Definition 1.2.1 (Topological and bi-lipschitz equivalence). *Two distance ρ_1 and ρ_2 on the same no empty set X are called*

i) *topologically equivalent if $\tau_{\rho_1} = \tau_{\rho_2}$;*

ii) *bi-lipschitz equivalent if there exist two real constant $c_1, c_2 > 0$ such that*

$$c_1\rho_2(x,y) \leq \rho_1(x,y) \leq c_2\rho_2(x,y)$$

for all $x, y \in X$.

Of course, the above definitions applies to metrics as well, as particular cases of quasi-ultrametrics and note that without loss of generality, we can take $c_1 = 1$ by replacing the distance ρ_2 by the distance $c_1\rho_2$.

Proposition 1.2.2. *Let (X, ρ_1) and (X, ρ_2) be distance spaces on the same set X. If ρ_1 and ρ_2 are bi-lipschitz equivalent then ρ_1 and ρ_2 are also topologically equivalent.*

Proof. By definition of bi-lipschitz equivalence there exist two real constant $c_1, c_2 > 0$ such that $c_1\rho_2(x,y) \leq \rho_1(x,y) \leq c_2\rho_2(x,y)$ for all $x, y \in X$. Let $x \in X$, let $\varepsilon > 0$, and let $B_{\rho_1}(x, c_1\varepsilon)$ denote the open ball of center x and radius c_ε with respect to distance ρ_1. If $y \in B_{\rho_1}(x, c_1\varepsilon)$ then $\rho_1(x,y) < c_1\varepsilon$, and, by hypothesis, $\rho_2(x,y) \leqslant \frac{\rho_1(x,y)}{c_1} < \varepsilon$, so $y \in B_{\rho_2}(x,\varepsilon)$ and $B_{\rho_2}(x, c_1\varepsilon) \subseteq B_{\rho_2}(x,\varepsilon)$ due to the arbitrariness of y. Similarly if $y \in B_{\rho_2}(x, \frac{\varepsilon}{c_2})$ we have $\rho_2(x,y) < \frac{\varepsilon}{c_2}$ and $\rho_1(x,y) \leqslant c_2\frac{\rho_2(x,y)}{c_1} < \varepsilon$, so $y \in B_{\rho_1}(x,\varepsilon)$ and $B_{\rho_2}(x, \frac{\varepsilon}{c_2}) \subseteq B_{\rho_1}(x,\varepsilon)$. Now suppose $A \in \tau_{\rho_1}$, i.e. it is an open set with respect to the distance ρ_1. If $x \in A$ then there exists $\varepsilon > 0$ such that $B_{\rho_1}(x,\varepsilon) \subseteq A$ thus $B_{\rho_2}(x, \frac{\varepsilon}{c_2}) \subseteq B_{\rho_1}(x,\varepsilon) \subseteq A$ and so A is open with respect to distance ρ_2, i.e. $A \in \tau_{\rho_2}$, Mutatis mutandis, if $A \in \tau_{\rho_2}$ it follows that $A \in \tau_{\rho_1}$ so $\tau_{\rho_1} = \tau_{\rho_2}$. □

Remark 1.2.3. In general proposition 1.2.2 can not be inverted, i.e. topological equivalence does not imply bi-lipschitz equivalence. A reason is that bounded sets under one metric are also bounded under an bi-lipschitz equivalent metric, but not necessarily under a topologically equivalent metric.

Example 1.2.3. *Let (X, d) be an unbounded metric space. Then (X, ρ) with $\rho(x,y) = \frac{d(x,y)}{1+d(x,y)}$ for all $x, y \in X$ is also a metric space which is bounded, and so not bi-lipschitz equivalent with respect to (X, d), because there is no constant $c_2 > 0$ such that $d_1(x,y) \leq c_2 d_2(x,y)$ as the inequality $t \leqslant c_2\frac{t}{1+t}$ does not hold for $t > c_2$. Anyway ρ induces the same topology on X. Indeed to show this one can simply use the fact that d and ρ agree on small balls. If $A \in \tau_d$, then for $x \in A$ there exists a radius $r > 0$ such that $B_d(x,r) \subseteq A$. Supposing that for some $y \in A$ and $r > 0$ it holds that $B_\rho(y,r) \subseteq A$ then for any $q \in (0,r)$ it holds that $B_\rho(y,q) \subseteq A$. In particular, it holds for $q = min\left\{\frac{1}{2}, r\right\} < 1$ but then $B_\rho(t,q) = B_d(y,q)$.*

Example 1.2.4. *Let $X = (0,1]$ and let $d_1(x,y) = |x-y|$, $d_2(x,y) = \left|\frac{1}{x} - \frac{1}{y}\right|$ for all $x, y \in X$. Then (X, d_1) and (X, d_2) are metric spaces, with d_1 and d_2 topologically equivalent distances but not bi-lipschitz equivalent distances as $x, y \to 0$.*

In the case of quasi-norms anyway topological equivalence is equivalent to bi-lipschitz equivalence as shown by next proposition.

Proposition 1.2.3. *Let $(V, \|\cdot\|_1)$ and $(V, \|\cdot\|_2)$ be quasi-normed spaces on the same linear space V. If $\|\cdot\|_1$ and $\|\cdot\|_2$ are topologically equivalent then $\|\cdot\|_1$ and $\|\cdot\|_2$ are also bi-lipschitz equivalent.*

Proof. If the topologies induced by two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on the vector space V are identical, there exists open balls such that we have the inclusion $B_{\|\cdot\|_1}(\mathbf{0}, r) \subseteq B_{\|\cdot\|_2}(\mathbf{0}, 1)$ for some $r > 0$. Let $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ and set $\mathbf{w} = \frac{r}{2} \frac{\mathbf{v}}{\|\mathbf{v}\|_1}$. Then $\|\mathbf{w}\|_1 = \frac{r}{2} < r$, hence $\|\mathbf{w}\|_2 < 1$ which shows $\|\mathbf{v}\|_2 \leqslant \frac{2}{r}\|\mathbf{v}\|_1$. Finally by symmetry we have that $\|\cdot\|_1$ and $\|\cdot\|_2$ are bi-lipschitz equivalent. □

If the topology τ_ρ generated by a distance ρ on a set X is metrizable this means that there exists a metric d on X topologically equivalent to ρ. The problem of the existence of a metric that is bi-lipschitz equivalent with respect to a distance, also called metric boundedness property, is solved [6, 25] by the next theorem of considerable importance because unexpectedly characterizes such distance as those that satisfy a relaxed polygonal inequality.

Theorem 1.2.2. *Let (X, ρ) be a distance space. Then ρ is bi-lipschitz equivalent with respect to a metric d if and only if ρ satisfies a relaxed polygonal inequality* (1.11) *for some constant $c \geq 1$.*

Proof. Let ρ be bi-lipschitz equivalent with respect to a metric d, i.e. there exists a constant $c \geq$ such that $d(x,y) \leq \rho(x,y) \leq cd(x,y)$ for all $x, y \in X$. Then we have $\rho(x,y) \leq cd(x,y) \leq c\sum_{i=1}^{n} \rho(z_{i-1}, z_i)$ for all $n \in \mathbb{N}$ with $n \geq 1$ and for all $z_1, \ldots, z_n \in X$ with $z_0 = x$, $z_n = y$ because d is a metric and therefore satisfies a relaxed polygonal inequality with constant 1; then by $d(x,y) \leq \rho(x,y)$ we have $\rho(x,y) \leq c\sum_{i=1}^{n} \rho(z_{i-1}, z_i)$, i.e. ρ satisfies a relaxed polygonal inequality with constant $c \geq 1$. For the converse, let ρ be satisfying a relaxed polygonal inequality with constant $c \geq 1$, i.e.

$$\rho(x,y) \leq c \sum_{i=1}^{n} \rho(z_{i-1}, z_i) \tag{1.16}$$

for all $n \in \mathbb{N}$ with $n \geq 1$ and for all $z_1, \ldots, z_n \in X$ with $z_0 = x$, $z_n = y$ and $d_\rho = \inf \left\{ \sum\limits_{i=1}^{n} \rho(z_{i-1}, z_i) : z_0, z_1, \ldots, z_{n-1}, z_n \in X, z_0 = x, z_n = y, n \in \mathbb{N} \right\}$ the Frink's map (1.12) induced by ρ. By definition of d_ρ it follows that $d_\rho(x,y) \leq$

$\rho(x,y)$ for all $x,y \in X$ and taking infimum on all chain with fixed extremes x,y in (1.16) we have $\rho(x,y) \le c d_\rho(x,y)$ for all $x,y \in X$, so it follows that

$$d_\rho(x,y) \le \rho(x,y) \le d_\rho(x,y) \tag{1.17}$$

for all $x,y \in X$. We have that d_ρ is a semimetric on X by lemma 1.1.1 , and by (1.17) if $x \neq y$ then $d_\rho(x,y) \neq 0$ because ρ is a distance by hypothesis, i.e. d_ρ is a metric on X. We conclude that there exists a metric $d = d_\rho$ on X such that ρ is bi-lipschitz equivalent with respect to d. $\square$

If now we introduce for any nonempty set X the set

$$\mathcal{D} = \{d : X \times X \to [0,+\infty[\,,\ d \ \text{ is a metric}\}$$

and, for any distance ρ on X the set

$$\mathcal{D}_\rho = \Big\{d \in \mathcal{D}\,,\ \exists H \ge 1\ :\ \frac{1}{H}\,\rho(x,y) \le d(x,y) \le \rho(x,y),\ \forall x,y \in X\Big\},$$

then by theorem 1.2.2 immediately follows the next interesting proposition

Proposition 1.2.4. *Let (X,ρ) be a semimetric space. Then the following facts are equivalent:*

j) $\exists\, c \in [1,+\infty[$ *such that*

$$\rho(x,y) \le c \sum_{i=1}^{n} \rho(z_{i-1},z_i) \qquad \textit{(relaxed polygonal inequality)}$$

$\forall n \in \mathbb{N}\,,\ n \ge 1\,,\ \forall z_1,\dots,z_n \in X$ *with* $z_0 = x,\ z_n = y$;

jj) $d_\rho \in \mathcal{D}_\rho$, *where d_ρ is the Frink's map 1.12;*

jjj) $\mathcal{D}_\rho \neq \emptyset$.

From here on we collect main results and conditions, both for quasi-metrics and for quasi-ultrametrics, related to the existence of a bi-lipschitz equivalent metric in terms of the existence of a relaxed polygonal inequality. First we present a new and very simple proof of a classic result of Frink about the existence of a bi-lipschitz equivalent metric with respect to a given quasi-ultrametric in the special case of a constant $K = 2$ and extended later for $K \le 2$.

Theorem 1.2.3. *Let (X, ρ) be a quasi-ultrametric space with constant $K \leq 2$. Then*

$$\frac{1}{2K}\,\rho(x,y) \leq d_\rho(x,y) \leq \rho(x,y) \quad \forall x,y \in X.$$

Proof. According to proposition 1.2.4 we have to prove *j)* with constant $2K$ and, to this aim, we prove by induction on n that for any given sequence of $n+1$ points $x_0 = x, x_1, \ldots, x_{n-1}, x_n = y \in X$, $n \in \mathbb{N}$, $n \geq 1$ it results

$$\rho(x,y) \leq K\left(\rho\left(x,x_1\right) + 2\sum_{i=2}^{n-1}\rho\left(x_{i-1},x_i\right) + \rho\left(x_{n-1},y\right)\right), \tag{1.18}$$

where the sum is absent for $n = 1,2$. Inequality (1.18) is trivial for $n = 1,2$, so given $n \geq 3$ let us assume the formula true for any sequence of points $x = x_0, x_1, \ldots, x_{m-1}, x_m = y \in X$ with $1 \leq m < n$. If $\frac{\rho(x,y)}{K} \leq \rho(x_{n-1},y)$ then (1.18) is true. Observing that thanks to hypothesis for any $x,y,z \in X$ the inequalities

$$\rho(x,z) < \frac{\rho(x,y)}{K} \quad \text{and} \quad \rho(y,z) < \frac{\rho(x,y)}{K} \tag{1.19}$$

cannot occur simultaneously, if $\rho(x_{n-1},y) < \frac{\rho(x,y)}{K}$ then it follows $\frac{\rho(x,y)}{K} \leq \rho(x,x_{n-1})$, so there exists $m \in \mathbb{N}$ such that $1 \leq m \leq n-1$ and

$$\rho(x,x_m) < \frac{\rho(x,y)}{K} \leq \rho(x,x_{m+1});$$

in the same way by (1.19) and by induction we have

$$\frac{\rho(x,y)}{K} \leq \rho(x_m,y) \leq K\left(\rho\left(x_m,x_{m+1}\right) + 2\sum_{i=m+2}^{n-1}\rho\left(x_{i-1},x_i\right) + \rho\left(x_{n-1},y\right)\right),$$

which added with the

$$\frac{\rho(x,y)}{K} \leq \rho(x,x_{m+1}) \leq K\left(\rho\left(x,x_1\right) + 2\sum_{i=2}^{m}\rho\left(x_{i-1},x_i\right) + \rho\left(x_m,x_{m+1}\right)\right)$$

implies

$$\frac{2\rho(x,y)}{K} \leq K\left(\rho\left(x,x_1\right) + 2\sum_{i=2}^{n-1}\rho\left(x_{i-1},x_i\right) + \rho\left(x_{n-1},y\right)\right).$$

The thesis follows noting that $\frac{K^2}{2} \leq K$ if $K \leq 2$. $\square$

Previous theorem is no more true when $K > 2$. In [4] the author exhibits, for any $K > 2$, a non trivial example of a space (X, ρ) for which theorem 1.2.3 does not work, more precisely showing that for every $\varepsilon > 0$ there exists a quasi-metric ρ with constant $A = 1 + \varepsilon$ such that the Frink's map is not a metric. Other example for the limits of Frink's method was given [31]. Actually we propose a very simple and general example of these limits. As shown in Example 1.1.23, the quasi-ultrametric $\rho \equiv d^\alpha$, with constant 2^α, does not satisfy a relaxed polygonal inequality for any $\alpha \in]1, +\infty[$ so, for any fixed $K > 1$ assuming $\alpha = \log_2 K$ we have a quasi-ultrametric with a prefixed constant that does not satisfy the relaxed polygonal inequality. The same is valid in the case of quasi-metric space, so we have the following fact.

Proposition 1.2.5. *For any fixed real constant $K > 2$ ($A > 2$) quasi-ultrametric (quasi-metric) spaces exists such that the induced Frink's map is not a metric.*

The failure, in some case, for d_ρ to be a metric yields immediately the following known equivalences [5], which adapt proposition 1.2.4 to the case of quasi-ultrametrics.

Theorem 1.2.4. *Let X be a nonempty set and $K \in \mathbb{R}$. The following facts are equivalent:*

a) $K \leq 2$;

b) for any quasi-ultrametric ρ on X with constant K, then ρ satisfies the relaxed polygonal inequality;

c) for any quasi-ultrametric ρ on X with constant K, then $d_\rho \in \mathcal{D}_\rho$;

d) for any quasi-ultrametric ρ on X with constant K, then $\mathcal{D}_\rho \neq \emptyset$;

e) for any quasi-ultrametric ρ on X with constant K, then $d_\rho \in \mathcal{D}$;

Proof. *a)* $\Longrightarrow$ *b)* This is the first part of the proof of theorem 1.2.3; *b)* $\Longrightarrow$ *c)* this is *j)* $\Longrightarrow$ *jj)* of proposition 1.2.4; *c* $\Longrightarrow$ *d)* trivial; *d)* $\Longrightarrow$ *e)* let ρ be a quasi-ultrametric on X with constant K, moreover let $d \in \mathcal{D}_\rho$ and, finally, let $x, y \in X$, $n \in \mathbb{N}$, $n \geq 1$, and $x_0 = x, x_1, \ldots, x_{n-1}, x_n = y \in X$: for some $H > 0$ it follows that

$$\rho(x,y) \leq H\,d(x,y) \leq H\sum_{i=1}^{n} d(x_{i-1},x_i) \leq H\sum_{i=1}^{n} \rho(x_{i-1},x_i);$$

then the relaxed polygonal inequality is satisfied and the thesis follows by *j)* $\Longrightarrow$ *jj)* of proposition 1.2.4 ; *e)* $\Longrightarrow$ *a)* let ρ be a given quasi-ultrametric on X with constant K: proposition 1.2.5 implies the thesis immediately. $\square$

We conclude this subsection with some remarks on the constant in $\mathcal{D}_\rho$, for a given quasi-ultrametric ρ on a nonempty set X. More precisely, in [5] the authors show that a metric with the better constant K^2 exists in the equivalence relation of theorem 1.2.3 . They also show that for any function $\varphi : [1,2] \to [0,+\infty[$ such that for any $K \in [1,2]$ and any quasi-ultrametric space (X,ρ) with constant K and any $d \in \mathcal{D}_\rho$, such that

$$\frac{1}{\varphi(K)}\,\rho(x,y) \leq d(x,y) \leq \rho(x,y) \quad \forall x,y \in X,$$

then $\varphi(K) \geq K^2$. Nevertheless for a given space the constant K^2 may not be the best possible, as shown in the following proposition and example.

Proposition 1.2.6. *Let (X,ρ) be a quasi-ultrametric space with* $\operatorname{card}(X) < +\infty$. *Then $\mathcal{D}_\rho \neq \emptyset$.*

Proof. Setting $S = \rho\left((X \times X) \setminus \Delta(X)\right)$, $m = \min S$, $M = \max S$ and $H = \frac{M}{m}$, the position $d = m\,\bar{d}$, with $\bar{d}$ the discrete metric on X, obtain us the requested metric on X. We note also that proposition immediately follows by theorem 1.1.2, i.e. if X is a finite set then it also is bounded from below and above. $\square$

Example 1.2.5. *Let $X = \{x_1,x_2,x_3\}$ and $\rho : X \times X \to [0,+\infty[$ assumed to be symmetric and zero only on the diagonal, defined as follows: $\rho(x_1,x_2) = a$, $\rho(x_1,x_3) = ab$ and $\rho(x_2,x_3) = Kab$, with $a > 0$, $b,K \geq 1$, suitable chosen in the sequel. The map ρ is a quasi-ultrametric on X with constant K but if we choose $K > 1$ and $b > \frac{1}{K-1}$ then $\rho(x_2,x_3) > \rho(x_1,x_2) + \rho(x_1,x_3)$, i.e. ρ is not a metric. By proposition 1.2.6 there exists a metric d on X equivalent to ρ with constant $H = Kb$ and it is $H < K^2$ if $\frac{1}{K-1} < K$, i.e. $K \in]\varphi,2[$, with $\varphi = \frac{\sqrt{5}+1}{2}$ the golden ratio.*

A particular class of bi-lipschitz equivalent metrics, in a suitable sense, can also be obtained through a special type of function $\varphi : [0,+\infty[\to [0,+\infty[$, more precisely the ones satisfying the so called Δ_2 condition.

Definition 1.2.2 (Δ_2 condition). *A function $\varphi : [0, +\infty[\to [0, +\infty[$ satisfies condition Δ_2 with a constant D if $\varphi(2t) \leq D\varphi(t)$ for all $t \geq 0$.*

We note that if φ is a non-decreasing function satisfying condition Δ_2 with a constant D then by definition we have $D \geq 1$.

Theorem 1.2.5. *Let (X, ρ) be a quasi-ultrametric space with a constant $K \geq 1$ and φ a non-decreasing function satisfying condition Δ_2 with a constant $D \geq 1$ and such that $\varphi(t) = 0$ if and only if $t = 0$. If $D \leqslant 2^{\frac{1}{\lceil \log_2 K \rceil}}$ then $\varphi(\rho)$ is a distance on X satisfying a relaxed polygonal inequality.*

Proof. It is easy to check that $\varphi(\rho)$ is a distance on X. For all $x, y, z \in X$ we have $\rho(x, y) \leqslant K \max\{\rho(x, z), \rho(z, y)\}$ because ρ is a quasi-ultrametric on X with a constant K, and by $K = 2^{\log_2 K} \leqslant 2^{\lceil \log_2 K \rceil}$ we have $\rho(x, y) \leqslant 2^{\lceil \log_2 K \rceil} \max\{\rho(x, z), \rho(z, y)\}$. By properties of φ it follows

$$\begin{aligned} \varphi(\rho(x,y)) &\leqslant \varphi(2^{\lceil \log_2 K \rceil} \max\{\rho(x,z), \rho(z,y)\}) \\ &\leqslant D^{\lceil \log_2 K \rceil} \max\{\varphi(\rho(x,z)), \varphi(\rho(z,y))\}, \end{aligned}$$

i.e. $\varphi(\rho)$ is a quasi-ultrametric on X with a constant $\bar{K} = D^{\lceil \log_2 K \rceil}$. By $D \leqslant 2^{\frac{1}{\lceil \log_2 K \rceil}}$ then $\bar{K} \leq 2$, thus metrization theorems on quasi-ultrametric spaces for this special value of the constant give that $\varphi(\rho)$ satisfies a relaxed polygonal inequality with a constant c depending on K. $\square$

Remark 1.2.4. The function $\varphi(t) = t^\alpha$ with $\alpha > 0$ is a non-decreasing function satisfying condition Δ_2 with a constant $D = 2^\alpha$ and such that $\varphi(t) = 0$ if and only if $t = 0$. Then the condition $D \leqslant 2^{\frac{1}{\lceil \log_2 K \rceil}}$ is satisfied by assuming $\alpha \leq \frac{1}{\lceil \log_2 K \rceil}$, so recovering also bi-lipschitz equivalent metrics with respect to the α-power of a general quasi-ultrametric.

Up to now we have collected main results and ideas on the existence of a bi-lipschitz equivalent metric with respect to a suitable quasi-ultrametric ρ, using standard techniques to prove the validity of the relaxed polygonal inequality. In what follows we exhibit analogous results adapting topological techniques involving uniformity-type families generated by the given distance ρ; more precisely, in Theorem 1.2.6 we prove the existence of a bi-lipschitz equivalent distance f to a given quasi-ultrametric ρ, with f satisfying the relaxed polygonal inequality: it comes to be that both d_ρ and d_f are bi-lipschitz equivalent metrics to ρ. First let us introduce some notations.

Le X be a set and let A and B subsets of $X \times X$, i.e. relations on the set X. The inverse relation of A is $-A = \{(x,y) : (y,x) \in A\}$; then a relation A is symmetric if $A = -A$ holds. The composition of A and B is denoted by $A + B$ and we have

$$A + B = \{(x,y) : \exists z \in X : (x,z) \in A, (z,y) \in B\};$$

the composition is associative, i.e. $(A+B)+C = A+(B+C)$, but it is not commutative, i.e., generally $A + B \neq B + A$. Finally, for $U \subseteq X \times X$ and a natural number $n \geq 1$ the set $nU \subseteq X \times X$ is defined inductively by the formulas $1U = U$ and $nU = (n-1)U + U$.

Now we can start with the following lemma.

Lemma 1.2.1. *Let X be a nonempty set and $\{U_i\}_{i\in\mathbb{Z}} \subseteq X \times X$ such that*

- *U_i is symmetric $\forall i \in \mathbb{Z}$;*
- *$\Delta(X) \subseteq U_i \quad \forall i \in \mathbb{Z}$;*
- *$3\, U_{i+1} \subseteq U_i \quad \forall i \in \mathbb{Z}$;*
- *$\bigcup_{i\in\mathbb{Z}} U_i = X \times X$.*

Then, for any $L \in]1,2]$, the function $f : X \times X \to [0,+\infty[$ defined by the position

$$f(x,y) = \begin{cases} 0 & if \quad (x,y) \in \bigcap_{i\in\mathbb{Z}} U_i \\ \frac{1}{L^i} & if \quad (x,y) \in U_{i-1} \setminus U_i \end{cases}$$

is a semi-distance on X satisfying a relaxed polygonal inequality with constant L.

Proof. Let us verify that f satisfies a relaxed polygonal inequality with constant L, i.e. $f(x,y) \leq L \sum_{i=1}^{n} f(x_{i-1},x_i)$ for $x_0 = x, x_1, ..., x_n = y \in X$. For $n = 1$ the thesis is trivial. Let us assume the thesis true for some $n \geq 1$, and let $a = \sum_{i=1}^{n+1} f(x_{i-1},x_i)$. If $a = 0$ it trivially follows that $f(x,y) = 0$. If $a > 0$, up to a rearrangement of the chain, let m be the largest integer such that $\sum_{i=1}^{m} f(x_{i-1},x_i) \leqslant \frac{a}{2}$, then $\sum_{i=m+2}^{n+1} f(x_{i-1},x_i) < \frac{a}{2}$; obviously $f(x_m,x_{m+1}) \leqslant a$, moreover, by inductive hypotheses it follows that $f(x_0,x_m) \leqslant \frac{L}{2}a$ and

$f(x_{m+1}, x_{n+1}) < \frac{L}{2}a$; if s is the least integer involved in the definition of the above values of f, it results

$$\frac{1}{L^s} \leqslant \max\left\{\frac{L}{2}a, a\right\}\ ; \tag{1.20}$$

so that $f(x_0, x_m), f(x_m, x_{m+1}), f(x_{m+1}, x_{n+1}) \leqslant \frac{1}{L^s}$ and consequently

$$(x_0, x_m), (x_m, x_{m+1}), (x_{m+1}, x_{n+1}) \in U_{s-1},$$

and $(x_0, x_{n+1}) \in 3\,U_{s-1} \subseteq U_{s-2}$, i.e. $f(x_0, x_{n+1}) \leqslant \frac{1}{L^{s-1}} = L\frac{1}{L^s}$. By (1.20) it follows that

$$f(x_0, x_{n+1}) \leqslant L \max\left\{\frac{L}{2}, 1\right\} a :$$

if $\max\left\{\frac{L}{2}, 1\right\} \leqslant 1$, i.e. $1 < L \leq 2$, the thesis holds. $\square$

Theorem 1.2.6. *Let (X, ρ) be a quasi-ultrametric space with constant $K \in]1, \sqrt{2}]$. Then ρ satisfies a relaxed polygonal inequality with constant $C = K^4$ and there exist more than one metric $d \equiv d(K)$ on X such that for any $x, y \in X$,*

$$\frac{1}{K^4}\, \rho(x, y) \leq d(x, y) \leq \rho(x, y).$$

Proof. Let $L = K^2$. For any $i \in \mathbb{Z}$ let $\{U_i\}_{i \in \mathbb{Z}} \subseteq X \times X$ defined as $U_i = \left\{(x, y) : \rho(x, y) < K^{-2i}\right\}$: it is easy to check that the family $\{U_i\}_{i \in \mathbb{Z}}$ satisfies the hypotheses of Lemma 1.2.1, so the related function $f : X \times X \to [0, +\infty[$ satisfies a relaxed polygonal inequality with constant L. Then, if $\rho(x, y) = 0$, i.e. $(x, y) \in \bigcap_{i \in \mathbb{Z}} U_i$, the thesis is trivially true; otherwise there exists $i \in \mathbb{Z}$ such that $(x, y) \in U_{i-1} \backslash U_i$, so $\frac{1}{K^{2i}} \leq \rho(x, y) < \frac{2}{K^{2i}}$, i.e. $f(x, y) \leqslant \rho(x, y) < K^2 f(x, y)$ for any $x, y \in X$, and ρ satisfies a relaxed polygonal inequality with constant $C = K^4$: then the semi-metrics $d : X \times X \to [0, +\infty[$, where $d = d_f, d_\rho$, are defined by the chain approach, are metrics on X such that $d(x, y) \leq \rho(x, y) \leq K^4\, d(x, y)$, i.e. thesis. $\square$

We observe that weakening the hypotheses of Lemma 1.2.1 the semi-distance f satisfies the weaker condition *v)* as shown in the following lemma.

Lemma 1.2.2. *Let X be a nonempty set and $\{U_i\}_{i \in \mathbb{Z}} \subseteq X \times X$ such that*

- *U_i is symmetric $\forall i \in \mathbb{Z}$;*
- *$\Delta(X) \subseteq U_i \quad \forall i \in \mathbb{Z}$;*

- $2\,U_{i+1} \subseteq U_i \quad \forall i \in \mathbb{Z}$;
- $\bigcup_{i\in\mathbb{Z}} U_i = X \times X$.

Then, for any $L > 1$, the function $f : X \times X \to [0, +\infty[$ defined by the position

$$f(x,y) = \begin{cases} 0 \quad if \quad (x,y) \in \bigcap_{i\in\mathbb{Z}} U_i \\ \frac{1}{L^i} \quad if \quad (x,y) \in U_{i-1} \setminus U_i \end{cases}$$

is a semi-distance on X satisfying v) with constant L.

Proof. Let $x, y, z \in X$. If $f(x,z) \neq 0$ and $f(z,y) \neq 0$ by definition of f there exist integers k, h, for instance $k \leqslant h$, such that $f(x,z) = \frac{1}{L^k}$ and $f(z,y) = \frac{1}{L^h}$, so $(x,z) \in U_{k-1}$, $(z,y) \in U_{h-1} \subseteq U_{k-1}$, and $(x,y) \in 2\,U_{k-1} \subseteq U_{k-2}$ so that $f(x,y) \leqslant \frac{1}{L^{k-1}} = L\frac{1}{L^k} = L\,f(x,z)$, i.e. $f(x,y) \leqslant L \max\{f(x,z), f(z,y)\}$. If $f(x,z) = f(z,y) = 0$ by the definition of f we have $(x,z), (z,y) \in U_i$ for all $i \in \mathbb{Z}$, so $(x,z), (z,y) \in U_{i+1}$ and $(x,y) \in 2\,U_{i+1} \subseteq U_i$ for all $i \in \mathbb{Z}$, i.e. $f(x,y) = 0$ and $f(x,y) \leqslant L \max\{f(x,z), f(z,y)\}$. If $f(x,z) \neq 0$ and $f(z,y) = 0$ then there exists an integer k such that $f(x,z) = \frac{1}{L^k}$ with $(x,z) \in U_{k-1}$ and $(z,y) \in U_i$ for all $i \in \mathbb{Z}$, so also $(z,y) \in U_{k-1}$; then $(x,y) \in 2\,U_{k-1} \subseteq U_{k-2}$ and $f(x,y) \leqslant \frac{1}{L^{k-1}} = L\frac{1}{L^k} = L\,f(x,z)$ and again $f(x,y) \leqslant L \max\{f(x,z), f(z,y)\}$, that concludes the proof. □

As a byproduct of Lemma 1.2.2 we obtain directly a bi-lipschitz equivalent metric with respect to a given suitable distance, as in the original paper of Frink (see [3]), with the same constant equal to 4. More precisely let (F) the following condition:

(F) for every $\varepsilon > 0$, if $\rho(x,z) < \varepsilon$ and $\rho(z,y) < \varepsilon$, then $\rho(x,y) < 2\varepsilon$.

Then arguing by contradiction Frink obtains the existence of a bi-lipschitz equivalent metric d_ρ with respect to the distance ρ; indeed, by applying Lemma 1.2.2 to the distance ρ with $U_i = \{(x,y) : \rho(x,y) < 2^{-i}\}$, $i \in \mathbb{Z}$, arguing as in Theorem 1.2.6, the semi-distance f comes to be bi-lipschitz equivalent to ρ, as well as the metric d_f, with the same constant 4. Clearly, any quasi-ultrametric ρ with constant $K \leq 2$ satisfies the (F) condition but, in general, Lemma 1.2.2 allows us to generate a bi-lipschitz equivalent quasi-ultrametric with respect to a given quasi-ultrametric, maintaining the same constant, arguing as before with the sets U_i defined as $U_i = \{(x,y) : \rho(x,y) < K^{-i}\}$, Lemma

1.2.2 provides a quasi-ultrametric f such that, for all $x, y \in X$,

$$\frac{1}{K}\rho(x,y) \leq f(x,y) \leq \rho(x,y).$$

The following proposition concludes the overview concerning the existence of a bi-lipschitz equivalent metric with respect to a given quasi-ultrametric: an interesting characterization of a "n-angular" inequality is shown.

Proposition 1.2.7. *Let X be a nonempty set, $L \in]1, +\infty[$ and let $n \geq 2$ an integer. Then the following facts are equivalent:*

l) *There exists a semi-distance ρ on X such that for any $x_0 = x, x_1, ..., x_n = y \in X$ it results*

$$\rho(x,y) \leqslant L \max\{\rho(x,x_1), \ldots, \rho(x_{n-1},y)\};$$

ll) *There exists $\{U_i\}_{i\in\mathbb{Z}} \subseteq X \times X$ such that*

- *U_i is symmetric $\forall i \in \mathbb{Z}$;*
- *$\Delta(X) \subseteq U_i \quad \forall i \in \mathbb{Z}$;*
- *$n\, U_{i+1} \subseteq U_i \quad \forall i \in \mathbb{Z}$;*
- *$\bigcup_{i\in\mathbb{Z}} U_i = X \times X$.*

Proof. l) $\Longrightarrow$ ll) Let $\{U_i\}_{i\in\mathbb{Z}}$ be a collection of subsets of $X \times X$ defined as

$$U_i = \left\{(x,y) : \rho(x,y) < \frac{1}{L^i}\right\};$$

then, it is easy to check that U_i satisfy all condition of *ll);*

ll) $\Longrightarrow$ l) let $\rho : X \times X \to [0, +\infty[$ be a function defined as in Lemma 1.2.2: then, with similar arguments, the thesis holds. $\square$

In conclusion, we want to emphasize that in the existence of a bi-lipschitz equivalent metric with respect to a given quasi-ultrametric, the condition $K \leq 2$ seems to play a crucial role, probably due to deeper geometric reasons. Also, the chain approach technique seems to be particularly relevant, probably because of the strict relationship with the relaxed polygonal inequality condition. Moreover it is not so clear what kind of relationship holds between d_f in Theorem 1.2.6 and the original Frink's metric d_ρ; furthermore, it is not clear if a given distance ρ satisfying the condition (F), jointly with the relationship

$3\,U_{i+1} \subseteq U_i$ for the sets $U_i = \{(x,y) : \rho(x,y) < 2^{-i}\}$, $i \in \mathbb{Z}$, is actually a quasi-ultrametric for some constant $K \leq 2$; finally, if ρ is a distance on X satisfying a n-angular inequality like in Proposition 1.2.7, it is not that obvious if ρ satisfies a more general version of inequality (1.18) for particular values of the constant L.

1.3 Equivalent metrics on quasi-metric spaces

It is clear that not every quasi-metric with constant A is a quasi-ultrametric with constant A, so general results for quasi-ultrametric do not hold for a given quasi-metric.

Remark 1.3.1. No analogous of theorem 1.2.3 can be proved for the general case of a set X endowed with a quasi-metric ρ with constant $A \in]1, +\infty[$: in Example 1.1.23 the quasi-metric $\rho \equiv d^\alpha$, with constant $A = 2^{\alpha-1}$, do not satisfy the relaxed polygonal inequality for any $\alpha \in]1, +\infty[$.

In the case of a quasi-metric space (X, ρ) it is however possible to prove the existence of a metric which is bi-lipschitz equivalent metric with respect to the power of a suitable power of ρ. A very simple proof of this fact is easily obtained in the following proposition.

Proposition 1.3.1. *Let be (X, ρ) a quasi-metric space with constant A. Then for any $\alpha \in \left]0, \frac{1}{1+\log_2 A}\right]$ it is*

$$\frac{1}{2(2A)^\alpha}\,\rho^\alpha(x,y) \leqslant d_{\rho^\alpha}(x,y) \leqslant \rho^\alpha(x,y) \quad \forall x, y \in X. \tag{1.21}$$

Proof. Let $\alpha \in \left]0, \frac{1}{1+\log_2 A}\right]$, then ρ^α is a quasi-ultrametric with constant $(2A)^\alpha$ thanks to proposition 1.1.6 : so the thesis follows immediately by theorem 1.18 . □

For the sake of completeness we show the classical result about the existence of a bi-lipschitz equivalent metric with respect to the power of a given quasi-metric. Theorem 1.3.1, which, in its turn, relies on Lemma 1.3.1, of which we exhibit two proofs [3, 7, 4, 5, 32, 26, 33, 28], based on the validity of a relaxed polygonal inequality.

Lemma 1.3.1. *Let (X, ρ) be a quasi-metric space with constant A. Then there exists a non-increasing continuous function $\alpha = \alpha(k)$, $]1, +\infty[\ni k \to \alpha(k) \in]0, 1[$,* *such that*

$$\rho^\alpha(x,y) \leqslant k \sum_{i=1}^{n} d_{\rho^\alpha}(x_{i-1},x_i)\,, \tag{1.22}$$

for any $x_0 = x, x_1, \ldots, x_{n-1}, x_n = y \in X$.

First Proof. Inequality 1.22 is easily true when $n = 1$ for any $k > 1$. By induction, for $n \geq 2$, let us assume 1.22 true for $x = x_0, x_1, \ldots, x_{m-1}, x_m = y \in X$, with $m < n$. Then, for $x \neq y$, if $a = \sum_{i=1}^{n} \rho(x_{i-1},x_i)^\alpha > 0$, it is either $\rho(x_0,x_1)^\alpha \leq \frac{a}{2}$ or $\rho(x_{n-1},x_n)^\alpha \leq \frac{a}{2}$. Let us suppose that $\rho(x_0,x_1)^\alpha \leq \frac{a}{2}$ (the other case is similar): let j be the greatest index such that $\sum_{i=1}^{j} \rho(x_{i-1},x_i)^\alpha \leq \frac{a}{2}$. Clearly $j < n$ so that $\sum_{i=1}^{j+1} \rho(x_{i-1},x_i)^\alpha > \frac{a}{2}$. Suppose that $j+1 < n$; in this case it is $\sum\limits_{i=j+2}^{n} \rho(x_{i-1},x_i)^\alpha < \frac{a}{2}$. Thanks to the inductive hypothesis we therefore have: $\rho(x_0,x_j)^\alpha \leq k\frac{a}{2}$, $\rho(x_{j+1},x_n)^\alpha < k\frac{a}{2}$ and obviously $\rho(x_j,x_{j+1})^\alpha \leq a$. Furthermore, the previous inequalities continue to be true even when $j+1 = n$. Given such a j by the inequality

$$\rho\left(x,y\right) \leq A\rho\left(x_0,x_j\right) + A\rho\left(x_j,x_n\right)$$

it follows that either

$$\rho\left(x,y\right) \leq 2A\rho\left(x_0,x_j\right) \quad \text{or} \quad \rho\left(x,y\right) \leq 2A\rho\left(x_j,x_n\right).$$

Then, if $\rho\left(x,y\right) \leq 2A\rho\left(x_0,x_j\right)$ it is $\rho(x,y)^\alpha \leq \frac{(2A)^\alpha}{2}ka$, so (1.22) follows by choosing $\alpha \leq \frac{1}{\log_2(2A)}$, for any $k > 1$; while if $\rho\left(x,y\right) \leq 2A\rho\left(x_j,x_n\right)$, if $j+1 = n$, $\rho\left(x_j,x_{j+1}\right) \leq a^{\frac{1}{\alpha}}$ implies $\rho(x,y)^\alpha \leq (2A)^\alpha a$ and (1.22) follows by choosing $\alpha \leq \frac{\log_2 k}{\log_2(2A)}$, for any $k > 1$. If $\rho\left(x,y\right) \leq 2A\rho\left(x_j,x_n\right)$ with $j+1 < n$ by the inequality

$$\rho\left(x_j,x_n\right) \leq A\rho\left(x_j,x_{j+1}\right) + A\rho\left(x_{j+1},x_n\right),$$

it still follows that either

$$\rho\left(x_j,x_n\right) \leq 2A\rho\left(x_j,x_{j+1}\right) \quad \text{or} \quad \rho\left(x_j,x_n\right) \leq 2A\rho\left(x_{j+1},x_n\right).$$

By $\rho\left(x_j,x_n\right) \leq 2A\rho\left(x_j,x_{j+1}\right)$ we have $\rho\left(x,y\right) \leq (2A)^2\rho\left(x_j,x_{j+1}\right)$ and by $\rho\left(x_j,x_{j+1}\right) \leq a^{\frac{1}{\alpha}}$ we have $\rho(x,y)^\alpha \leq (2A)^{2\alpha}a$, so (1.22) follows by choosing $\alpha \leq \frac{\log_2 k}{2\log_2(2A)}$, for any $k > 1$. If $\rho\left(x_j,x_n\right) \leq 2A\rho\left(x_{j+1},x_n\right)$, by $\rho\left(x_{j+1},x_n\right) < \left(k\frac{a}{2}\right)^{\frac{1}{\alpha}}$ we have that $\rho(x,y)^\alpha < \frac{(2A)^{2\alpha}}{2}ka$, and (1.22) follows by choosing $\alpha \leq$

$\frac{1}{2\log_2(2A)}$, for any $k > 1$. Finally noting that

$$\min\left\{\frac{1}{2\log_2(2A)}, \frac{1}{\log_2(2A)}, \frac{\log_2 k}{2\log_2(2A)}, \frac{\log_2 k}{\log_2(2A)}\right\} =$$

$$= \begin{cases} \frac{\log_2 k}{2\log_2(2A)} & \text{if } k \in]1,2] \\ \frac{1}{2\log_2(2A)} & \text{if } k \in [2,+\infty[\end{cases},$$

we conclude that the inequality (1.22) is verified for any $h > 1$ as long as $\alpha(k)$ is chosen as follow

$$\alpha(k) = \begin{cases} \frac{\log_2 k}{2\log_2(2A)} & k \in]1,2] \\ \\ \frac{1}{2\log_2(2A)} & k \in [2,+\infty[\end{cases}.$$

Second Proof. It suffices to verify that fixed $n \geq 2$, for any sequence of $n+1$ points $x_0 = x, x_1, \ldots, x_{n-1}, x_n = y \in X$ and for any $h \in]1,+\infty[$, there exists $\alpha \in]0,1[$ such that

$$\rho(x,y)^\alpha \leq h\left(\rho(x,x_1)^\alpha + 2\sum_{i=2}^{n-1}\rho(x_{i-1},x_i)^\alpha + \rho(x_{n-1},y)^\alpha\right), \tag{1.23}$$

for all $x, y \in X$, where if $n = 2$ the sum in the right-hand side of the inequality is null; more precisely it follows that for any $h \in]1,+\infty[$ there exists $\alpha = \alpha(h) \in]0,1[$ such that

$$\rho(x,y)^\alpha \leq 2h\sum_{i=1}^{n}\rho(x_{i-1},x_i)^\alpha,$$

and the thesis is obtained by setting $k = 2h$. So, for $n = 2$ (1.23) becomes $\rho(x,y)^\alpha \leq h\left(\rho(x,x_1)^\alpha + \rho(x_1,y)^\alpha\right)$ and the thesis follows by remark 1.1.2 and triangular inequality

$$\rho(x,y) \leq A\left(\rho(x,x_1) + \rho(x_1,y)\right),$$

with $\alpha \leq \frac{\log_2 h}{\log_2(2A)}$. By induction, if $n > 2$, let the inequality (1.23) be true for any sequence of points $x_0 = x, x_1, \ldots, x_{m-1}, x_m = y \in X$, with $m < n$. Let j be the maximum index such that $\rho(x,y)^\alpha \leq (2A)^\alpha \rho(x_j,y)^\alpha$, so $\rho(x,y)^\alpha > (2A)^\alpha \rho(x_{j+1},y)^\alpha$. In correspondence of x_{j+1} the triangular inequality

$$\rho(x,y) \leq A\left(\rho(x,x_{j+1}) + \rho(x_{j+1},y)\right),$$

yields

$$\rho(x,y)^\alpha \leq (2A)^\alpha \max\left\{\rho\left(x,x_{j+1}\right)^\alpha, \rho\left(x_{j+1},y\right)^\alpha\right\},$$

so that $\rho(x,y)^\alpha \leq (2A)^\alpha \rho\left(x,x_{j+1}\right)^\alpha$. Therefore, by adding we obtain

$$\rho(x,y)^\alpha \leq \frac{(2A)^\alpha}{2}\left(\rho\left(x,x_{j+1}\right)^\alpha + \rho\left(x_j,y\right)^\alpha\right),$$

and the inductive hypothesis applied to the sequences $x_0 = x, x_1, \ldots, x_{m+1}$ and $x_m, x_{m+1}, \ldots, x_{n+1} = y$ ensures

$$\rho(x,y)^\alpha \leq \frac{(2A)^\alpha}{2} h\left(\rho(x,x_1)^\alpha + 2\sum_{i=2}^{n-1}\rho(x_{i-1},x_i)^\alpha + \rho(x_{n-1},y)^\alpha\right),$$

by which (1.23) follows for any $h > 1$ by choosing $\alpha \leq \frac{1}{\log_2(2A)}$.

Finally, noting that

$$\min\left\{\frac{1}{\log_2(2A)}, \frac{\log_2 h}{\log_2(2A)}\right\} = \begin{cases} \frac{\log_2 h}{\log_2(2A)} & \text{if } h \in]1,2] \\ \frac{1}{\log_2(2A)} & \text{if } h \in [2,+\infty[\end{cases},$$

we conclude that the inequality (1.23) is verified for any $h > 1$ as long as $\alpha(h)$ is chosen as follow

$$\alpha(h) = \begin{cases} \frac{\log_2 h}{\log_2(2A)} & h \in]1,2] \\ \\ \frac{1}{\log_2(2A)} & h \in [2,+\infty[\end{cases}, \tag{1.24}$$

concluding so the proof of the lemma. □

Theorem 1.3.1. *Let (X,ρ) be a quasi-metric space with constant A. Then there exist a non-increasing continuous function $\alpha = \alpha(k)$,*

$$]1,+\infty[\ \ni k \to \alpha(k) \in]0,1[$$

such that

$$\frac{1}{k}\rho^\alpha(x,y) \leqslant d_{\rho^\alpha}(x,y) \leqslant \rho^\alpha(x,y) \quad \forall x,y \in X. \tag{1.25}$$

Proof. Let $k > 1$: then lemma 1.3.1 yields the existence of a non-increasing continuous function $\alpha :]1,+\infty[\to]0,1[$, $\alpha = \alpha(k)$, such that (1.22) holds; so the right-hand side in (1.25) is apparent by the very definition of d_{ρ^α} and, finally, the thesis immediately follows from proposition 1.2.4. $\square$

Remark 1.3.2. It is known that the position $\rho' = d_{\rho^\alpha}{}^{\frac{1}{\alpha}}$ where d_{ρ^α} is the metric in (1.25) defines a continuous quasi-metric on X.

Proposition 1.3.2. *Let (X,ρ) be a quasi-metric space with constant A. Then the quasi-metric ρ' in the remark 1.3.2 satisfies the inequality*

$$\left|\rho'(x,y) - \rho'(y,z)\right| \leqslant \frac{1}{\alpha} \max\left\{\rho'(x,z), \rho'(y,z)\right\}^{1-\alpha} \rho'(x,y)^\alpha \tag{1.26}$$

for all $x, y, z \in X$. Furthermore for all $c \in X$, $r > 0$ the balls $B_{\rho'}(c,r)$ are open sets in the topology equivalently induced by ρ', ρ or d.

Proof. By lemma 1.1.6 and according to what already proved ρ' is a quasi-metric on X that ρ induces so the topology τ_ρ. Then, fixed $x, y, z \in X$, for all $r \geq \max\left\{\rho'(x,z), \rho'(z,y)\right\}$ by classical Lagrange's theorem we have

$$\begin{aligned}
|\rho'(x,z) - \rho'(z,y)| &= |d(x,z)^{\frac{1}{\alpha}} - d(z,y)^{\frac{1}{\alpha}}| \\
&\leq \frac{1}{\alpha} r^{1-\alpha} \, |d(x,z) - d(z,y)| \\
&\leq \frac{1}{\alpha} r^{1-\alpha} \, d(x,y) \\
&= \frac{1}{\alpha} r^{1-\alpha} \, \rho'(x,y)^\alpha,
\end{aligned}$$

i.e. the (1.26). Finally by this inequality it is easy to check that for all $c \in X$, $r > 0$ the balls $B(c,r)$ are open in the topology τ_ρ. Indeed if $B_{\rho'}(x,r)$ is a ball, we have to show that for every $y \in B_{\rho'}(x,r)$ there exists $r' > 0$ such that $B_{\rho'}(y,r') \subseteq B_{\rho'}(x,r)$, that is,

$$\rho'(y,z) < r' \Rightarrow \rho'(x,z) < r,$$

for any $z \in X$. Supposing that ρ' satisfies an relaxed triangle inequality for some constant $A' \geq 1$, choose first $0 < r' < r$. Then we have

$$\rho'(x,z) \leq A'(\rho'(x,y) + \rho'(y,z)) < 2A'r,$$

and by (1.26),

$$\begin{aligned}\rho'(x,z) &\leqslant \rho'(x,y) + \left|\rho'(x,z) - \rho'(x,y)\right| \\ &< \rho'(x,y) + \frac{1}{\alpha}\max\left\{\rho'(x,z), \rho'(x,y)\right\}^{1-\alpha}\rho'(y,z)^{\alpha} \\ &< \rho'(x,y) + \frac{1}{\alpha}(2A'r)^{1-\alpha}(r')^{\alpha}.\end{aligned}$$

Choosing $0 < r' < r$ such that $\frac{1}{\alpha}(2A'r)^{1-\alpha}(r')^{\alpha} < r - \rho'(x,y)$ then $\rho'(x,z) < r$ holds, concluding the proof. $\square$

Chapter 2

A brief glimpse into doubling spaces and applications

2.1 Homogeneity and geometric doubling properties

A first organic exposition of quasi-metric spaces, and in particular of so called *spaces of homogeneous type* appears in [7]; we stress that a particular class of such spaces are the ones endowed with a doubling measure. In this monograph the authors present these spaces as a natural setting for standard real analysis, i.e. covering theorems, Lebesgue's theorems, Whitney's decompositions theorems, strong and weak estimates for maximal operators, Calderòn-Zygmund's singular integral operators, fractional integrals, weights, representation formulas, and so on; since then several studies have appeared, up to nowadays: first order differential calculus in metric spaces, Sobolev–Poincarè inequalities and Sobolev spaces and, finally, geometric measure theory. We recall that spaces of homogeneous type include in particular Carnot-Carathéodory spaces and Carnot groups. Last but not least, in the last decades, several applications of metric and ultrametric spaces, and doubling condition, have been done to computer science, i.e. graph theory and networks, embedding theorems, machine learning, big data and so on.

Let us now recall some definitions.

Definition 2.1.1 (Space of homogeneous type). *A quasi-metric space* (X, ρ) *is called a space of homogeneous type if there exists* $N \in \mathbb{N}$, $N \geq 1$ *such that one of the two following condition is satisfied:*

1_{HT} *for all* $x \in X$, $r > 0$, $\{x_i\}_{i\in I} \subseteq B(x,r)$ *such that* $\rho(x_i, x_j) \geq \frac{r}{2}$ *for all* $i, j \in I$, $i \neq j$, *it follows that* $|I| \leq N$;

2_{HT} *for all* $x \in X$, $r > 0$, $n \in \mathbb{N}$, $n \geq 1$, $\{x_i\}_{i\in I} \subseteq B(x,r)$ *such that* $\rho(x_i, x_j) \geq \frac{r}{2^n}$ *for all* $i, j \in I$, $i \neq j$ *it follows that* $|I| \leq N^n$;

The equivalence of the two above condition is proved in [7], jointly with other properties, and authors refer to these conditions as *homogeneity properties*.

Theorem 2.1.1. *Let* (X, ρ) *be a quasi-metric space. Then conditions* 1_{HT} *and* 2_{HT} *in definition 2.1.1 are equivalent.*

Proof. Assuming that 1_{HT} is true, we proceed by induction on n. For $n = 1$ obviously 2_{HT} holds. For $n > 1$ we suppose 2_{HT} valid for n and fixed $x \in X$, $r > 0$ we consider a set of points $\{x_i\}_{i\in I} \subseteq B(x,r)$ such that $\rho(x_i, x_j) \geq \frac{r}{2^{n+1}}$ for all $i, j \in I$ with $i \neq j$. If $x_{i_0}, x_{i_1}, x_{i_2}, \ldots$ is a maximal set of points such that $\rho(x_{i_j}, x_{i_k}) \geq \frac{r}{2}$ when $j \neq k$, it can not contain more than N points. Then any other point in the set belongs to one of the balls $B(x_{i_j}, \frac{r}{2})$ and by induction it follows that each ball does not contain more than N^n points. For the converse, condition 1_{HT} is obtained by 2_{HT} for $n = 1$. $\square$

These spaces are intimately connected to the geometrically doubling spaces, as in the following definition [34, 35].

Definition 2.1.2 (Geometrically doubling space)**.** *A quasi-metric space* (X, ρ) *is called a geometrically doubling space if there exists a constant* $D \in \mathbb{N}$ *such that one of the two following equivalent condition is satisfied:*

1_D $\forall\, x \in X, \forall\, r > 0, \exists\, x_1, \ldots, x_k \in X$ *with* $k \leq D$, *such that*

$$B(x,r) \subseteq \bigcup_{1\leq i\leq k} B\left(x_i, \frac{r}{2}\right);$$

2_D $\forall\, x \in X, \forall\, r > 0, \forall n \in \mathbb{N}, n \geq 1, \exists\, x_1, \ldots, x_k \in X$ *with* $k \leq D^n$, *such that*

$$B(x,r) \subseteq \bigcup_{1\leq i\leq k} B\left(x_i, \frac{r}{2^n}\right).$$

Remark 2.1.1. Clearly previous definitions can be formulated in a weaker sense with, respectively, "$\geq$" instead of "$>$" in Definition 2.1.1 , and with closed balls instead of open balls in Definition 2.1.2 : *all these definitions are* *qualitatively equivalent*, all proofs rest on typical covering arguments and yield quantitative equivalence between various absolute constant.

Remark 2.1.2. It is no hard to verify that every space of homogeneous type is separable [34].

Now we propose one of the possible proposition that provides quantitative estimates of some equivalences.

Proposition 2.1.1. *Let (X, ρ) be a quasi-metric space with constant $A \in [1, +\infty[$. Then the following facts are equivalent:*

i) $\exists\, M \in \mathbb{N}$, $M \geq 1$ *such that* $\forall\, x \in X$, $\forall\, r > 0$, $\exists\, z_1, \ldots, z_k \in X$ *with* $k \leq M$, *such that*

$$B(x,r) \subseteq \bigcup_{1 \leq i \leq k} \overline{B}\left(z_i, \frac{r}{2}\right);$$

ii) $\exists\, M \in \mathbb{N}$, $M \geq 1$ *such that* $\forall\, x \in X$, $\forall\, r > 0$, $\forall n \in \mathbb{N}$, $n \geq 1$, $\exists\, z_1, \ldots, z_k$ $\in X$ *with* $k \leq M^n$, *such that*

$$B(x,r) \subseteq \bigcup_{1 \leq i \leq k} \overline{B}\left(z_i, \frac{r}{2^n}\right);$$

iii) $\exists\, K \in \mathbb{N}$, $K \geq 1$ *such that* $\forall\, x \in X$, $\forall\, r > 0$, $\forall\, m \in \mathbb{N}$, $m \geq 1$, $\forall$ $\{y_j\}_{j \in J_m} \subseteq B(x,r)$ *such that* $\coprod_{j \in J_m} \overline{B}(y_j, \frac{r}{2^m}) \subseteq B(x,r)$, *then*

$$|J_m| \leq K^m;$$

iv) $\exists\, H \in \mathbb{N}$, $H \geq 1$ *such that* $\forall\, x \in X$, $\forall\, r > 0$, $\forall n \in \mathbb{N}$, $n \geq 1$, $\forall$ $\{x_l\}_{l \in L_n} \subseteq B(x,r) : \rho(x_{l_1}, x_{l_2}) > \frac{r}{2^n}$ *for* $l_1 \neq l_2$, *then*

$$|L| \leq H^{2\log_2 A} H^{n+1};$$

v) $\exists\, H \in \mathbb{N}$, $H \geq 1$ *such that* $\forall\, x \in X$, $\forall\, r > 0$, $\forall\, \{x_l\}_{l \in L_n} \subseteq B(x,r)$: $\rho(x_{l_1}, x_{l_2}) > \frac{r}{2}$ *for* $l_1 \neq l_2$, *then*

$$|L| \leq H^{2\log_2 A} H.$$

Proof. *i)* $\Longrightarrow$ *ii)* Obvious by induction;

ii) $\Longrightarrow$ *iii)* Let $x \in X$, $r > 0$, $m \in \mathbb{N}$ with $m \geq 1$ be fixed and let $\{y_j\}_{j \in J_m} \subseteq B(x,r)$ with $\coprod_{j \in J_m} \overline{B}(y_j, \frac{r}{2^m}) \subseteq B(x,r)$; obviously for all $j \in J_m$, $m \in \mathbb{N}$ with $m \geq 1$, $i = 1, ..., k$ there exists i such that $y_j \in \overline{B}(z_j, \frac{r}{2^n})$, for every choice of $n \in \mathbb{N}$; thus for $n = m$ if $j_1, j_2 \in J_m$ exist such that $y_{j_1}, y_{j_2} \in \overline{B}(z_i, \frac{r}{2^m})$ it follows that $\rho(y_{j_1}, z_i) \leq \frac{r}{2^m}$ and $\rho(y_{j_2}, z_i) \leq \frac{r}{2^m}$, i.e. $z_i \in \overline{B}(y_{j_1}, \frac{r}{2^m}) \cap \overline{B}(y_{j_2}, \frac{r}{2^m})$, false: so $|J_m| \leq N^m$, i.e. the thesis, by assuming $K = M$;

iii) $\Longrightarrow$ *iv)* Let $x \in X$, $r > 0$, $n \in \mathbb{N}$ with $n \geq 1$, $\{x_l\}_{l \in L_m} \subseteq B(x,r)$ with $\rho(x_{l_1}, x_{l_2}) > \frac{r}{2^n}$ for $l_1 \neq l_2$; we choice $m \in \mathbb{N}$ such that $m = n + 1 + 2\log_2 A$, i.e. $2^{m-n-1} \geq A^2$; then it follows that $\coprod_{l \in L_m} \overline{B}(x_l, \frac{Ar}{2^m}) \subseteq B(x, Ar)$, in fact: for all $l \in L_n$, $y \in \overline{B}(x_l, \frac{Ar}{2^m})$ it is

$$\begin{aligned}
\rho(y,x) &\leq A\left(\rho(y,x_l) + \rho(x_l,x)\right) \\
&\leq 2A \max\left\{\rho(y,x_l), \rho(x_l,x)\right\} \\
&\leq 2A \max\left\{\frac{Ar}{2^m}, r\right\} \\
&< 2Ar;
\end{aligned}$$

if then, for some $l_1, l_2 \in L_n$ there exists $c \in X$ such that $c \in \overline{B}(x_{l_1}, \frac{Ar}{2^m}) \cap \overline{B}(x_{l_2}, \frac{Ar}{2^m})$, it follows that

$$\begin{aligned}
\frac{r}{2^n} &< \rho(x_{l_1}, x_{l_2}) \\
&\leq A\left(\rho(x_{l_1}, c) + \rho(c, x_{l_2})\right) \\
&\leq 2A \max\left\{\rho(x_{l_1}, c),\, \rho(c, x_{l_2})\right\} \\
&\leq \frac{A^2 r}{2^{m-1}},
\end{aligned}$$

false. Thus we have $|L_n| \leqslant |J_m| \leqslant K^m = K^{\log_2 A^2} M^{n+1}$, i.e. the thesis assuming $H = K$;

iv) $\Longrightarrow$ *v)* Obvious;

v) $\Longrightarrow$ *i)* Let $x \in X$, $r > 0$ be fixed and $\{x_l\}_{l \in L} \subseteq B(x,r)$ with $\rho(x_{l_1}, x_{l_2}) > \frac{r}{2}$ for $l_1 \neq l_2$; then it follows that or $B(x,r) \subseteq \bigcup_{l \in L} \overline{B}(x_l, \frac{r}{2})$ or there exists z such that $z \in B(x,r) \setminus \bigcup_{l \in L} \overline{B}(x_l, \frac{r}{2})$: in this case it follows that $\rho(z, z_l) > \frac{r}{2}$ for all $l \in L$, and iterating after at most $H^{\log_2 A^2} H$ iterations the ball $B(x,r)$ is covered by at most $H^{\log_2 A^2} H$ closed balls of radius $\frac{r}{2}$, i.e. the thesis assuming $M = H^{\log_2 A^2} H$. $\square$

The adopted terminology may generate some slight ambiguity; in several papers homogeneous spaces are called *doubling spaces*, meaning that the quasi-metric space is endowed with a a so called doubling measure according to the next definition. Note that if X is a topological space with $\mathcal{B}(X)$ we indicate the family of Borel sets of X, i.e. the elements belonging to the σ -algebra generated by the open sets of X.

Definition 2.1.3 (Doubling measure)**.** *A doubling measure μ on a quasi-metric space (X, ρ) is a Borel measure $\mu : \mathcal{B}(X) \to [0, +\infty]$ such that for all balls $B(c, r)$ it is $\mu(B(c, r)) \in]0, +\infty[$ and there exists a constant $C_\mu \geq 1$ such that*

$$\mu(B(c, 2r)) \leqslant C_\mu \mu(B(c, r)) \tag{2.1}$$

for all $x \in X$ and $r > 0$. A doubling measure space (X, ρ, μ) is a quasi-metric space (X, ρ) equipped with a doubling measure μ and C_μ is the doubling constant of the space.

Example 2.1.1. *Lebesgue measure is a doubling measure. Indeed for the Lebesgue measure m on $\mathbb{R}^2$ it is immediately that $m(B(c, 2r)) = 4m(B(c, r))$ holds for all $c \in \mathbb{N}$ and $r > 0$. Generally for all $n \in \mathbb{N}$, $n \geq 1$ the Lebesgue measure m on $\mathbb{R}^n$ satisfies* (2.1) *with constant $C_\mu = 2^n$.*

About the doubling constant C_μ appearing in (2.1), a nice result exists for its optimal value, that is $C_\mu = \sup\limits_{c \in X, r > 0} \frac{\mu(B(c,2r))}{\mu(B(c,r))}$, in a sense that we now clarify, introducing for convenience the following definition.

Definition 2.1.4 (Least doubling constant)**.** *The least doubling constant $C_{(X,\rho)}$ of a quasi-metric space (X, ρ) is the constant*

$$C_{(X,\rho)} = \inf \left\{ C_\mu : \mu \textit{ is a doubling measure on } (X, \rho) \right\},$$

with the convention that if no doubling measure exists on (X, ρ) it is $C_{(X,\rho)} = +\infty$.

Almost all references in the literature place the constant $C_{(X,\rho)}$ in the interval $[1, \infty)$ and one can easily check, unless the metric space reduces to a singleton, that $C_{(X,\rho)} > 1$. Anyway an elementary argument shows that the lower bound $C_{(X,\rho)} \geq \varphi$ always holds, with $\varphi = \frac{\sqrt{5}+1}{2}$ the golden ratio.

Proposition 2.1.2. *Let (X, ρ) be a quasi-metric space with at least two points. Then $C_{(X,\rho)} \geq \varphi = \frac{\sqrt{5}+1}{2}$.*

Proof. Let $x, y \in X$ be with $x \neq y$ and $r = \rho(x, y) > 0$. For some $\lambda > 0$ suppose first

$$\mu\left(B(x, \frac{2r}{3})\right) \leqslant \lambda\mu\left(B(y, \frac{r}{3})\right). \tag{2.2}$$

It is easy to check that $B(x, \frac{2r}{3}) \cap B(y, \frac{r}{3}) = \emptyset$ and $B(x, \frac{2r}{3}) \cup B(y, \frac{r}{3}) \subseteq B(x, \frac{4r}{3})$, so

$$C_\mu\mu\left(B(x, \frac{2r}{3})\right) \geqslant \mu\left(B(x, \frac{4r}{3})\right) \geqslant (1 + \frac{1}{\lambda})\mu\left(B(x, \frac{2r}{3})\right)$$

holds and we have $C_\mu \geqslant (1 + \frac{1}{\lambda})$. Similarly, if

$$\mu\left(B(y, \frac{2r}{3})\right) \leqslant \lambda\mu\left(B(x, \frac{r}{3})\right) \tag{2.3}$$

then one also gets $C_\mu \geqslant (1 + \frac{1}{\lambda})$. Finally assuming that neither (2.2) nor (2.3) hold, then we have

$$\begin{aligned} C_\mu\mu\left(B(x, \frac{r}{3})\right) + \mu\left(B(y, \frac{r}{3})\right) &\geqslant \mu\left(B(x, \frac{2r}{3})\right) + \mu\left(B(y, \frac{2r}{3})\right) \\ &> \lambda(\mu\left(B(y, \frac{r}{3})\right) + \mu\left(B(x, \frac{r}{3})\right)), \end{aligned}$$

which gives $C_\mu > \lambda$. Since this works for any $\lambda > 0$, in any case we get

$$C_\mu \geqslant \sup_{\lambda > 0} \min\left\{\lambda, 1 + \frac{1}{\lambda}\right\},$$

and optimizing in $\lambda > 0$ the thesis follows. $\square$

Remark 2.1.3. The estimate in proposition 2.1.2 can be improved [36]. Indeed it should be noted that, despite the apparently irrelevant choice of radii of the balls considered in the proof, any other combination actually yields the weaker estimate $C_{(X,\rho)} \geq 2$.

An interesting characterization of these measures that offers the idea for further considerations is that expressed by the following theorem. [37, 38].

Theorem 2.1.2. *Let (X, ρ) be a quasi-metric space with constant $A \geq 1$ and μ a Borel measure $\mu : \mathcal{B}(X) \to [0, +\infty]$ such that for all balls $B(c, r)$ it is $\mu(B(c, r)) \in]0, +\infty[$. Then μ is a doubling measure if and only if there exist two constant $s, C' > 0$ such that*

$$\frac{\mu\left(B(x,r)\right)}{\mu\left(B(y,R)\right)} \geq C'\left(\frac{r}{R}\right)^{s} \tag{2.4}$$

for all $x, y \in X$ and $R \geq r > 0$ with $x \in B(y,R)$).

Proof. Assuming that μ is a doubling measure, for $i = 1, 2 \ldots,$ we define $R_i = 2^i r$ and

$$j = \min\left\{i : B(y,R) \subseteq B(x,R_i)\right\}.$$

Iterating j times (2.1) we have $\mu\left(B(y,R)\right) \leqslant \mu\left(B(x,R_i)\right) \leqslant C_\mu^j \mu\left(B(x,r)\right)$ and then

$$\frac{\mu\left(B_r\left(x\right)\right)}{\mu\left(B_R\left(y\right)\right)} \geqslant C_\mu^{-j}.$$

We first note that by $B(y,R) \not\subset B(x,R_{j-1})$ and $x \in B(y,R)$ it follows that $R_{j-1} \leq 2AR$. Indeed assuming $R_{j-1} > 2AR$ if $z \in B(y,R)$ and $x \in B(y,R)$ it follows that $\rho(z,y) < R$ and $\rho(x,y) < R$, so by relaxed triangular inequality we have $\rho(z,x) \leq A(\rho(z,y) + \rho(y,x)) < 2AR < R_{j-1}$ and finally $z \in B(x,R_{j-1})$, a contradiction. By $2^{j-1}r \leq R_{j-1} \leq 2AR$ the maximum value of j can be obtained as

$$j \leq \log_2\left(\frac{4AR}{r}\right). \tag{2.5}$$

Now choosing $C' = C_\mu^{-2-\log_2 A}$ we can determinate $s > 0$ such that $C_\mu^{-j} \geq C'\left(\frac{r}{R}\right)^s$, since (2.4) is satisfied. Indeed we have $j-2-\log_2 A \leqslant s\log_{C_\mu}\left(\frac{R}{r}\right)$ and by (2.5) $j-2 \leqslant \log_2\left(\frac{R}{r}\right)$ holds. To prove the thesis it is enough to choose s such that $s\log_{C_\mu}\left(\frac{R}{r}\right) \geqslant \log_2\left(\frac{R}{r}\right)$, for instance $s = \log_2 C_\mu$. For the converse, assuming $x = y$ e $R = 2r$ immediately we have

$$\mu\left(B(x,2r)\right) \leqslant \frac{1}{C'}\left(\frac{1}{2}\right)^s \mu\left(B(x,r)\right).$$

concluding the proof. □

Theorem 2.1.2 naturally suggests the introduction of a parameter associated with a given doubling measure μ, i.e. the exponent s. However it is always possible to take a larger constant C_μ, so as the proof of theorem shows this parameter is not uniquely definible even if it can not be arbitrarily small. By this considerations the definition of homogeneous dimension can be given.

Definition 2.1.5 (Homogeneous dimension)**.** *The homogeneous dimension of a doubling measure space* (X, ρ, μ) *is the constant* $s = \log_2 C_\mu$.

Considerations about s suggest also the idea of defining the doubling constant C_μ as the lower bound of the constants in (2.1) so that the homogeneous dimension can be uniquely fixed. Anyway sometimes according to context additional conditions can be required for this parameter. Typically the condition $s > 1$ is necessary in the discussion of the theorems of Sobolev's immersion where s plays the same role of the dimension of $\mathbb{R}^n$, while in the case of Lebesgue measure on $\mathbb{R}^n$ it is proved precisely that $C_\mu = 2^n$ and $s = n$. Then for a measure the property of being doubling is closely related to geometry of the space and not all spaces can be equipped with a doubling measure. The following proposition using theorem 2.1.2 offers a necessary condition on the space so that it can support a doubling measure by which follows quite easily that a doubling metric space (X, ρ, μ) is also a space of homogeneous type [7]. The reverse is not true in general [37] and however it is also true that a complete doubling metric space carries a doubling measure [39, 40].

Theorem 2.1.3. *Let* (X, ρ, μ) *be a doubling measure space with doubling constant* C_μ *and let* A *be the constant of quasi-metric* ρ*. Then there exists a constant* $C > 0$*, dependent only on* C_μ *and* A*, such that the minimum number of balls with radius* $\frac{r}{2}$ *necessary to cover the ball* $B(x, r)$ *is not greater than* C *for all* $x \in X$ *and* $r > 0$ *.*

Proof. Let $\{x_i\}_{i \in I} \subseteq B(x, r)$ be a maximal set of points such that $\rho\left(x_i, x_j\right) > \frac{r}{2}$ for all $i, j \in I$ with $i \neq j$. We have that $B(x, r) \subset \bigcup_{i \in I} B(x_i, \frac{r}{2})$, indeed if $z \in B(x, r)$ and $z \notin B(x_i, \frac{r}{2})$ for all $i \in I$ then $\rho(z, x_i) > \frac{r}{2}$ with $z \neq x_i$ for all $i \in I$, a contradiction because $\{x_i\}_{i \in I}$ is a maximal set. Now it is easy to check that the balls $B(x_i, \frac{r}{4A})$ are disjoint, in fact fixed $i, j \in I$ with $i \neq j$, if there exists $z \in B(x_i, \frac{r}{4A}) \cap B(x_j, \frac{r}{4A})$, then we have $\rho(z, x_i) < \frac{r}{4A}$ and $\rho(z, x_j) < \frac{r}{4A}$ so by relaxed triangular inequality it follow that $\rho(x_i, x_j) \leqslant A(\rho(x_i, z) + \rho(z, x_j)) < A(\frac{r}{4A} + \frac{r}{4A}) = \frac{r}{2}$, a contradiction. Furthermore the balls $B(x_i, \frac{r}{4A})$ are such that $\bigcup_{i \in I} B(x_i, \frac{r}{4A}) \subseteq B(x, 2Ar)$, in fact by $x_i \in B(x, r)$ we have $\rho(x_i, x) < r$ and if $z \in B(x_i, \frac{r}{4A})$ for some $i \in I$, it follows that $\rho(z, x_i) < \frac{r}{4A}$ so by relaxed triangular inequality $\rho(z, x) \leqslant A(\rho(z, x_i) + \rho(x_i, x)) < A(\frac{r}{4A} + r) \leqslant 2Ar$ holds, i.e. $z \in B(x, r)$. Then by (2.1) it follows that

$$\sum_{i\in I}\mu\left(B(x_i,\frac{r}{4A})\right)\leqslant\mu\left(B(x,2Ar)\right)$$
$$\leqslant C_\mu\mu\left(B(x,Ar)\right),$$

and (2.4) implies that

$$C_\mu\geqslant\sum_{i\in I}\frac{\mu\left(B(x_i,\frac{r}{4A})\right)}{\mu\left(B(x,Ar)\right)}$$
$$\geqslant\sum_{i\in I}C'\left(\frac{\frac{R}{4A}}{AR}\right)^s$$
$$=\sum_{i\in I}C'\frac{1}{\left(4A^2\right)^s}.$$

Finally assuming $C'=C_\mu^{-2-\log_2 A}$ and $s=\log_2 C_\mu$ as in the proof of theorem 2.1.2 it follows

$$|I|\leqslant C_\mu^{5+\log_2 A}A^{2\log_2 C_\mu},$$

giving the thesis for $C=C_\mu^{5+\log_2 A}A^{2\log_2 C_\mu}$. □

Anyway it is true that every doubling measure space is also a space of homogeneous type.

Theorem 2.1.4. *Let (X,ρ,μ) be a doubling measure space with doubling constant C_μ and let A be the constant of quasi-metric ρ. Then $(X,\rho,)$ is a space of homogeneous type.*

Proof. Fixed $x\in X$ and $r>0$, let $x_i\in B(x,r)$, $i=1,2,\dots,N$, be N points such that $\rho(x_i,x_j)\geq\frac{r}{2}$ with $i\neq j$. Then the balls $B(x_i,\frac{r}{4A})$ are disjoint and it is easy to check that $B(x_i,\frac{r}{4A})\subseteq B(x,R)$ for any $i=1,2,\dots,N$ if $R=(A+\frac{1}{4})r$, so we have $\sum_{i=1}^{N}\mu\left(B(x_i,\frac{r}{4A})\right)\leqslant\mu\left(B(x,R)\right)$. On the other hand it is also $B(x,R)\subseteq B(x_i,\bar{R})$ if $\bar{R}=A(A+\frac{5}{4})r$, so that we have $\mu\left(B(x,R)\right)\leqslant\mu\left(B(x_i,\bar{R})\right)$. For any fixed $x\in X$ and a,b such that $0<b<a$, by (2.1) it follows that $\mu\left(B(x,a)\right)\leqslant A^{\log_2\frac{a}{b}+1}\mu\left(B(x,b)\right)$, so with $a=\bar{R}$ and $b=\frac{r}{4A}$ we obtain $\mu\left(B(x,R)\right)\leqslant\mu\left(B(x_i,\bar{R})\right)\leqslant A^{\log_2\left(4A^3+5A^2\right)+1}\mu\left(B(x_i,\frac{r}{4A})\right)$ for any

$i = 1, 2, \ldots, N$, that together with previous inequality $\sum_{i=1}^{N} \mu\left(B(x_i, \frac{r}{4A})\right) \leqslant \mu\left(B(x, R)\right.$ gives $N \leq A^{\log_2\left(4A^3+5A^2\right)+1}$, concluding the proof. □

Thanks to the previous considerations, since a doubling measure space is also a space of homogeneous type, clearly different examples can be founded among the latter.

Example 2.1.2. *Fixed $n \in \mathbb{N}$ with $n \geq 1$ consider the application $\rho : \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty[$ given by $\rho(x, y) = \sum_{i=1}^{n} |x_i - y_i|^{\alpha_i}$ for all $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, with $\alpha_i > 0$ for all i. It is no hard to check that ρ is a quasi-metric on $\mathbb{R}^n$ though not in general a metric. Then $(\mathbb{R}^n, \rho, m)$, with m the Lebesgue measure on $\mathbb{R}^n$, is a space of homogeneous type because m is a doubling measure.*

Example 2.1.3. *Let M be a C^∞ compact Riemannian manifold and ρ the Riemannian metric on M. Then (M, ρ, μ) is a space of homogeneous type if μ is the volume measure on the Borel sets of M.*

Example 2.1.4. *Let G be a discrete group with a finite generating set $\{g_i\}$, i.e. every $g \in G$ may be represented in at least one way as a finite product*

$$g = g_{i_1}^{\pm 1} g_{i_2}^{\pm 1} \cdots g_{i_k}^{\pm 1}. \tag{2.6}$$

Denoting by e the identity element of G we consider the application $|\cdot| : G \to [0, +\infty[$ defined as follows: $|g| = 0$ if $g = e$, otherwise $|g|$ is the minimal value of k in all representations (2.6) *of g as finite products of the generators and their inverses if $g \neq e$. Then define the application $\rho : G \times G \to [0, +\infty[$ such that $\rho(g, h) = \left|g^{-1}h\right|$, a left-invariant metric on G. If μ is the counting measure on G, then (G, ρ, μ) may or may not be a space of homogeneous type, depending on the algebraic structure of G which regulates the rate of growth of the cardinalities of the balls $B(e, k)$ as $k \to +\infty$. If G is an abelian group, or more generally is a nilpotent group or has a nilpotent subgroup of finite index, then it can be proved that $C > 0$ there exists such that $\mu\left(B(e, k)\right) \leqslant Ck^C$ as $k \to +\infty$ and moreover the doubling property* (2.1) *for μ holds, hence (G, ρ, μ) is indeed a space of homogeneous type. In all other cases a theorem of Gromov shows that $\mu\left(B(e, k)\right)$ grows faster than any power of k, so the doubling property for μ fails.*

Example 2.1.5. *Let $\Omega \subseteq \mathbb{N}$ be an open set with $n \in \mathbb{N}$, $n \geq 1$, and let $X_1, \ldots, X_k$ be C^∞ vector fields in Ω. Suppose that $\{X_i\}$ satisfy the so called condition of Hörmander, which is that they, together with all their commutators of all orders, span the tangent space to $\mathbb{N}$ at each $x \in \Omega$. Then say that a Lipschitz curve $\gamma : [0, r] \to \Omega$ is admissible if for almost every t it follows that*

$$\frac{d\gamma}{dt} = \sum_{i=1}^{k} c_i(t) X_i\left(\gamma(t)\right)$$

where $\sum\limits_{i=1}^{k} |c_i(t)|^2 \leq 1$. Defining the application $\rho : \Omega \times \Omega \to [0, +\infty[$ such that $\rho(x, y)$ is the infimum of the set of all r for which there exists an admissible curve with $\gamma(0) = x$ and $\gamma(r) = y$ it is possible to prove that such an admissible curve exists for any $x, y \in \Omega$ provided Ω is connected. Then ρ is clearly a metric on Ω and if one stays away from the boundary of Ω, then (Ω, ρ, m) with m Lebesgue measure on Ω is a space of homogeneous type.

We conclude this section recalling that the existence of a doubling measure μ imposes some restrictions, both on the measure μ and on the space. For instance if (X, d) is a metric space and the measure μ satisfies the doubling condition, then $\mu(\{x\}) = 0$ for every non-isolated point $x \in X$. Also the space (X, d) must be separable and every bounded subset of X is totally bounded. If X is further complete, then every closed bounded subset of X is compact. Finally not any space of homogeneous type carries a doubling measure: for instance there is a bounded Jordan domain of $\mathbb{R}^n$ for each $n > 2$ which does not carry a doubling measure with respect to the Euclidean metric [41].

2.2 Partition of unity

Partition of unity is usually associated in the literature to covering lemmas which, in their turn, relies either to a on a dyadic type decomposition, or on a Vitali type covering argument or also on the existence of maximal families through Zorn's lemma [42, 43, 44, 45, 14, 46]. Here we exhibit a quite simple proof in the general setting of a space of homogeneous type according to Definition 2.1.1, in the general case of a quasi-metric ρ with the additional hypothesis that $\mathcal{D}_\rho \neq \emptyset$, i.e. ρ do admits an equivalent metric $d \in \mathcal{D}_\rho$ (in particular ρ do verify the relaxed polygonal inequality with constant H); note that this fact occurs when the quasi-metric ρ is on its own a quasi-ultrametric with constant $K \leq 2$.

For completeness we recall that in the general contest of topological spaces a paracompact space, i.e. a space in which every open cover has an open refinement that is locally finite in the sense that each point in the space has a neighbourhood that intersects only finitely many of the sets in the cover, then it is also normal space, i.e. a space in witch every two disjoint closed sets have disjoint open neighborhoods. Then classical Urysohn's lemma the existence of continuous functions that separate disjoint closed convex sets follows for this spaces. But paracompact spaces can be also characterized by a stronger property, i.e. the existence of partitions of unity.

If (X, τ) is a topological space, the support $\mathrm{supp}(f)$ of a function $f : X \to \mathbb{R}$ is the set defined by

$$\mathrm{supp}(f) = \mathrm{cl}\left(\{x \in X : f(x) \neq 0\}\right),$$

where with "cl" we indicate the closure of a set, and the general definition of partition of unity can be given as follows.

Definition 2.2.1 (Partition of unity). *A partition of unity of a topological space (X, τ) is a family $\mathcal{F} = \{\varphi_i : i \in I\}$ of continuous functions $\varphi_i : X \to [0, 1]$ such that*

i) for every $x \in X$ there exists a neighborhood U of x where $\varphi_i = 0$ for all but finitely many $i \in I$;

ii) $\sum\limits_{i \in I} \varphi_i(x) = 1$ for all $x \in X$.

If $\mathcal{A}$ is an open cover of X then a partition of unity $\{\varphi_i : i \in I\}$ is subordinated to $\mathcal{A}$ if the cover $\{\mathrm{supp}(\varphi_i) : i \in I)\}$ of X refines $\mathcal{A}$ and it is called locally finite if the family $\{\mathrm{supp}(\varphi_i) : i \in I)\}$ is locally finite.

Partitions of unity will serve to "glue" local constructions to a global one. For example, assume that $\mathcal{F} = \{\varphi_i : i \in I\}$ is a partition of unity and for each i the function f_i is defined on some set containing $\{x \in X : \varphi_i(x) > 0\}$ and has a certain property $\mathcal{P}$. Then the function $\sum\limits_{i \in I} \varphi_i f_i$ is well defined on X and is a locally finite convex combination of the f_i's, and for certain properties $\mathcal{P}$, e.g., continuity, non negativity, boundedness by a uniform constant, this suffices to ensure that also $\sum\limits_{i \in I} \varphi_i f_i$ has $\mathcal{P}$.

Remark 2.2.1. Let (X, τ) be a topological space and $\{A_i : i \in I)\}$ a locally finite open cover of X. Then there exists a partition of unity $\{\varphi_i : i \in I\}$ such that $\mathrm{supp}(\varphi_i) \subseteq A_i$ for all $i \in I$.

Now to show the next result first we need some notations. For a fixed subset $F \subseteq X$, for any $x \in X$ and for any $\beta > 0$ we set $r_x = \rho(x, F) = \inf\{\rho(x, y) : y \in F\}$, $B(x, r) = \{y \in X : \rho(y, x) < r\}$, $\beta B = B(x, \beta r)$ and $B_x = B(x, r_x)$.

Lemma 2.2.1. *Let (X, ρ) be a quasi-metric space. Then, for any closed set $F \subseteq X$, for any $l \in [1, +\infty[$, for any $\alpha \in \left]0, \frac{1}{l\,A^2}\right[$ and for any $\delta \in]0, \alpha]$, there exists a family $\{x_i\}_{i \in I}$ of points in $X \setminus F$ such that:*

1) $X \setminus F = \bigcup\limits_{i \in I} \delta B_{x_i} = \bigcup\limits_{i \in I} \alpha B_{x_i} = \bigcup\limits_{i \in I} l\alpha B_{x_i}$;

Moreover let (X, ρ) be a space of homogeneous type, then for $k = \frac{A + l\,\alpha\,A^3}{1 - l\,\alpha\,A^2}$, for any $x \in X \setminus F$, $l \in \left[1, +\infty\right[$ and $\alpha \in]0, \frac{1}{lA^2}[$, if $I_x^{l\alpha} = \left\{i \in I : l\alpha B_{x_i} \cap l\alpha B_x \neq \emptyset\right\}$ then we have the following facts:

2) *for any $i \in I_x^{l\alpha}$ it is $\frac{1}{k}\, r_x \le r_{x_i} \le k\, r_x$, in particular for any $i, j \in I_x^{l\alpha}$ with $i \neq j$ it is $\frac{1}{k^2} r_{x_i} \le r_{x_j} \le k^2 r_{x_i}$;*

3) *there exists a positive integer $n = n(A, l)$ such that* card $\left(I_x^{l\alpha}\right) \le N^n$, *where N is the absolute constant in Definition 2.1.1;*

4) *the covering $\{l\alpha B_{x_i}\}_{i \in I}$ is numerable and locally uniformly finite;*

5) *let $d \in \mathcal{D}_\rho$; by assuming $l > H$ there exists a lipschitz continuous partition of unity $\{\varphi_i\}_{i \in I}$ subordinated to the family $\{l\,\alpha B_{x_i}\}_{i \in I}$, and then numerable and locally uniformly finite, i.e. such that*

- $0 \le \varphi_i(x) \le 1$, *for any $x \in X \setminus F$ and for any $i \in I$;*
- φ_i *is continuous and* $\mathrm{supp}(\varphi_i) \subseteq l\,\alpha\,B_{x_i}$, *for any $i \in I$;*
- $\sum_{i \in I} \varphi_i(x) = 1$ *for any $x \in X \setminus F$.*

Proof. Let $\delta > 0$ to be chosen better later. Thanks to Zorn's lemma, fixed a maximal family $\{x_i\}_{i \in I}$ of point in $X \setminus F$ such that, for any $i, j \in I$, $i \neq j$, we have $\rho(x_i, x_j) \ge \delta \, \min\{r_{x_i}, r_{x_j}\}$, we can achieve the thesis.

1) If for some $x \in X \setminus F$ and for all $i \in I$ it were $x \notin \delta B_{x_i}$, then $\rho(x, x_i) \ge \delta \, \min\{r_x, r_{x_i}\}$ for all $i \in I$, and the family $\{x_i\}_{i \in I}$ would not be maximal, so

$X \setminus F \subseteq \bigcup_{i \in I} \delta B_{x_i}$ holds; moreover, if $x \in \bigcup_{i \in I} l\alpha B_{x_i}$, $\exists i \in I$ such that $x \in l\alpha B_{x_i}$, i.e. $\rho(x, x_i) < l\alpha r_{x_i} = l\alpha\rho(x_i, F)$ and $\rho(x, x_i) < l\alpha A(\rho(x_i, x) + \rho(x, F))$; so, for any $l \in [1, +\infty[$, by choosing $\alpha \in]0, 1/lA[$ we have $\rho(x, F) > 0$, i.e. $x \in X \setminus F$.

2) Fixed $l \in \left[1, +\infty\right[$ and $x \in X \setminus F$, for any $i \in I_x^{l\alpha}$, with z fixed in $l\alpha B_{x_i} \cap l\alpha B_x$ and for any $\bar{z}$ in $l\alpha B_x$ we have

$$
\begin{aligned}
r_{x_i} &\leq A\left(\rho(x_i, x) + \rho(x, F)\right) \\
&< A^2(\rho(x_i, z) + \rho(z, x)) + A r_x \\
&< l\alpha A^2 r_{x_i} + A^2\rho(z, x) + A r_x \\
&< l\alpha A^2 r_{x_i} + A^3(\rho(z, \bar{z}) + \rho(\bar{z}, x)) + A r_x \\
&< l\alpha A^2 r_{x_i} + A^3\rho(z, \bar{z}) + l\alpha A^3 r_x + A r_x,
\end{aligned}
$$

and then $(1 - l\alpha A^2) r_{x_i} - (A + l\alpha A^3) r_x < A^3 \rho(z, \bar{z})$ follows, so that, choosing here and in what follows $\alpha < \frac{1}{l A^2} \left(\leq \frac{1}{Al} \right)$, taking infimum on $\bar{z} \in l\alpha B_x$ the right-hand side of the thesis follows by setting $k = \frac{A + l\alpha A^3}{1 - l\alpha A^2}$. The left-hand side then follows immediately by interchanging in the above calculations r_{x_i} and r_x. Finally, the others inequalities immediately can be verified. Note that $k > A$ for any $\alpha \in \left]0, \frac{1}{l A^2}\right[$ and any $l \in [1, +\infty[$.

3) According again to Definition 2.1.1 , fixed $l \in \left[1, +\infty\right[$, if we choose a positive integer n large enough such that $2^n > \frac{l\alpha k}{\delta}$, then, for $x \in X \setminus F$ and $i, j \in I_x^{l\alpha}$, $i \neq j$, thanks to *2)* we have

$$
\begin{aligned}
\rho\left(x_i, x_j\right) &\geq \delta \min\left\{r_{x_i}, r_{x_j}\right\} \\
&\geq \frac{\delta}{k} r_x \\
&> \frac{l\alpha r_x}{2^n},
\end{aligned}
$$

and since the ball $l\alpha B_x$ contains at most N^n different points whose distance is bigger then $\frac{l\alpha r_x}{2^n}$, it follows that card $\left(I_x^{l\alpha}\right) \leq N^n$, i.e. the thesis.

4) Fixed $l \in \left[1, +\infty\right[$ and choosing, for any $x \in X \setminus F$, the ball $l\alpha B_x$ as neighborhood of x, taking into account the estimate just obtained, jointly with the separability of X, we have card $(I) \leq \aleph_0$, and the proof is completed.

5) Fixed $l > H$, let $g : [0, +\infty[\to [0, 1]$ the function

$$g(t) = \begin{cases} 1 & \text{if } t \in [0,1] \\ \frac{Ht-l}{H-l} & \text{if } t \in [1, \frac{l}{H}] \\ 0 & \text{if } t \in [\frac{l}{H}, +\infty[\end{cases} \quad ;$$

then, for any $x \in X$ and any $i \in I$, if we set $\sigma_i(x) = g\big(\frac{d(x,x_i)}{\alpha\, r_{x_i}}\big)$ it is

$$\sigma_i(x) = \begin{cases} 1 & \text{if } x \in \alpha\, B_{x_i} \\ 0 & \text{if } x \notin l\,\alpha B_{x_i} \end{cases} .$$

In particular, fixed $x \in X \setminus F$ we have that, thanks to *1)*, $\sigma_i(x) = 1$ for at least one value of $i \in I_x^{l\alpha}$, and $\sigma_i(x) = 0$ for any $i \in I \setminus I_x^{l\alpha}$; so if $\sigma(x) = \sum_{i \in I} \sigma_i(x)$, it is $1 \leq \sigma(x) = \sum_{i \in I_x^{l\alpha}} \sigma_i(x) \leq N^n < +\infty$ and by setting $\varphi_i(x) = \frac{\sigma_i(x)}{\sigma(x)}$, we obtain the desired continuous locally uniformly finite partition of unity concluding the proof. $\square$

2.3 Lipschitz and hölder continuous functions

We consider some classes of functions specific to distance spaces recalling main definitions and properties. Later we will focus on the special class of lipschitz functions.

Definition 2.3.1 (Continuity). *Let (X_1, ρ_1) and (X_2, ρ_2) be distance spaces. A function $f : X_1 \to X_2$ is called*

i) continuous if for any $x \in X_1$ and any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon, x) > 0$ such that

$$\rho_1(x, y) < \delta \Longrightarrow \rho_2(f(x), f(y)) < \varepsilon \quad \textit{for all} \quad y \in X_1; \tag{2.7}$$

ii) uniformly continuous if for any $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that

$$\rho_2(f(x), f(y)) < \varepsilon \quad \textit{for all} \quad x, y \in X_1 \quad \textit{with} \quad \rho_1(x, y) < \delta_\varepsilon; \tag{2.8}$$

iii) lipschitz continuous if there exists a constant $L \geq 0$ such that

$$\rho_2(f(x), f(y)) \leq L\rho_1(x,y) \quad \textit{for all} \quad x, y \in X_1. \tag{2.9}$$

By definition it is clear that a uniformly continuous function is continuous. Also, it is easy to check that a L-lipschitz continuous function is uniformly continuous, indeed, if $f : X_1 \to X_2$ is L-lipschitz continuous for some constant $L > 0$, then for an arbitrary $\varepsilon > 0$ taking $\delta_\varepsilon = \frac{\varepsilon}{L+1}$ we have $\rho_2(f(x), f(y)) \leq L\rho_1(x,y)$ for all $x, y \in X_1$ with $\rho_1(x,y) < \delta_\varepsilon$. The constant L in (2.9) is called a lipschitz constant for the function f, and usually one says that f is L-lipschitz continuous to emphasize this specific constant that obviously it is not unique, i.e. if $L' > L$ then also L' is a lipschitz constant for the function. Therefore it makes sense to consider the smallest of these constants, i.e. the so called lipschitz norm $L(f)$ of a lipschitz continuous function $f : X_1 \to X_2$ given by

$$L(f) = \inf\left\{L \geqslant 0 : L \quad \text{is a lipschitz constant for} f\right\}.$$

of which some elementary properties are expressed by the following proposition.

Proposition 2.3.1. *Let (X_1, ρ_1) and (X_2, ρ_2) be distance spaces and $f : X_1 \to X_2$ a lipschitz continuous function. Then we have*

$$L(f) = \sup\left\{\frac{\rho_2(f(x), f(y))}{\rho_1(x,y)} : x, y \in X_1, x \neq y\right\}. \tag{2.10}$$

Furthermore the lipschitz norm $L(f)$ is the least lipschitz constant for f, in the sense that

i) $\rho_2(f(x), f(y)) \leq L(f)\rho_1(x,y)$ *for all* $x, y \in X_1$*;*

ii) *if* L *is a lipschitz constant for* f*, then* $L(f) \leq L$*.*

Proof. Let $\mathcal{L}(f)$ be the set of all lipschitz constants of function f, so we have that $L(f) = \inf \mathcal{L}(f)$, and let γ be the supremum in the right hand-side of (2.10). If $L \in \mathcal{L}(f)$ then $\frac{\rho_2(f(x),f(y))}{\rho_1(x,y)} \leqslant L$ for all $x, y \in X_1$, $x \neq y$. We have that $\gamma \leq L$ for all $L \in \mathcal{L}(f)$ and then $\gamma \leq L(f)$. By definition of γ it follows that $\rho_2(x,y) \leq \gamma\rho_1(x,y)$ for all $x, y \in X_1$ so $\gamma \in \mathcal{L}(f)$ and $L(f) \leq \gamma$, finally giving $\gamma = L(f)$, i.e. the (2.10) holds. Then to prove *i)* and *ii)*, just note that $\gamma \in \mathcal{L}(f)$ and $\gamma = \inf \mathcal{L}(f)$. $\square$

Together with the lipschitz continuous functions it is also possible to consider the hölder continuous functions which are a generalization of the previous ones.

Definition 2.3.2 (Hölder continuity). *Let α be a constant such that $0 < \alpha \leq 1$ and let (X_1, ρ_1) and (X_2, ρ_2) be distance spaces. A function $f : X_1 \to X_2$ is called hölder continuous if there exists a constant $L \geq 0$ such that*

$$\rho_2(f(x), f(y)) \leqslant L(\rho_1(x, y))^\alpha \tag{2.11}$$

for all $x, y \in X_1$.

A function f satisfying (2.11) with the constant L is also called as an L-hölder continuous function of order α. Noting that assuming $\alpha = 1$ in condition (2.11) then in becomes the condition (2.9), so we have that the lipschitz continuous function are a special case of hölder. The class of hölder continuous function of a given order α can be denoted by $\mathrm{Lip}_\alpha(X_1, X_2)$ for $0 < \alpha \leq 1$ with the convention that $\mathrm{Lip}_1(X_1, X_2) = \mathrm{Lip}(X_1, X_2)$ denotes the class of lipschitz continuous function. In a natural way, for an $f \in \mathrm{Lip}_\alpha(X_1, X_2)$ an α-hölder norm can be defined as

$$\begin{aligned} L_\alpha(f) &= \inf\{L \geqslant 0 : f \quad \text{is} \quad L\text{-hölder of order} \quad \alpha\} \\ &= \sup_{x \neq y} \frac{\rho_2(f(x), f(y))}{(\rho_1(x, y))^\alpha}, \end{aligned}$$

assuming that $L_1(f) = L(f)$ for consistency with definition of lipschitz norm.

For a fixed α with $0 < x\alpha \leq 1$, considering the distance $\rho_1^\alpha : X_1 \times X_2 \to [0, +\infty]$ we also have the equivalence

$$f \quad \text{is} \quad L\text{-hölder of order} \quad \alpha \Longleftrightarrow f \quad \text{is} \quad L\text{-}(\rho_1^\alpha, \rho_2)\text{-lipschitz},$$

that is to say

$$\mathrm{Lip}_\alpha((X_1, \rho_1), (X_2, \rho_2)) = \mathrm{Lip}((X_1, \rho_1^\alpha), (X_2, \rho_2)),$$

with the preservation of the α-hölder and lipschitz norms, therefore in substance their study coincides with that of lipschitz class. About the chance to take $\alpha > 1$ in the condition (2.11), the following fact should be emphasized if involved distance are metric. Before we give a definition relating to one of the possible notions of convexity in a general metric space.

Definition 2.3.3 (Midpoint convexity). *A metric space (X, d) is called midpoint convex if for all $x, y \in x$ with $x \neq y$ there exists $z \in X$ such that*

$$d(x,z) = d(z,y) = \frac{1}{2}d(x,y).$$

Theorem 2.3.1. *Let (X_1, d_1) be a midpoint convex metric space and (X_2, d_2) an arbitrary metric space. If, for some $L > 0$ and $\alpha > 1$, the function $f : X_1 \to X_2$ satisfies the condition*

$$d_2(f(x), f(y)) \leq L d_1^{\alpha}(x,y),$$

for all $x, y \in X_1$, then f is a constant function.

Proof. First by induction we prove that for all $n \in \mathbb{N}$ the inequality

$$d_2(f(x), f(y)) \leqslant \frac{L}{2^{n(\alpha-1)}} d_1^{\alpha}(x,y) \tag{2.12}$$

holds for all $x, y \in X_1$ with $x \neq y$. Indeed for $n = 1$, fixed $x, y \in X_1$ with $x \neq y$, by midpoint convexity, there exists a point $z \in X_1$ such that $d_1(x,z) = d_1(z,y) = \frac{1}{2}d_1(x,y)$. Then using triangular inequality we have

$$\begin{aligned} d_2(f(x), f(y)) &\leqslant d_2(f(x), f(z)) + d_2(f(z), f(y)) \\ &\leqslant L(d_1^{\alpha}(x,z) + d_1^{\alpha}(z,y)) \\ &= \frac{2L}{2^{\alpha}} d_1^{\alpha}(x,y) \\ &= \frac{L}{2^{\alpha-1}} d_1^{\alpha}(x,y), \end{aligned}$$

i.e. the (2.12) holds. Now let us suppose that (2.12) is true for $k > 1$ and prove it for $k+1$. Again, fixed $x, y \in X_1$ with $x \neq y$, by midpoint convexity, there exists a point $z \in X_1$ such that $d_1(x,z) = d_1(z,y) = \frac{1}{2}d_1(x,y)$. For the conditions (2.12) applied to the pairs of points (x,z) and (z,y) it follows that

$$\begin{aligned} d_2(f(x), f(z)) &\leqslant \frac{L}{2^{k(\alpha-1)}} d_1^{\alpha}(x,z) \\ &= \frac{L}{2^{k(\alpha-1)}} \frac{1}{2^{\alpha}} d_2^{\alpha}(x,z) \end{aligned}$$

and

$$\begin{aligned} d_2(f(z), f(y)) &\leqslant \frac{L}{2^{k(\alpha-1)}} d_1^{\alpha}(z, y) \\ &= \frac{L}{2^{k(\alpha-1)}} \frac{1}{2^{\alpha}} d_2^{\alpha}(z, y), \end{aligned}$$

so finally

$$\begin{aligned} d_2(f(x), f(y)) &\leqslant d_2(f(x), f(z)) + d_2(f(z), f(y)) \\ &\leqslant \frac{L}{2^{(k+1)(\alpha-1)}} d_1^{\alpha}(x, y). \end{aligned}$$

Now by (2.12), for all $x, y \in X_1$ with $x \neq y$, observing that it is $2^{(\alpha-1)}$ we have

$$d_2(f(x), f(y)) \leqslant \frac{L}{2^{n(\alpha-1)}} d_1^{\alpha}(x, y) \to 0$$

as $n \to 0$, i.e. $f(x) = f(y)$ for all $x, y \in X_1$, that is to say f is a constant function on X_1 concluding the proof. $\square$

Remark 2.3.1. Let $I \subseteq \mathbb{R}$ an interval and $f : I \to \mathbb{R}$ an L-hölder continuous function of order α with $\alpha > 1$. Then by

$$\left| \frac{f(x+h) - f(x)}{h} \right| \leqslant L|h|^{\alpha-1},$$

we have $f'(x) = 0$ as $h \to 0$, i.e. f is a constant function on I.

2.4 Lipschitz extension theorems

The problem of finding lipschitz continuous extensions of a given lipschitz continuous function is well known and much discussed in the literature, with applications and repercussions in very different areas as geometry [47], computer science, image processing [48, 49], elasticity [50] and medicine [51]. We first recall various classical extension results for lipschitz functions obtained and then present an original result in the special case of spaces of homogeneous type with a quasi-metric satisfying a relaxed polygonal inequality using an improvement of the partition unity lemma.

Let us start with the proof of classical McShane's extension theorem for real lipschitz continuous function.

Theorem 2.4.1 (McShane's theorem). *Let (X, d) be a metric space, $S \subseteq X$ a non-empty set and $f : S \to \mathbb{R}$ a L - lipschitz continuous function. Then the functions $g, h : X \to \mathbb{R}$ given by*

$$g(x) = \sup_{y \in S} \{f(y) - Ld(x,y)\} \quad \text{and} \quad h(x) = \inf_{y' \in S} \{f(y') - Ld(x,y')\} \tag{2.13}$$

for all $x \in X$ are L-lipschitz continuous extensions of f to X. Furthermore other L-lipschitz continuous extension $e : X \to \mathbb{R}$ of f to X satisfies the inequalities

$$g(x) \leq e(x) \leq h(x) \tag{2.14}$$

for all $x \in X$.

Proof. For arbitrary $x \in X$ and $y, y' \in S$, by lipschitz condition for f and triangular inequality we have

$$\begin{aligned} f(y) - f(y') &\leqslant Ld(y,y') \\ &\leqslant Ld(x,y) + Ld(x,y') \end{aligned}$$

and then

$$f(y) - Ld(x,y) \leqslant f(y') + Ld(x,y'). \tag{2.15}$$

It follows that the functions $g, h : X \to \mathbb{R}$ given by (2.13) are well-defined and

$$g(x) \leq h(x) \tag{2.16}$$

for all $x \in X$. Let us show that g, h are L-lipschitz continuous extensions of f. First we prove that g and h extend f on x. Indeed if $x \in S$ then assuming $y = y' = x$ in (2.15) and by definition (2.13) of the functions g, h and the inequality (2.16) we obtains

$$\begin{aligned} f(x) &\leq g(x) \\ &\leq h(x) \\ &\leq f(x), \end{aligned}$$

showing that $g(x) = h(x) = f(x)$, i.e. g and h are extensions of f. Now prove that g and h are L-lipschitz continuous functions. In fact if $x, x' \in X$ we have

$$\begin{aligned} h(x) &\leq f(y') + Ld(x,y') \\ &\leq f(y) + Ld(x',y') + Ld(x,x'), \end{aligned}$$

for all $y' \in S$, implying

$$h(x) \leq h(x') + Ld(x,x') \iff h(x) - h(x') \leq Ld(x,x'),$$

and interchanging the roles of x and x' we have $h(x') - h(x) \leq Ld(x,x')$ so

$$\left|h(x') - h(x)\right| \leq Ld(x,x')$$

holds, i.e. h is a L-lipschitz continuous function. Then starting from the inequalities

$$\begin{aligned} g(x) &\geq f(y) - Ld(x,y) \\ &\geq f(y) - Ld(x',y) - Ld(x,x') \end{aligned}$$

similarly it can be proved that the function g is L-lipschitz continuous too. Finally we show that any other L-lipschitz continuous extension e of f satisfies the inequalities (2.14). For $x \in X$ and $y \in S$, we have $e(x) - e(y) = e(x) - f(y)$. Since e is L-lipschitz continuous, it follows that

$$\begin{aligned} -Ld(x,y) &\leq e(x) - f(y) \\ &\leq Ld(x,y), \end{aligned}$$

and equivalently,

$$\begin{aligned} f(y) - Ld(x,y) &\leq e(x) \\ &\leq f(y) + Ld(x,y) \end{aligned} \tag{2.17}$$

Finally taking the supremum with respect to $y \in S$ in the left hand side of inequalities (2.17) and the infimum in the right hand-side, we obtains

$$\begin{aligned} f(x) &\leq g(x) \\ &\leq h(x) \\ &\leq f(x), \end{aligned}$$

concluding the proof. □

In the case of a complex lipschizt continuous function an extension with the same lipschitz constant not always exists. Anyway the following result can be easily proved.

Corollary 2.4.1. *Let (X, d) be a metric space, $S \subseteq X$ a non-empty set and $f : S \to \mathbb{C}$ a complex L-lipschitz continuous function. Then there exists a $\bar{L}$-lipschitz continuous extensions of f to X such that $\bar{L} \leqslant \sqrt{2}L$.*

Proof. Let $f_1, f_2 : S \to \mathbb{R}$ be real and imaginary part of f respectively, i.e. $f(x) = f_1(x) + if_2(x)$ for all $x \in S$. Then f_1, f_2 are both L-lipschitz continuous function, and by theorem 2.4.1 have L-lipschitz continuous extensions to X, $e_1, e_2 : X \to \mathbb{R}$. Setting $e(x) = e_1(x) + ie_2(x)$ for all $x \in X$ it follows that

$$|e(x) - e(y)| = \sqrt{(e_1(x) - e_2(y))^2 + (e_2(x) - e_2(y))^2} \leqslant \sqrt{2}Ld(x, y)$$

for all $x, y \in X$, i.e. the function $e : X \to \mathbb{C}$ is a $\bar{L}$-lipschitz continuous extensions of f to X with $\bar{L} \leqslant \sqrt{2}L$. □

Let us remember now that the property for an lipschitz continuous extension to preserve the norm can be proved in the important case of lipschitz continuous functions defined in a subset of an real Hilbert space to another real Hilbert space, according to the well known classical Kirszbraun's theorem. To give a proof of this theorem we need some preliminary definitions and results.

First, considering two metric spaces (X_1, d_1) and (X_2, d_2) and a nonempty subset S of X_1, we say that a function $f : S \to X_2$ is called a contraction or a nonexpansive map if it is L-lipschitz continuous with $L \in [0, 1]$, that is

$$d_2(f(x), f(y)) \leq d_1(x, y),$$

for all $x, y \in S$. Then we say also that a pair $((X_1, d_1), (X_2, d_2))$ of metric spaces has the contraction extension property if for any subset S of X then any contraction mapping $f : S \to X_2$ has a contraction extension $e : X_1 \to X_2$, and that a pair $((X_1, d_1), (X_2, d_2))$ has the lipschitz extension property if any lipschitz continuous function $f : S \to X_2$ has an extension $e : X_1 \to X_2$ with the same lipschitz constant, i.e. there exists an extension e with $L(e) = L(f)$.

Considering now two families of closed balls

$$\bar{B}_i(x_i, r_i) = \{x \in X_1 : d_1(x_i, x) \leqslant r_i\}, \quad i \in I$$

and

$$\bar{B}'_i(x'_i, r_i) = \left\{x' \in X_2 : d_2(x'_i, x') \leqslant r_i\right\}, \quad i \in I$$

we say that a pair of metric spaces $((X_1, d_1), (X_2, d_2))$ has the Kirszbraun property if for any families $\{\bar{B}_i(x_i, r_i) : i \in I\}$ and $\left\{\bar{B}'_i(x'_i, r_i) : i \in I\right\}$ of such closed balls that satisfy the condition

$$d_2(x'_i, x'_j) \leq d_1(x_i, x_j),$$

for all $i, j \in I$, the implication

$$\bigcap_{i \in I} \bar{B}_i(x_i, r_i) \neq \emptyset \Rightarrow \bigcap_{i \in I} \bar{B}'_i(x'_i, r_i) \neq \emptyset.$$

holds. Now we can show some fundamental equivalent facts about these notions according to the next theorem.

Theorem 2.4.2. *Let $((X_1, d_1), (X_2, d_2))$ be a pair of metric spaces. Then the following facts are equivalent.*

1) *the pair $((X_1, d_1), (X_2, d_2))$ has the contraction extension property;*

2) *the pair $((X_1, d_1), (X_2, d_2))$ has the Kirszbraun property.*

Moreover, if X_1 and X_2 are normed spaces, then each of the above properties is equivalent to the following one.

3) *the pair $((X_1, d_1), (X_2, d_2))$ has the contraction extension property.*

Proof.

1) $\Rightarrow$ *2)* Assuming $x_i \in X_1$, $x'_i \in X_2$, $r > 0$, $i \in I$ such that

$$d_2(x'_i, x'_j) \leqslant d_1(x_i, x_j) \quad \text{for all} \quad i, j \in I, \quad \text{with} \quad \bigcap_{i \in I} \bar{B}_i(x_i, r_i) \neq \emptyset,$$

then the map $f : \{x_i : i \in I\} \to \{x'_i : i \in I\}$ given by

$$f(x_i) = x'_i, \quad i \in I,$$

satisfies the inequalities

$$d_2(f(x_i), f(x_j)) \leqslant d_1(x_i, x_j),$$

for all $i, j \in I$, so it is a contraction. Then, by hypothesis, there exists a contraction extension $e : X_1 \to X_2$ such that $g(x_i) = f(x_i) = x'_i$ for all $i \in I$. Finally, if $x \in \bigcap_{i \in I} \bar{B}_i(x_i, r_i)$, then

$$\begin{aligned} d_2(g(x), x'_i) &= d_2(g(x), e(x_i)) \\ &\leq d_1(x, x_i), \end{aligned}$$

for all $i \in I$, showing that $e(x) \in \bigcap_{i \in I} \bar{B}'_i(x_i, r_i)$, i.e. $\bigcap_{i \in I} \bar{B}'_i(x'_i, r_i) \neq \emptyset$.

2) $\Rightarrow$ *1)* Let us consider the set

$$\mathcal{M} = \{(g, U) : S \subseteq U \subseteq X_1, g : U \to X_2 \text{ is a contraction extension of } f\},$$

with the order given by

$$(g_1, U_1) \leq^* (g_2, U_2) \iff U_1 \subseteq U_2 \text{ and } g_2|_{U_1} = g_1.$$

We note that if (g_i, U_i), $i \in I$ is a totally ordered subset of $\mathcal{M}$, then putting $U = \bigcup_{i \in I} U_i$ and defining $g : U \to X_2$ by $g(u) = g_i(u)$ with $i \in I$ such that $u \in U_i$, it is easy to check that map e is well defined, $(e, U) \in \mathcal{M}$ and $(e_i, U_i) \leq^* (g, U)$ for all $i \in I$, so the set $(\mathcal{M}, \leq^*)$ is inductively ordered, and consequently, thanks to Zorn's lemma, it contains a maximal element (e, Y). Now let us suppose that there exists a point $x_0 \in X_1\ Y$; setting

$$r_x = d(x, x_0), \tag{2.18}$$

for all $x \in Y$ then we consider the balls $\bar{B}(x, r_x) \subseteq X_1$ and $\bar{B}'(e(x), r_x) \subseteq X_2$, for all $x \in Y$. Because e is a contraction on Y we have

$$d_2(e(x), e(x')) \leq d_1(x, x'),$$

for all $x, x' \in Y$ and by (2.18) it follows that $x_0 \in \bigcap_{x \in Y} \bar{B}(x, r_x)$ so $\bigcap_{x \in Y} \bar{B}(x, r_x) \neq \emptyset$ and by hypothesis there exists a point $y_0 \in \bigcap_{x \in Y} \bar{B}'(e(x), r_x)$. Setting $Z = Y \cup \{x_0\}$ we define the map $h : Z \to X_2$ by $h(x) = e(x)$ for $x \in Y$ and

$e(x_0) = y_0$, so that $h|_Y = f$ and $y_0 \in \bigcap_{x\in Y} \bar{B}'(e(x), r_x)$ is equivalent to

$$\begin{aligned} d_2(h(x\_0), h(x)) &= d_2(y_0, e(x)) \\ &\leqslant r_x \\ &= d_1(x_0, x), \end{aligned}$$

for all $x \in Y$, proving that h is a contraction, that is $(h, Z) \in \mathcal{M}$. Finally, $(h, Z) \neq (e, Y)$ and $(e, Y) \leq^* (h, Z)$ we obtain a contradiction to the maximality of (e, Y).

Suppose now that (X_1, d_1) and (X_2, d_2) are normed spaces.

3) $\Rightarrow$ *1)* It is obvious for arbitrary metric spaces (X_1, d_1), (X_2, d_2).

1) $\Rightarrow$ *3)* Let Y be some subset of X_1 and let $f : Y \to X_2$ be an L-lipschitz continuous map. If $L = 0$, then f is a constant function, i.e. $f(x) = c$, for all $x \in Y$, for some $c \in X_2$, which automatically extends to $e(x) = c$, for all $x \in X_1$. If $L > 0$ then the map $g = L^{-1}$ is a contraction which, by hypothesis, has an extension to a contraction $h : X_1 \to X_2$, and $e = Lh$ is an L-lipschitz continuous function extending f. $\square$

Before giving a proof of the Kirszbraun's theorem, let us also briefly recall some facts concerning topologies in a Hilbert space. If $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ is a Hilbert space, then it has the so called weak topology for which all the functionals $\mathbf{x} \mapsto \langle \mathbf{x}, \mathbf{y}\rangle$, for $\mathbf{y} \in \mathcal{H}$, are continuous, and every ball $B = \{\mathbf{x} \in \mathcal{H} : \|\mathbf{x}\| \leqslant \alpha\}$, for $\alpha \geq 0$, is compact. This is really a special case of the Banach-Alaoglu theorem, according to which the closed unit ball of the dual space of a normed vector space is compact in the so called weak* topology. Indeed if V is a topological vector space over the field $\mathbb{K}$ of real or complex numbers, i.e. V is a $\mathbb{K}$ linear space equipped with a topology so that vector addition and scalar multiplication are continuous, we may define a possibly different topology on V using the topological or continuous dual space V^*, which consists of all linear functionals from X into the base field $\mathbb{K}$ that are continuous with respect to the given topology. Recalling that $\langle\cdot,\cdot\rangle$ is the canonical evaluation map defined by $\langle \mathbf{x}, \mathbf{x}'\rangle = \mathbf{x}'(\mathbf{x})$ for all $\mathbf{x} \in V$ and $\mathbf{x}' \in V^*$, where in particular, $\langle\cdot, \mathbf{x}'\rangle = \mathbf{x}'(\cdot) = x'$, then the weak topology on V is the weakest topology on V making all maps $\mathbf{x}' = \langle\cdot, \mathbf{x}'\rangle : V \to \mathbb{K}$ continuous, as $\mathbf{x}'$ ranges over V^*, and the weak topology on V^* is the weakest topology on V^* making all maps $\langle \mathbf{x}, \cdot \rangle : X^* \to \mathbb{K}$ continuous, as $\mathbf{x}$ ranges over V, also called the weak* topology.

Remark 2.4.1. Many results for convex sets in a topological vector space can be proved using the classic Hahn-Banach extension theorem which allows to extend bounded linear operators defined on a subspace of some vector space to the whole space. In particular, the following separation result is very useful in applications: let (X, τ) be a topological vector space over the field $\mathbb{K}$ of real or complex numbers and A, B disjoint nonempty convex subsets of X. If A is open, then there exist a continuous linear functional $x^* \in X^*$ and $\alpha, \beta \in \mathbb{R}$ such that $\operatorname{Re} \mathbf{x}^*(\mathbf{x}) < \alpha \leqslant \operatorname{Re} \mathbf{x}^*(\mathbf{y})$ for all $\mathbf{x} \in A$ and $\mathbf{y} \in B$. Then if A is compact, B closed and X is locally convex, i.e. every point in X has a neighborhood basis formed of convex sets, then there exist a continuous linear functional $\mathbf{x}^* \in X^*$ and $\alpha, \beta \in \mathbb{R}$ such that $\operatorname{Re} \mathbf{x}^*(\mathbf{x}) < \alpha < \beta < \operatorname{Re} \mathbf{x}^*(\mathbf{y})$ for all $\mathbf{x} \in A$ and $\mathbf{y} \in B$.

Now we can give a proof of the Kirszbraun's theorem in the following general version.

Theorem 2.4.3 (Kirszbraun's theorem)**.** *Let $(\mathcal{H}_1, \langle \cdot, \cdot \rangle_1)$ and $(\mathcal{H}_2, \langle \cdot, \cdot \rangle_2)$ be real Hilbert spaces, $S \subseteq \mathcal{H}_1$ a non-empty set and $f : S \to \mathcal{H}_2$ a L - lipschitz continuous function. Then there exists a lipschitz continuous function $e : \mathcal{H}_1 \to \mathcal{H}_2$ extending f to $\mathcal{H}_2$ and such that $L(e) = L(f)$.*

Proof. Thanks to theorem 2.4.2 it suffices to show that the pair of metric spaces $((\mathcal{H}_1, d_1), (\mathcal{H}_2, d_2))$ has the Kirszbraun property, with $d_i(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_i$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{H}_i$, and $\|\mathbf{v}\|_i = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle_i}$ for all $\mathbf{v} \in \mathcal{H}_i$, $i = 1, 2$, as usual in Hilbert spaces.

So if $\bar{B}_i(\mathbf{x}_i, r_i)$ and $\bar{B}'_i(\mathbf{y}_i, r_i)$, $i \in I$, are two families of closed balls in $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, such that

$$\left\|\mathbf{y}_i - \mathbf{y}_j\right\|_2 \leq \left\|\mathbf{x}_i - \mathbf{x}_j\right\|_1 \text{ for all } i, j \in I, \text{ and } \bigcap_{i \in I} \bar{B}_i(\mathbf{x}_i, r_i) \neq \emptyset,$$

we have to prove that the condition

$$\bigcap_{i \in I} \bar{B}'_i(\mathbf{y}_i, r_i) \neq \emptyset$$

holds. So let us assume that

$$\bigcap_{j\in J} \bar{B}'_j(\mathbf{y}_j, r_j) \neq \emptyset, \tag{2.19}$$

for every nonempty finite subset J of I. Fix an element $i_0 \in J$ and consider the sets $C_i = \bar{B}'_{i_0}(\mathbf{y}_{i_0}, r_{i_0}) \cap \bar{B}'_i(\mathbf{y}_i, r_i)$, $i \in I$, observing that since all the balls $\bar{B}'_i(\mathbf{y}_i, r_i)$ are weakly compact, it follows that the sets C_i are weakly compact, hence weakly closed subsets of the weakly compact set $\bar{B}'_{i_0}(\mathbf{y}_{i_0}, r_{i_0})$. By (2.19) this family has the finite intersection property, so it has a nonempty intersection, hence

$$\bigcap_{i\in I} \bar{B}'_i(\mathbf{y}_i, r_i) = \bigcap_{i\in I} C_i \neq \emptyset$$

holds. Now suppose that $\mathbf{w} \in \bigcap_{i\in I} \bar{B}_i(\mathbf{x}_i, r_i) =$ and let J be a finite subset of I, observing that to simplify the notation we can suppose $J = \{1, 2, \ldots, n\}$ for same $n \in \mathbb{N}$. If $\mathbf{w} = \mathbf{x}_i$ for some $i \in J$, then $\left\|\mathbf{y}_j - \mathbf{y}_i\right\|_2 \leqslant \left\|\mathbf{x}_j - \mathbf{x}_j\right\|_1 \leqslant r_j$ for all $j \in J$, so that $\mathbf{y}_i \in \bigcap_{j\in J} \bar{B}'_j(\mathbf{y}_j, r_j)$. If $\mathbf{w} \notin \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, then we consider the function $g : \mathcal{H}_2 \to \mathcal{R}$ given by

$$g(\mathbf{y}) = \max\left\{\frac{\|\mathbf{y} - \mathbf{y}_i\|}{\|\mathbf{w} - \mathbf{x}_i\|} : 1 \leqslant i \leqslant n\right\}, y \in \mathcal{H}_2,$$

and set $\lambda = \inf g(\mathcal{H}_2)$. Because $\lim\limits_{\|\mathbf{y}\|_2 \to 0} g(\mathbf{y}) = +\infty$ there exists $r > 0$ such that $g(\mathbf{y}) > \lambda + 1$ for all $\mathbf{y} \in \mathcal{H}_2$ with $\|\mathbf{y}\|_2 > r$. As the function g is weakly lower semicontinuous and the ball $rB_{\mathcal{H}_2} = r\{\mathbf{t} \in \mathcal{H}_2 : \|\mathbf{t}\|_2 \leqslant 1\}$ is weakly compact, there exists $\mathbf{v} \in rB_{\mathcal{H}_2}$ such that

$$0 < g(\mathbf{v}) = \inf g(\mathcal{H}_2) = \lambda.$$

Numbering the points $\mathbf{x}_i$ we can suppose that it is

$$\begin{aligned} \|\mathbf{v} - \mathbf{y}_i\|_2 &= \lambda\|\mathbf{w} - \mathbf{x}_i\|_1 \text{ for } 1 \leq i \leq k, \\ \|\mathbf{v} - \mathbf{y}_i\|_2 &< \lambda\|\mathbf{w} - \mathbf{x}_i\|_1 \text{ for } k \leq i \leq n, \end{aligned} \tag{2.20}$$

for some $k \in \{1, 2, \ldots, n\}$. Now we can show that $\mathbf{v}$ is a convex combination of vectors $\{\mathbf{y}_1, \ldots, \mathbf{y}_k\}$, i.e.

$$\mathbf{v} \in \operatorname{conv}(\{\mathbf{y}_1, \ldots, \mathbf{y}_k\}). \tag{2.21}$$

Indeed if (2.21) is not true then according the separation results in remark

2.4.1, we have that $\mathbf{v}$ can be strictly separated from conv $(\{\mathbf{y}_1, \ldots, \mathbf{y}_k\})$, so there exists a vector $\mathbf{u} \in \mathcal{H}_2$ with $\|\mathbf{u}\|_2 = 1$ and such that

$$\langle \mathbf{v}, \mathbf{u} \rangle_2 < \langle \mathbf{y}_i, \mathbf{u} \rangle_2 \text{ for } 1 \leqslant i \leqslant k,$$

that is

$$\langle \mathbf{v} - \mathbf{y}_i, \mathbf{u} \rangle_2 < 0 \text{ for } 1 \leqslant i \leqslant k. \tag{2.22}$$

For $\mathbf{z} = \mathbf{v} + t\mathbf{u}$, with $t > 0$, then we have

$$\|\mathbf{z} - \mathbf{y}_i\|_2^2 = \|\mathbf{v} - \mathbf{y}_i\|_2^2 + t\left(2\langle \mathbf{v} - \mathbf{y}_i, \mathbf{u} \rangle_2 + \|\mathbf{u}\|_2^2\right),$$

for all $\leq i \leq k$, and by the second of (2.20) and (2.22) we can choose a sufficiently small $t > 0$ such that $2\langle \mathbf{v} - \mathbf{y}_i, \mathbf{u} \rangle_2 + \|\mathbf{u}\|_2^2 < 0$ for $1 \leq i \leq k$ and $\|\mathbf{z} - \mathbf{y}_i\|_2 < \|\mathbf{w} - \mathbf{x}_i\|_1$ for $k < 1 \leq n$, so that $g(\mathbf{z}) < g(\mathbf{v})$, in contradiction to the choice of $\mathbf{v}$. By (2.21) there exists $\lambda_1, \lambda_2, \ldots, \lambda_k \geqslant 0$ with $\lambda_1 + \lambda_2 + \ldots + \lambda_k = 1$ such that

$$\mathbf{v} = \lambda_1 \mathbf{y}_1 + \lambda_2 \mathbf{y}_2 + \ldots + \lambda_k \mathbf{y}_k, \tag{2.23}$$

and we have

$$\begin{aligned} \left\|\mathbf{y}_i - \mathbf{y}_j\right\|_2^2 &= \left\|(\mathbf{y}_i - \mathbf{v}) + (\mathbf{v} - \mathbf{y}_j)\right\|_2^2 \\ &= \|\mathbf{y}_i - \mathbf{v}\|_2^2 + \left\|\mathbf{y}_j - \mathbf{v}\right\|_2^2 + 2\langle \mathbf{y}_i - \mathbf{v}, \mathbf{v} - \mathbf{y}_j \rangle_2, \end{aligned} \tag{2.24}$$

and similarly

$$\left\|\mathbf{x}_i - \mathbf{x}_j\right\|_1^2 = \|\mathbf{x}_i - \mathbf{w}\|_1^2 + \left\|\mathbf{x}_j - \mathbf{w}\right\|_1^2 + 2\langle \mathbf{x}_i - \mathbf{w}, \mathbf{w} - \mathbf{x}_j \rangle_1. \tag{2.25}$$

Replacing (2.24) and (2.25) in the inequality $\left\|\mathbf{y}_i - \mathbf{y}_j\right\|_2^2 \leqslant \left\|\mathbf{x}_i - \mathbf{x}_j\right\|_1^2$ we obtain

$$\begin{aligned} \|\mathbf{y}_i - \mathbf{v}\|_2^2 + \left\|\mathbf{y}_j - \mathbf{v}\right\|_2^2 + 2\langle \mathbf{v}_i - \mathbf{v}, \mathbf{v} - \mathbf{y}_j \rangle_2 \leq \|\mathbf{x}_i - \mathbf{w}\|_1^2 + \|\mathbf{x}_i - \mathbf{w}\|_1^2 \\ + 2\langle \mathbf{x}_i - \mathbf{w}, \mathbf{w} - \mathbf{x}_j \rangle_1, \end{aligned} \tag{2.26}$$

for all $1 \leq i, j \leq k$.

If $\lambda > 1$, by the first of (2.20) we have

$$\begin{aligned} -\|\mathbf{y}_i - \mathbf{v}\|_2^2 &< -\|\mathbf{x}_i - \mathbf{w}\|_1^2, \\ -\left\|\mathbf{y}_j - \mathbf{v}\right\|_2^2 &< -\left\|\mathbf{x}_j - \mathbf{w}\right\|_1^2, \end{aligned} \tag{2.27}$$

and adding these inequalities to (2.26) we obtain

$$2\langle \mathbf{y}_i - \mathbf{v}, \mathbf{v} - \mathbf{y}_j \rangle_2 < 2\langle \mathbf{x}_i - \mathbf{w}, \mathbf{w} - \mathbf{x}_j \rangle_1,$$

that is

$$\langle \mathbf{y}_i - \mathbf{v}, \mathbf{y}_j - \mathbf{v} \rangle_2 > \langle \mathbf{x}_i - \mathbf{w}, \mathbf{x}_j - \mathbf{w} \rangle_1 \tag{2.28}$$

for all $1 \leq i, j \leq k$.

Observing that (2.23) implies

$$\begin{aligned} \sum_{i,j=1}^{k} \lambda_i \lambda_j \langle \mathbf{y}_i - \mathbf{v}, \mathbf{y}_j - \mathbf{v} \rangle_2 &= \sum_{j=1}^{k} \lambda_j \langle \sum_{i=1}^{k} \lambda_i \mathbf{y}_i - \mathbf{v}, \mathbf{y}_j - \mathbf{v} \rangle_2 \\ &= 0, \end{aligned} \tag{2.29}$$

and that, by direct calculation, it is

$$\sum_{i,j=1}^{k} \lambda_i \lambda_j \langle \mathbf{x}_i - \mathbf{w}, \mathbf{x}_j - \mathbf{w} \rangle_1 = \left\| \sum_{i=1}^{k} \lambda_i (\mathbf{x}_i - \mathbf{w}) \right\|_1^2, \tag{2.30}$$

multiplying the inequality (2.28) by $\lambda_i \lambda_j$, summing for $i, j = 1, \ldots, k$ and taking into account the equalities (2.29) and (2.30), we have

$$\begin{aligned} 0 &> \left\| \sum_{i=1}^{k} \lambda_i (\mathbf{x}_i - \mathbf{w}) \right\|_1^2 \\ &\geq 0, \end{aligned} \tag{2.31}$$

i.e. a contradiction, so it must be $\lambda \leq 1$. Finally the relations (2.20) show that

$$\mathbf{v} \in \bigcap_{i=1}^{n} \bar{B}(\mathbf{y}_i, r_i),$$

concluding the proof. $\square$

As it is clear by the previous argument, we recall that the main tool in the proof of the existence of the extension is a property expressed in terms of the intersection of some families of balls, i.e. the Kirszbraun property. Similar but less general proofs exist, for instance in the special case for $X_1 = X_2 = \mathbb{R}$ [52] and in the Hilbert case for $X_1 = X_2 = \mathcal{H}$ [53, 54, 55].

2.5 Extension of lipschitz functions on space of homogeneous type

We apply previous results to the extension of normed valued lipschitz functions defined on a closed subset of a space of homogeneous type endowed with a suitable quasi-metric. As a first result we prove that the partition of lemma 2.2.1 consists of lipschitz continuous functions, i.e., it is a lipschitz partition of unity according to the following natural definition.

Definition 2.5.1 (Lipschitz partition of unity)**.** *A lipschitz partition of unity subordinated to the open cover $\{A_i : i \in I)\}$ of a topological space (X, τ) is a family $\mathcal{F} = \{\varphi_i : i \in I\}$ of continuous functions $f : X \to [0, 1]$ satisfying conditions of definition 2.2.1 and such that φ_i is a lipschitz continuous function for all $i \in I$.*

Lemma 2.5.1. *Under the same hypotheses, the partition of unity $\{\varphi_i\}_{i\in I}$ defined on X founded in Lemma 2.2.1 is also a lipschitz continuous partition of unity in $X \setminus F$.*

Proof. We observe that the function g is $\frac{H}{l-H}$ - lipschitz, so that for any $i \in I$ the σ_i is lipschitz with the same constant and the function $\sigma(x) = \sum_{i\in I_x^{l\alpha}} \sigma_i(x)$ is $\frac{H}{l-H}N^n$-lipschitz. Hence for any $x, y \in X \setminus F$ and for any $i \in I$ we have

$$\begin{aligned}
|\varphi_i(x) - \varphi_i(y)| &= \left|\frac{\sigma_i(x)}{\sigma(x)} - \frac{\sigma_i(y)}{\sigma(y)}\right| \\
&= \frac{|\sigma_i(x)\,\sigma(y) - \sigma_i(y)\,\sigma(x)|}{\sigma(x)\,\sigma(y)} \\
&\leq |(\sigma_i(x) - \sigma_i(y))\,\sigma(y) + (\sigma(y) - \sigma(x))\,\sigma_i(y)| \\
&\leq |\sigma_i(x) - \sigma_i(y)| + \sigma_i(y)\,|\sigma(y) - \sigma(x)| \\
&\leq \frac{H}{l-H}(1 + N^n)\,d(x, y) \\
&\leq \frac{H}{l-H}(1 + N^n)\,\rho(x, y),
\end{aligned}$$

i.e. is $\frac{H}{l-H}(1 + N^n)$ - lipschitz for any $i \in I$ giving the thesis. $\square$

Now in the following lemma we refine *5)* of previous Lemma 2.2.1, so that lipschitz functions of the partition of unity depends on the radius of the single ball αB_{x_i}: this lemma will be useful in the lipschitz extension results [56, 19].

Lemma 2.5.2. *Under the same hypotheses of Lemma 2.2.1 there exists a lipschitz continuous partition of unity $\{\varphi_i\}_{i\in I}$ defined on X and subordinated to the family*

$\{\alpha B_{x_i}\}_{i\in I}$*, numerable and locally uniformly finite, such that 1), 2), 3) and 4) hold, jointly with the following:*

6) $\mathrm{supp}(\varphi_i) \subseteq \alpha B_{x_i}, \forall i \in I$*;*

7) there exists an absolute constant $C > 0$ *such that* φ_i *is* $\frac{C}{r_{x_i}}$ *- lipschitz* $\forall\, i \in I$*.*

Proof. 6) Fixed $\delta \in]0, \frac{\alpha}{2H}]$ and $l > \max\{H, 2A\}$, for any $i \in I$ let $g_i : [0,+\infty[\to [0,1]$ be such that

$$g_i(t) = \begin{cases} 1 & \text{if } t \in [0, \frac{r_{x_i}}{2H}] \\ 2\frac{r_{x_i} - Ht}{r_{x_i}} & \text{if } t \in [\frac{r_{x_i}}{2H}, \frac{r_{x_i}}{H}] \\ 0 & \text{if } t \in [\frac{r_{x_i}}{H}, +\infty[\end{cases}$$

and $\sigma_i(x) = g_i(\frac{d(x,x_i)}{\alpha})$, where $d \in \mathcal{D}_\rho$. So if $I_x^\alpha = \{i \in I : x \in \alpha B_{x_i}\} \subseteq I_x^{l\alpha}$, for any $x \in X \setminus F$ we have $\sigma_i(x) = 0$ for any $i \in I \setminus I_x^\alpha$, and in force of $\delta \le \frac{\alpha}{2H}$ and *1)* of Lemma 2.2.1, also $\sigma_i(x) = 1$ for at least one value of $i \in I$. Then if we set $\sigma(x) = \sum_{i\in I} \sigma_i(x)$ for any $x \in X \setminus F$, it is $1 \le \sigma(x) = \sum_{i\in I_x^\alpha} \sigma_i(x) \le \sum_{i\in I_x^{l\alpha}} \sigma_i(x) \le N^n < +\infty$, i.e. this sum is positive and, thanks to *c)* of Lemma 2.2.1, uniformly finite for any $x \in X \setminus F$. By setting $\varphi_i(x) = \frac{\sigma_i(x)}{\sigma(x)}$, we obtain the desired continuous locally uniformly finite partition of unity.

7) Let $i \in I$. We distinguish three cases : $x, y \in X \setminus F$, $x \in F$ and $y \in X \setminus F$, $x, y \in F$. In the first case we have

$$\begin{aligned} |\varphi_i(x) - \varphi_i(y)| &= \left| \frac{\sigma_i(x)}{\sigma(x)} - \frac{\sigma_i(y)}{\sigma(y)} \right| \\ &= \frac{|\sigma_i(x)\sigma(y) - \sigma_i(y)\sigma(x)|}{\sigma(x)\sigma(y)} \\ &= |(\sigma_i(x) - \sigma_i(y))\sigma(y) + (\sigma(y) - \sigma(x))\sigma_i(y)| \\ &\le |\sigma_i(x) - \sigma_i(y)|\sigma(y) + |\sigma(y) - \sigma(x)|\sigma_i(y)\ , \end{aligned}$$

and the $\frac{2H}{\alpha r_{x_i}}$ - lipschitz continuity of σ_i jointly to $\sigma(y) \le N^n$ yields

$$|\varphi_i(x) - \varphi_i(y)| \le \frac{2H}{\alpha r_{x_i}} N^n \rho(x,y) + R_i(x,y),$$

where $R_i(x,y) = |\sigma(x) - \sigma(y)|\,\sigma_i(y)$. If $y \notin \alpha B_{x_i}$ then $\sigma_i(y) = 0$ and also $R_i(x,y) = 0$. So let us assume that $y \in \alpha B_{x_i}$: then

$$R_i(x,y) \le (\sigma(x) + \sigma(y))\,\sigma_i(y) \tag{2.32}$$
$$\le 2N^n. \tag{2.33}$$

Now we have either $x \notin l\alpha B_{x_i}$ or $x \in l\alpha B_{x_i}$; in the first case it is $l\alpha r_{x_i} \le \rho(x, x_i)$ so that $l\alpha r_{x_i} \le \rho(X \setminus l\alpha B_{x_i}, x_i)$; then for any $z \in \alpha B_{x_i}$ it is $l\alpha r_{x_i} \le A\,(\rho(X \setminus l\alpha B_{x_i}, z) + \rho(z, x_i))$ so that

$$0 < \alpha\left(\frac{l}{A} - 1\right) r_{x_i} \tag{2.34}$$
$$\le \rho(X \setminus l\alpha B_{x_i}, \alpha B_{x_i})\,; \tag{2.35}$$

it follows that either

$$\rho(x,y) \ge \alpha\left(\frac{l}{A} - 1\right) r_{x_i} \quad \text{or} \quad \rho(x,y) < \alpha\left(\frac{l}{A} - 1\right) r_{x_i}.$$

The second case cannot occur because if $\rho(x,y) < \alpha\left(\frac{l}{A} - 1\right) r_{x_i}$ then $\rho(x,y) < \rho(X \setminus l\alpha B_{x_i}, \alpha B_{x_i})$ and it would follow that $x \in l\alpha B_{x_i}$; so $1 \le \frac{A}{\alpha(l-A) r_{x_i}}\rho(x,y)$, and

$$R_i(x,y) \le 2N^n \frac{A}{\alpha\,(l-A)\,r_{x_i}}\rho(x,y)\,,$$

which gives thesis in this case. So let $x \in l\alpha B_{x_i}$. By introducing the characteristic functions of $l\alpha B_{x_i}$, noting that $\mathbb{1}_{l\alpha B_{x_i}}(x) = \mathbb{1}_{l\alpha B_{x_i}}(y) = 1$, so that

$$|\sigma_i(x) - \sigma_i(y)| \le \frac{2H}{\alpha r_{x_i}}\rho(x,y)\left(\mathbb{1}_{l\alpha B_{x_i}}(x) + \mathbb{1}_{l\alpha B_{x_i}}(y)\right),$$

we have

$$\begin{aligned} R_i(x,y) &\le \sum_{j\in I}\left|\sigma_j(x) - \sigma_j(y)\right| = \sum_{j\in I}\left|\sigma_j(x) - \sigma_j(y)\right|\mathbb{1}_{l\alpha B_{x_i}}(x)\mathbb{1}_{l\alpha B_{x_i}}(y) \\ &\le \left(\sum_{j\in I}\frac{2H}{\alpha r_{x_j}}\left(\mathbb{1}1_{l\alpha B_{x_j}}(x) + \mathbb{1}1_{l\alpha B_{x_j}}(y)\right)\rho(x,y)\right)\mathbb{1}1_{l\alpha B_{x_i}}(x)\mathbb{1}1_{l\alpha B_{x_i}}(y) \\ &= \frac{2H}{\alpha}\left(\sum_{j\in I}\frac{1}{r_{x_j}}\mathbb{1}_{l\alpha B_{x_i}\cap l\alpha B_{x_j}}(x) + \sum_{j\in I}\frac{1}{r_{x_j}}\mathbb{1}_{l\alpha B_{x_i}\cap l\alpha B_{x_j}}(y)\right)\rho(x,y). \end{aligned}$$

The sum $\sum\limits_{j\in I}\frac{1}{r_{x_j}}\mathbb{1}_{l\alpha B_{x_i}\cap l\alpha B_{x_j}}(x)$ has non-null terms exactly when $x \in l\alpha B_{x_i} \cap$

$l\alpha B_{x_j}$, so by using *2)* of Lemma (2.2.1) it follows that $\sum_{j\in I} \frac{1}{r_{x_j}} \mathbb{1}_{l\alpha B_{x_i} \cap l\alpha B_{x_j}}(x) = \sum_{j\in I_x^{l\alpha}} \frac{1}{r_{x_j}} \leq \frac{k^2 N^n}{r_{x_i}}$ (analogous calculations for $\sum_{j\in I} \frac{1}{r_{x_j}} \mathbb{1}_{l\alpha B_{x_i} \cap l\alpha B_{x_j}}(y)$); so we obtain

$$R_i(x,y) \leq \frac{4Hk^2N^n}{\alpha r_{x_i}} \rho(x,y)\,;$$

so that

$$|\varphi_i(x) - \varphi_i(y)| \leq \frac{2N^n}{\alpha r_{x_i}} \left(H + \max\left\{ \frac{A}{l-A}, 2Hk^2 \right\} \right) \rho(x,y),$$

for any $x, y \in X \setminus F$. In the second case, by definition of φ_i we have $\varphi_i(x) = 0$ for $x \in F$ so, if $y \notin \alpha B_{x_i}$ then $\varphi_i(y) = 0$, while if $y \in \alpha B_{x_i}$, $x \in F$ implies $x \notin l\alpha B_{x_i}$, then (2.34) holds and $1 \leq \frac{A}{\alpha(l-A)r_{x_i}} \rho(x,y)$, so

$$\begin{aligned} |\varphi_i(x) - \varphi_i(y)| &= \varphi_i(y) \\ &\leq 1 \\ &\leq \frac{A}{\alpha\,(l-A)\,r_{x_i}} \rho(x,y)\,; \end{aligned}$$

finally, the third case is trivial because $|\varphi_i(x) - \varphi_i(y)| = 0$. All in all, by comparing lipschitz constants for the three cases we conclude that the function φ_i is $\frac{C}{r_{x_i}}$ - lipschitz on X, with an absolute constant

$$C = \frac{2N^n}{\alpha} \left(H + \max\left\{ \frac{A}{l-A}, 2Hk^2 \right\} \right). \quad \square \tag{2.36}$$

Due to the general lack of completeness of the quasi-metric space, we need also the following lemma, where a suitable family of point $\{p_i\}_{i\in I}$ is found as substitutes of projections on the family $\{x_i\}_{i\in I}$ on the set F.

Lemma 2.5.3. *Under the hypothesis of Lemma 2.2.1 there exists a family of point* $\{p_i\}_{i\in I} \subseteq F$ *such that,* $\forall\ i \in I$, $\forall\ y \in F$, *we have* $\rho(p_i, y) \leq 5A^4 \rho(x,y)$ $\forall\ x \in \alpha B_{x_i}$ *and* $\rho(p_i, x) \leqslant 2(l+1)A^3 r_{x_i}$ $\forall\ x \in l\alpha B_{x_i}$.

Proof. First we note that for any $\gamma \in]1, +\infty[$ it follows that $\gamma B_{x_i} \cap F \neq \emptyset$, for any $i \in I$. Indeed, if $\gamma B_{x_j} \cap F = \emptyset$ for some $\gamma > 1$ and some $j \in I$, then $x \notin \gamma B_{x_j}$ for any $x \in F$, so that $\rho\left(x, x_j\right) \geq \gamma r_{x_j}$ and passing to infimum on $x \in F$ a contradiction. Let so $\gamma \in]1, +\infty[$ to be fixed later: let us choose a family of points $\{p_i\}_{i\in I}$ such that $p_i \in \gamma B_{x_i} \cap F \ \forall i \in I$. For any $x \in \alpha B_{x_i}$, $y \in F$ and $p_i \in \gamma B_{x_i} \cap F$, in force of $\alpha < \gamma$, we have

$$\begin{aligned}\rho(p_i, x) &\leq A(\rho(p_i, x_i) + \rho(x_i, x)) \\ &< A(\gamma + \alpha) r_{x_i} \\ &< 2A\gamma r_{x_i},\end{aligned}$$

and also

$$\begin{aligned}\rho(p_i, y) &\leq A(\rho(p_i, x) + \rho(x, y)) \\ &< 2A^2 \gamma r_{x_i} + A\rho(x, y).\end{aligned}$$

Fixed $\bar{x} \in \eta B_{x_i}$ with $\eta < \alpha$, by

$$\begin{aligned}r_{x_i} =&\rho(F, x_i) \\ &\leq A(\rho(F, \bar{x}) + \rho(\bar{x}, x_i)) \\ &\leq A^2\rho(F, x) + A^2\rho(\bar{x}, x) + A\rho(\bar{x}, x_i) \\ &\leq A^2\rho(F, x) + A^3\rho(\bar{x}, x_i) + A^3\rho(x_i, x) + A\rho(\bar{x}, x_i),\end{aligned}$$

we obtain

$$r_{x_i} < A^2\rho(x, y) + A^3\alpha r_{x_i} + 2A^3\eta r_{x_i},$$

and

$$r_{x_i} < \frac{A^2}{1 - A^3\alpha}\rho(x, y) + \frac{2A^3}{1 - A^3\alpha}\eta r_{x_i}.$$

Finally we have

$$\rho(p_i, y) < \frac{2A^4}{1 - A^3\alpha}\gamma\rho(x, y) + \frac{4A^5}{1 - A^3\alpha}\gamma\eta r_{x_i} + A\rho(x, y)$$

for all $\eta \in]0, \alpha[$, so

$$\rho(p_i, y) \leq \frac{2A^4}{1 - A^3\alpha}\gamma\rho(x, y) + A\rho(x, y).$$

Observing that for $\alpha \in]0, \frac{1}{lA^2}[$ with $l > 2A$ it is $2(1 - A^3\alpha) \in]1, 2[$, by assuming $\gamma = 2(1 - \alpha A^3)$ we obtain

$$\rho\left(p_i, y\right) \leq (4A^4 + A)\rho\left(x, y\right) \leq 5A^4\rho\left(x, y\right),$$

i.e. the first inequality. For the second one, with $x \in l\alpha B_{x_i}$, $y \in F$ and $p_i \in \gamma B_{x_i} \cap F$, similarly, by

$$\begin{aligned}
\rho(p_i, x) &\leq A(\rho(p_i, x_i) + \rho(x, x_i)) \\
&\leq A\rho(p_i, x_i) + A^2\rho(x, \bar{x}) + A^2\rho(\bar{x}, x_i) \\
&\leq A\rho(p_i, x_i) + A^3\rho(x, x_i) + A^3\rho(x_i, \bar{x}) + A^2\rho(\bar{x}, x_i),
\end{aligned}$$

we obtain

$$\begin{aligned}
\rho(p_i, x) &< A\gamma r_{x_i} + A^3 l\alpha r_{x_i} + 2A^3\eta r_{x_i} \\
&< (l+1)A^3\gamma r_{x_i} + 2A^3\eta r_{x_i}
\end{aligned}$$

so that $\rho(p_i, x) < 2(l+1)A^3 r_{x_i} + 2A^3\eta r_{x_i}$ for all $\eta \in]0, \alpha[$, and the second inequality follows. $\square$

Theorem 2.5.1. *Let (X, ρ) be a space of homogeneous type, $(Y, \|\cdot\|)$ a normed space, $F \subseteq X$ a non-empty closed set and $f : F \to Y$ a L - lipschitz continuous function. Then if $\mathcal{D}_\rho \neq \emptyset$ there exists a lipschitz continuous extension of f on X.*

Proof. Let $\{p_i\}_{i\in I} \subseteq F$ as in Lemma (2.5.3) and define the function $\bar{f} : F \to Y$ as follow:

$$\bar{f}(x) = \begin{cases} \sum_{i\in I} \varphi_i(x) f(p_i) & \text{if } x \in X \setminus F \\ f(x) & \text{if } x \in F \end{cases}, \tag{2.37}$$

where $\{\varphi_i\}_{i\in I}$ are as in Lemma 2.5.2 . The function $\bar{f}$ is well posed and extends f to all X by construction: let us prove that there exists a constant $\bar{L} > 0$ such that $\|\bar{f}(x) - \bar{f}(y)\| \leq \bar{L}\rho(x, y)$ for any $x, y \in X$. Obviously $\bar{f}$ is lipschitz continuous with constant $\bar{L} = L$ for any $x, y \in F$. We distinguish two cases : $x \in X \setminus F, y \in F$ and $x, y \in X \setminus F$. In the first case, according to Lemma 2.5.2 , the properties of partition of unity $\{\varphi_i\}_{i\in I}$ subordinated to the family $\{\alpha B_{x_i}\}_{i\in I}$, recalling that $I_x^\alpha \subseteq I_x^{l\alpha}$, gives

$$\begin{aligned}
\bar{f}(x) - \bar{f}(y) &= \sum_{i\in I} \varphi_i(x) f(p_i) - f(y) \\
&= \sum_{i\in I_x^\alpha} \varphi_i(x) \left(f(p_i) - f(y)\right),
\end{aligned}$$

and then $\|\bar{f}(x)-\bar{f}(y)\| \leq \sum_{i\in I_x^\alpha} \|f(p_i)-f(y)\|$. Because of L - lipschitz continuity of f on F by hypothesis, and taking into account that card $(I_x^\alpha) \leq N^n$ and first inequality of Lemma 2.5.3 , we have

$$\begin{aligned} \|\bar{f}(x)-\bar{f}(y)\| &\leq L \sum_{i\in I_x^\alpha} \rho(p_i,y) \\ &\leq 5A^4LN^n\rho(x,y). \end{aligned}$$

In the second case, let $z \in F$ be such that $\rho(x,z) \leqslant 2\rho(x,F)$. Now for the properties of partition of unity we have

$$\bar{f}(x)-\bar{f}(y) = \sum_{i\in I} \left(f(p_i)-f(z)\right)\left(\varphi_i(x)-\varphi_i(y)\right),$$

and for the L - lipschitz continuity of f in F and the $\frac{C}{r_{x_i}}$ - lipschitz continuity of φ_i in X we have

$$\|\bar{f}(x)-\bar{f}(y)\| \leq LC\left(\sum_{i\in I}\frac{\rho(p_i,z)}{r_{x_i}}\left(\mathbb{1}_{\alpha B_{x_i}}(x)+\mathbb{1}_{\alpha B_{x_i}}(y)\right)\right)\rho(x,y), \tag{2.38}$$

where C is the constant in (2.36). Considering that

$$\sum_{i\in I}\frac{\rho(p_i,z)}{r_{x_i}}\mathbb{1}_{\alpha B_{x_i}}(x) = \sum_{i\in I_x^\alpha}\frac{\rho(p_i,z)}{r_{x_i}},$$

in the left-hand side of (2.38), by

$$\begin{aligned} \rho(x,z) &\leqslant 2\rho(x,F) \\ &\leqslant 2\rho(x,p_i) \end{aligned}$$

and

$$\rho(p_i,z) \leqslant A(\rho(x,p_i)+\rho(x,z)),$$

we have $\rho(p_i,z) \leq 3A\rho(x,p_i)$, and thanks to the second inequality of Lemma (2.5.3), if $i \in I_x^\alpha$ then we have

$$\rho\left(p_i,x\right) \leqslant 2(l+1)A^3 r_{x_i},$$

so that

$$\frac{\rho(z,p_i)}{r_{x_i}} \leqslant 6(l+1)A^4,$$

and therefore it is

$$\sum_{i\in I}\frac{\rho(p_i,z)}{r_{x_i}}\mathbb{1}_{\alpha B_{x_i}}(x)\leq 6(l+1)A^4N^n.$$

For the other sum

$$\sum_{i\in I}\frac{\rho(p_i,z)}{r_{x_i}}\mathbb{1}_{\alpha B_{x_i}}(y)=\sum_{i\in I_y}\frac{\rho(p_i,z)}{r_{x_i}},$$

we define $Q_i(x,y)=\frac{\rho(x,y)}{r_{x_i}}$ and assume that $i\in I$ is such that $y\in\alpha B_{x_i}$; if it occurs $Q_i(x,y)<\frac{\alpha}{k}$, thanks to *2)* of Lemma (2.2.1) we have

$$\begin{aligned}\rho(x,x_i)&\leq A(\rho(x,y)+\rho(y,x_i))\\&<A(Q_i(x,y)k+\alpha)r_{x_i}\\&<l\alpha r_{x_i},\end{aligned}$$

i.e. $x\in l\alpha B_{x_i}$ and, as in the first sum, it is $\frac{\rho(x,p_i)}{r_{x_i}}\leq 6(l+1)A^4$, concluding that the function $\bar{f}$ is lipschitz continuous with a constant equal to $12(l+1)LCA^4N^n$; otherwise, if $Q_i(x,y)\geq\frac{\alpha}{k}$, then it follows that $r_x\leq\frac{k}{\alpha}\rho(x,y)$ and by

$$\begin{aligned}\rho(x,z)&\leqslant 2\rho(x,F)\\&=2r_x,\end{aligned}$$

we obtain

$$\rho(x,z)\leq\frac{2k}{\alpha}\rho(x,y),$$

so that

$$\begin{aligned}\rho(y,z)&\leq A(\rho(x,y)+\rho(x,z))\\&\leq A\left(1+\frac{k}{\alpha}\right)\rho(x,y),\end{aligned}$$

and we can conclude that

$$\rho(x,z)+\rho(z,y)\leq\left(A+(A+1)\frac{2k}{\alpha}\right)\rho(x,y);\tag{2.39}$$

hence, we have

$$\left\|\bar{f}(x)-\bar{f}(y)\right\|\leqslant\left\|\bar{f}(x)-\bar{f}(z)\right\|+\left\|\bar{f}(z)-\bar{f}(y)\right\|,$$

with $x, y \in X \setminus F$ and $z \in F$, so, for the case just discussed, we obtain

$$\|\bar{f}(x) - \bar{f}(y)\| \leq 5A^4 N^n L \left(\rho(x,z) + \rho(z,y)\right),$$

so that, thanks to (2.39) , the function $\bar{f}$ results lipschitz continuous with constant equal to $5\ A^4\ N^n\ L\left(A + (A+1)\,\frac{2k}{\alpha}\right)$. Finally, all in all, we have that $\bar{f}$ is lipschitz continuous in X with a constant

$$\bar{L} = A^4 N^n L \max\left\{12(l+1)C, 5\left(A + (A+1)\,\frac{2k}{\alpha}\right)\right\}, \tag{2.40}$$

so completing the proof of theorem. □

Finally following corollary is then apparent.

Corollary 2.5.1. *Let (X, ρ) be a space of homogeneous type, $(Y, \|\cdot\|)$ a Banach space, $S \subseteq X$ a non-empty set and $f : S \to Y$ a L - lipschitz continuous function. Then, if $\mathcal{D}_\rho \neq \emptyset$ there exists a lipschitz continuous extension of f on X.*

Proof. By standard arguments the function f can be extended to an L - lipschitz function to the whole $\overline{S}$: the thesis then follows immediately from Theorem 2.5.1 . □

Chapter 3

Selected topics on calculus on Carnot groups

In this chapter we prove the well known Whitney's extension theorem for real valued functions defined on a step two Carnot group $\mathbb{G}$. Whitney's proof relies on two key tools: a particular open covering for open sets in $\mathbb{R}^n$ and an extension operator built up with the help of a suitable partition of unity subordinated to such a covering. We employ the partition of unity lemma in Chapter 2 adapted to the setting of the doubling metric space $\mathbb{H}^1$. For the sake of clarity we provide the quite detailed proof in $\mathbb{H}^1$, avoiding longer technicalities that naturally appear in the general setting of a step two Carnot group. So let us begin with basic definitions and properties.

3.1 Preliminaries and notations

A Lie group is a smooth manifold obeying the group properties and that satisfies the additional condition that the group operations are differentiable.The tangent space at the identity of a Lie group always has the structure of a Lie algebra, and this Lie algebra determines the local structure of the Lie group via the exponential map. A nilpotent Lie group is a Lie group which is connected and whose Lie algebra is a nilpotent Lie algebra [57, 58, 59]. A Carnot group $\mathbb{G}$ of step 2 is a connected and simply connected nilpotent Lie group whose Lie algebra, denoted by $\mathfrak{g}$, admits a suitable decomposition as direct sum of vector subspaces, i.e. there exists a vector subspace V_1 such that setting $V_2 = [V_1, V_1]$ we have $\mathfrak{g} = V_1 \oplus V_2 : V_1$ is called *the horizontal slice* and its vector fields the *horizontal vector fields* on $\mathbb{G}$. Such an algebra is nilpotent of step 2 by definition, i.e. V_2 is included in the center of $\mathfrak{g}$. The stratified structure of $\mathfrak{g}$ gives rise to a family $\{\gamma_\lambda\}_{\lambda\in\mathbb{R}}$, called *dilations*, defined on the generators by imposing, for any $\lambda \geq 0$, $\gamma_\lambda(X) = \lambda X$ whenever $X \in V_1$ and $\gamma_\lambda(X) = \lambda^2 X$ whenever $X \in V_2$; moreover we set $\gamma_{-1}(X) = -X$. For $\lambda \neq 0$ we then have a group of automorphisms of $\mathfrak{g}$. Thanks to the nilpotence, the exponential mapping $\exp : \mathfrak{g} \to \mathbb{G}$ is a global diffeomorphism so it is possible to push forward these dilations on $\mathbb{G}$ by the position $\delta_\lambda = \exp \circ \gamma_\lambda \circ \exp^{-1}$. Setting $\dim(V_1) = l$ and $\dim(V_2) = p$, one can choose a basis of $\mathfrak{g}$ adapted to the stratification by selecting an arbitrary basis of left invariant vector fields X_j, $j = 1, \dots, l+p$, assuming also that, thanks to the left invariance property, $X_j(e) = e_j$, where e is the identity of $\mathbb{G}$ and $\{e_j\}_{j=1,\dots,l+p}$ denotes the standard basis of $\mathbb{R}^{l+p}$; nevertheless, for our purposes, we choose another basis of $\mathfrak{g}$ selecting first a basis $X_1, \dots, X_l$ of V_1 as just described, and choosing the basis $\{T_1, \dots, T_p\}$, $p \leq \binom{l}{2}$, of V_2, between all elements of the kind $[X_{j_1}, X_{j_2}]$, $1 \leq j_1 < j_2 \leq l$. We stress that in general $T_j(e) \neq e_{l+j}$ for $j = l, \dots, p$. Relatively to this basis we introduce *Malcev's exponential graded coordinates of the first kind of* by setting $x = (x_1, \dots, x_l, t_1, \dots, t_p)$ where $x = \exp(\sum_{j=1}^{l} x_j X_j + \sum_{j=1}^{p} t_j T_j)$: in particular, when $\lambda \geq 0$, it results $\delta_\lambda(x) = = (\lambda x_1, \dots, \lambda x_l, \lambda^2 t_1, \dots, \lambda^2 t_p)$. In terms of these coordinates the product law of two elements $p, q \in \mathbb{G}$ is recovered through the Baker-Campbell-Dynkin-Hausdorff formula – BCDH for short – $P(R, S)$ between $R = \exp^{-1}(x)$ and $S = \exp^{-1}(y)$ so that it results $x\,y = \exp(P(R, S))$: the components of the exponential coordinates of the product are in general of polynomial kind and with respect to such a given system of coordinates, it is possible to prove some fundamental facts of the group law, that we collect in the following proposition whose proof can be adapted from the one of Proposition 2.1 in [60].

Proposition 3.1.1 (Structure of group's law). *Let us denote by $\mathbb{G} \ni (x, y) \to xy \in \mathbb{G}$ the group law in $\mathbb{G}$, and by "$+$" the usual Euclidean sum in $\mathbb{R}^{l+p}$. Then, there exists a polynomial vector function $Q : \mathbb{G} \times \mathbb{G} \to \mathbb{R}^n \equiv \mathbb{R}^l \oplus \mathbb{R}^p$, where $Q(x,y) = \big(Q_1(x,y), Q_2(x,y)\big)$, $Q_1(x,y) = \big(Q_1^1(x,y), \dots, Q_1^l(x,y)\big) \in \mathbb{R}^l$ and $Q_2(x,y) = \big(Q_2^1(x,y), \dots, Q_2^p(x,y)\big) \in \mathbb{R}^p$ are such that $xy = x + y + Q(x,y)$. Moreover, for any $i = 1, \dots, l$ and for any $j = 1, \dots, p$ we have:*

i) *for any $x, y \in \mathbb{G}$, it results $Q_1^i(x,y) = 0$ and $Q_2^j(x,0) = Q_2^j(0,x) = Q_2^j(x,x) = Q_2^j(x,-x) = 0$;*

ii) *each Q_2^j is a homogeneous polynomial of degree 2 with respect to the dilations δ_λ, i.e. for any $x, y \in \mathbb{G}$ and for any $\lambda \geq 0$, it results $Q_2^j\big(\delta_\lambda(x), \delta_\lambda(y)\big) = \lambda^2 Q_2^j(x,y)$;*

iii) *each Q_2^j depends only on the coordinates of the first slice i.e. for any $x, y \in \mathbb{G}$ it results $Q_2^j(x,y) = Q_2^j(x_1,\dots,x_l,y_1,\dots,y_l)$; more precisely each $Q_2^j(x,y)$ is a sum of terms, each of which contains a factor of the kind $(x_h y_k - x_k y_h)$, for some $1 \le h,k \le l$.*

From previous facts it follows that the identity e is exactly $0 \in \mathbb{R}^{l+p}$ and that the inverse of an element $x = (x_1,\dots,x_l,t_1,\dots,t_p)$ is exactly $-x = (-x_1,\dots,-x_l,-t_1,\dots,-t_p)$. So, roughly speaking, we can think of a Carnot group of step two as the set $\mathbb{R}^n$ endowed with a polynomial group law in each coordinate, and with a suitable family of endomorphisms $\{\delta_\lambda\}_{\lambda\in\mathbb{R}}$.

Definition 3.1.1 (Homogeneity). *We say a function $f : \mathbb{G} \to \mathbb{R}$ homogeneous of degree $\alpha \in \mathbb{R}$ if $f\big(\delta_\lambda(x)\big) = \lambda^\alpha f(x)$ for any $\lambda > 0$. We say a left invariant differential operator D on $\mathbb{G}$ homogeneous of degree $\alpha \in \mathbb{R}$ if $D(f \circ \delta_\lambda)(x) = \lambda^\alpha (Df \circ \delta_\lambda)(x)$, for any smooth function f and for any $\lambda > 0$.*

If D is a left invariant differential operator homogeneous of degree α, and f is function homogeneous of degree β, then Df, if defined, is a function homogeneous of degree $\beta - \alpha$ and fD is a left invariant differential operator homogeneous of degree $\alpha - \beta$; moreover if D_1, D_2 are left invariant differential operators, homogeneous of degree α and β respectively, then $D_1 D_2$ is a left invariant differential operator homogeneous of degree $\alpha + \beta$.

Arguing as in Proposition 2.2. of [60] we can prove the following properties on the left invariant vector fields $\{X_j\}$ of the basis previously chosen.

Proposition 3.1.2. *At each point $x \in \mathbb{G}$, it results*

$$X_j|_x = \partial_j + \sum_{r=1}^{p} \left[\frac{\partial Q_2^{l+r}}{\partial y_j}(x,y)\right]_{y=0} \partial_{l+h}\,, \qquad \text{for } j = 1,\dots,l,$$

$$T_j|_x = \sum_{i=1}^{l} a_i^j \partial_i + \sum_{i=1}^{p} a_{l+i}^j \partial_{l+i}\,,, \qquad \text{for } j = 1,\dots,p,$$

and where $T_j(0) = (a_1^j,\dots,a_n^j)$. Moreover each X_j is homogeneous of degree 1 and each T_j is homogeneous of degree 2.

Definition 3.1.2 (Degree of a multi index). *Let $I = (n_1,\dots,n_l,m_1,\dots,m_p) \in \mathbb{N}^p$ be a given $(l+p)$-tuple of non negative integers. We will call the numbers $d(I) = (\sum_{j=1}^{l} n_j) + 2(\sum_{j=1}^{p} m_j)$ and $|I| = \sum_{j=1}^{l} n_j + \sum_{j=1}^{p} m_j$ respectively the homogeneity degree and the order of the multi index I. If, moreover, $I_k = (j_1,\dots,j_k)$*

$\in \{1, \dots, l, l+1, \dots, l+p\}^k$, *setting* $\lambda_{i_j} = 1$ *if* $1 \le i_j \le l$, *and* $\lambda_{i_j} = 2$ *if* $l+1 \le i_j \le l+p$, *we will call the numbers* $d(I_k) = \sum_{j=1}^{k} \lambda_{i_j}$ *and* $|I_k| = k$, *respectively the homogeneity degree and the order of the multi index* I_k.

Example 3.1.1. *If* $I = (n_1, \dots, n_l, m_1, \dots, m_p) \in \mathbb{N}^p$, *then this* $(l+p)$*-tuple will refer to the operator* $X^I = X_1^{n_1} \cdots X_l^{n_l} T_1^{m_1} \cdots T_p^{m_p}$ *and we will call the numbers* $d(I)$ *and* $|I|$ *respectively the homogeneity degree and the order of such operator. If moreover* $I_k = (j_1, \dots, j_k) \in \{1, \dots, l, l+1, \dots, l+p\}^k$ *and* $(\alpha_1, \dots, \alpha_k) \in \mathbb{N}^k$, *setting* $h = \sum_{i=1}^{k} \alpha_i$, *then the iterated h-derivation* $Y_h = Y_{j_1}^{\alpha_1} \cdots Y_{j_k}^{\alpha_k}$ *will have and order h. In particular any horizontal k-derivation* X_{H_k} *has homogeneous degree as well as order k.*

Definition 3.1.3 (Derivations)**.** *Let* $I_k = (j_1, \dots, j_k) \in \{1, \dots, l, l+1, \dots, l+p\}^k$ *: a left invariant differential operator of the kind* $Y_{I_k} := Y_{j_1} \cdots Y_{j_k}$ *is called (iterated) k-derivation; if, in particular,* $I_k := H_k = (j_1, \dots, j_k) \in \{1, \dots, l\}^k$, *then the corresponding left invariant differential operator* $X_{H_k} = X_{j_1} \cdots X_{j_k}$ *is called horizontal k-derivation. A 0-derivation, denoted with the symbol* $X_{I_0} = X_{H_0}$, *is just the identity operator.*

Horizontal k-derivations are also called nonholonomic partial derivatives of order k [61].

Remark 3.1.1 (Birkhoff–Poincaré–Witt Theorem)**.** The Birkhoff–Poincaré–Witt theorem ensures that the set of all canonical derivations Z^I constitutes a basis for the algebra of all left invariant differential operators on $\mathbb{G}$ [59]. In particular, taking into account the stratified structure of $\mathfrak{g}$, the following spanning relationships easily hold

$$\begin{aligned} \operatorname{Span}_{\mathbb{R}}(\{Z_{I_k}\})_{d(I_k)=k} &= \operatorname{Span}_{\mathbb{R}}(\{Z_{H_k}\}) \\ &= \operatorname{Span}_{\mathbb{R}}(\{Z^I\})_{d(I)=k} \\ &= \operatorname{Span}_{\mathbb{R}}(\{Z_\alpha^k\})_{\alpha \in A_k}. \end{aligned}$$

Definition 3.1.4 (Polynomials)**.** *A function* $P : \mathbb{G} \to \mathbb{R}$ *is called a polynomial on* $\mathbb{G}$ *if* $P \circ \exp^{-1}$ *is a polynomial on the vector space* $\mathfrak{g}$.

If $\{\omega_j\}_{j=1,\dots,l+p}$ is the dual basis of $\{X_1, \dots, X_l, T_1, \dots, T_p\}$, then for any $x = \exp(\sum_{j=1}^{l} x_j X_j + \sum_{j=1}^{p} t_j T_j) \in \mathbb{G}$ it results $\omega_j \circ \exp^{-1}(x) = x_j$ if $j = 1, \dots, l$, $\omega_{l+j} \circ \exp^{-1}(x) = t_j$ if $j = 1, \dots, p$. So, if $I = (n_1, \dots, n_l, m_1, \dots, m_p)$ denotes a $(l+p)$-tuple of non negative integers, the general monomial has the form

$\mathtt{m}^I(x) = x^I = x_1^{n_1} \cdots x_l^{n_l} t_1^{m_1} \cdots t_p^{m_p}$. Clearly $\mathtt{m}^I$ has homogeneous degree $d(I)$. A basis for the real vector space of polynomials of degree m at most is the set of all monomials $\mathtt{m}^I$ for all n-tuple I such that $d(I) \leq m$, so that a general polynomial P_m of homogeneous degree m is $P_m = \sum_{d(I)\leq m} a_I \, \mathtt{m}^I$, $a_I \in \mathbb{R}$.

Definition 3.1.5 (Carnot-Carathéodory metric)**.** The distance between any two given points $x, y \in \mathbb{G}$ is the infimum of all T such that there exists an absolutely continuous *horizontal curve* joining the points, i.e. a curve $\gamma : [0, T] \to \mathbb{G}$ such that $\dot{\gamma} = \sum_{j=1}^{l} \mu_j X_j(\gamma)$ a.e., for some measurable vector function $\mu : [0, T] \to \mathbb{R}^l$, $\|\mu\|_\infty \leq 1$, and $\gamma(0) = x$, $\gamma(T) = y$. Such a curves there exist as a consequence of Chow-Rashevskiĭ Theorem, jointly with the Hörmander bracket generating conditions trivially satisfied by the horizontal vector fields $X_1, \ldots, X_l$. The distance d is finite, moreover for any fixed euclidean compact set $K \subset \mathbb{G}$, there exists a constant $C = C(K)$, such that

$$\frac{1}{C}\|x - y\| \leq d(x, y) \leq C\|x - y\|^{\frac{1}{k}}, \tag{D}$$

for any $x, y \in K$, where $\|\cdot\|$ is the euclidean norm. So the topology induced by d is the Euclidean one, but the two distances are not metrically equivalent. The metric d is (left) translation invariant and the homogeneous of degree one with respect to the dilations, i.e. δ_λ, i.e. $d(zx, zy) = d(x, y)$, for any $x, y, z \in \mathbb{G}$ and for any $\lambda \in \mathbb{R}$; through the distance d, it is possible to define a so called "homogeneous"norm through the position $|z|_d = d(z, 0)$: such a norm is not smooth in general but there exist several quasi norms N equivalent to $|\cdot|_d$, smooth and easier to handle, for instance – in a step two Carnot group – the following

$$N(z) = \left(x_1^4 + \cdots + x_l^4 + t_1^2 + \cdots + t_p^2\right)^{\frac{1}{4}}$$

for any $z \in \mathbb{G}$; indeed, setting

$$c_1 = \inf\left\{|z|_d : N(z) = 1\right\} \qquad c_2 = \sup\left\{|z|_d : N(z) = 1\right\}$$

we have $0 < c_1 < c_2 < +\infty$ and

$$c_1\, N(z) \leq |z|_d \leq c_2\, N(z) \tag{E}$$

for any $z \in \mathbb{G}$. From now we shall denote with d the Carnot-Carathéodory metric.

Definition 3.1.6 (Measure and Dimension). *The "homogeneous dimension" of $\mathbb{G}$ is the integer defined by the position*

$$Q = \sum_{i=1}^{2} i \, \dim(V_i) = l + 2p \, .$$

This integer is the Hausdorff dimension of $\mathbb{R}^n$ $(n = l + p)$ with respect to the distance d; moreover the n-dimensional Lebesgue measure $\mathcal{L}^n$ is the Haar measure of the group $\mathbb{G}$; then the translation and dilations invariant conditions read as follows; if $E \subset \mathbb{G}$ is a measurable set, $B_r(x)$ denotes the open ball of radius $r > 0$ about $x, z \in \mathbb{G}$ and $\lambda > 0$, then we have

$$\mathcal{L}^n(z\,E) = \mathcal{L}^n(E)\,, \quad \mathcal{L}^n\big(\delta_\lambda(E)\big) = \lambda^Q\,\mathcal{L}^n(E)\,, \quad \mathcal{L}^n\big(B_r(x)\big) = r^Q\,\mathcal{L}^n\big(B_1(0)\big)\,.$$

In particular $(\mathbb{G}, d)$ is a doubling space.

Definition 3.1.7 (Derivative of a function). *Let $f : \mathbb{G} \to \mathbb{R}$, $c \in \mathbb{G}$, $X \in V_1$, $v = \exp(X)$. The function f is differentiable along X at the point c if the function $\mathbb{R} \ni \lambda \to f\big(c\,\delta_\lambda(v)\big) \in \mathbb{R}$ is differentiable at the point $\lambda = 0$: in this case we write $Xf(c)$ for such a limit.*

Definition 3.1.8 (Functions of class $C^m_{\mathrm{H}}(\mathbb{G})$). *Let $f : \mathbb{G} \to \mathbb{R}$; we say that $f \in C^1_{\mathrm{H}}(\mathbb{G})$ if $X_j f$ exist and are continuous at each point of $\mathbb{G}$, for every $j = 1, \ldots, l$. Moreover, for any non negative integer m, we say that $f \in C^m_{\mathrm{H}}(\mathbb{G})$, if $X_{H_k} f$ exist and are continuous at each point of $\mathbb{G}$, for every horizontal k-derivation H_k such that $0 \le k \le m$.*

According to the paper [15] and for completeness, we report the following characterization of the Taylor polynomial.

Theorem 3.1.1. *Let m be a non negative integer, $c \in \mathbb{G}$, $f \in C^m_{\mathrm{H}}(\mathbb{G})$ and P a given polynomial of homogeneous degree m. The following facts are equivalent:*

i) *P is the m^{th} Taylor polynomial of f at c;*

ii) *$Z^i_\alpha(P - f)(c) = 0$ for any $i = 0, \ldots, m$ and for any $\alpha \in A_i$;*

iii) *$(P - f)(x) = o\big[d(x, c)\big]^m$ as $x \to c$.*

Proof. *i)* $\Longleftrightarrow$ *ii)* Follow immediately from Remark 3.1.1 *i)* $\Longrightarrow$ *iii)* Follows immediately from the stratified Taylor inequality proved in Theorem 1.42 of [62]. *iii)* $\Longrightarrow$ *i)* It suffices to verify that if Q is a polynomial of degree m

such that $Q(x) = o\big[d(x,0)\big]^m$ as $x \to 0$, than Q is identically zero. Indeed, let $Q = \sum_{d(I)\le m} a_I \, \mathtt{m}^I$: let us verify by induction that $a_I = 0$ for all I such that $d(I) \le m$. This is trivial when $d(I) = 0$; assuming that $a_I = 0$ for all I such that $d(I) = k < m$, we can write $Q = \sum_{d(I)=k+1} a_I \, \mathtt{m}^I + \sum_{k+1<d(I)\le m} a_I \, \mathtt{m}^I$. Fix $p \in \mathbb{G}$, $p \ne 0$. Then $\delta_\lambda(p) \to 0$ if and only if $\lambda \to 0$ and, recalling that d is homogeneous of degree one, we have

$$\frac{\sum_{d(I)=k+1} a_I \, \mathtt{m}^I(p)}{\big[d(p,0)\big]^{k+1}} = \frac{Q(\delta_\lambda(p))}{\big[d(\delta_\lambda(p),0)\big]^{k+1}} - \frac{\sum_{k+1<d(I)\le m} \lambda^{d(I)} a_I \, \mathtt{m}^I(p)}{\lambda^{k+1}\big[d(p,0)\big]^{k+1}} \to 0.$$

Thanks to the arbitrariness of p it follows that $a^I = 0$ for any I such that $d(I) = k+1$, as desired. To conclude the proof observe that if P' is the Taylor polynomial of f at the point c, then, denoting by τ_c the left translation pointed at c we have $(P - P') \circ \tau_c^{-1}(x) = o\big[d(x,0)\big]^m$ as $x \to 0$, i.e. $P = P'$. □

Thanks to the left invariance of the vector fields, for any $f \in C^m_{\mathbb{H}}(\mathbb{G})$, for any $c \in \mathbb{G}$ and for any $|I|$–derivation of the canonical basis of BPW, it results $X^I f(c) = X^I(f \circ \tau_c)(0)$ so, in the next sections, we will look for the Taylor polynomial at the identity of $\mathbb{G}$.

Example 3.1.2. The Heisenberg group $\mathbb{H}^n$ is the step two Carnot group associated to the Lie algebra $\mathfrak{h}^n = \mathfrak{h} \oplus \mathfrak{v}$ whose generators $X_1, \dots, X_n, Y_1, \dots, Y_n$, T satisfy the conditions $\mathfrak{h} = \mathrm{Span}_{\mathbb{R}}\left(\{X_j, Y_j\}_{j=1,\dots,n}\right)$, $\mathfrak{v} = \mathrm{Span}_{\mathbb{R}}\{T\}$ and the only non zero brackets are $[X_j, Y_j] = aT$, for any $j = 1, \dots, n$ and for some fixed real number $a \ne 0$. If $p = (x_1, \dots, x_n, y_1, \dots, y_n, t)$, $q = (x'_1, \dots, x'_n, y'_1, \dots, y'_n, t')$ $\in \mathbb{H}^n$, than the group product reads $p\,q = (x_1 + x'_1, \dots, x_n + x'_n, y_1 + y'_1, \dots, y_n + y'_n, t + t' + \frac{a}{2}\sum_{j=1}^n (x_j y'_j - x'_j y_j))$; it follows that $X_j = \frac{\partial}{\partial x_j} - \frac{a}{2} y_j \frac{\partial}{\partial t}$, $Y_j = \frac{\partial}{\partial y_j} + \frac{a}{2} x_j \frac{\partial}{\partial t}$ and $T = \frac{\partial}{\partial t}$. As in [63] we choose $a = -4$.

Let now recall the Taylor formula, in the setting of step two Carnot groups.

Definition 3.1.9. *Let $a_1, \dots, a_r$ be given elements of a (not necessarily commutative) ring A. Set*

$$\sigma(a_1, \dots, a_r) = \sum_{\pi \in S_r} a_{\pi(1)} \cdots a_{\pi(r)}, \qquad (\sigma)$$

where $\pi \in S_r$ denotes an element of the symmetric group over $\{1, \dots, r\}$.

Definition 3.1.9 is then useful to define the following differential operator. For any fixed $n_1, n_2, n_3 = 0, 1, \dots$, setting $X_1 = \cdots = X_{n_1} = X$, $X_{n_1+1} = \cdots =$

$X_{n_1+n_2} = Y$, $X_{n_1+n_2+1} = \cdots = X_{n_1+n_2+n_3} = T$, we define

$$\sigma(X^{n_1}, Y^{n_2}, T^{n_3}) = \sigma(X_1, \dots, X_{n_1+n_2+n_3}),$$

where, if some n_i is equal to zero, then the identity operator is omitted in the definition.

Then the Taylor polynomial in zero of a given $f \in C^m_{\mathbb{H}}(\mathbb{H}^1)$ is given as in the following definition.

Definition 3.1.10. *For any $f \in C^m_{\mathbb{H}}(\mathbb{H}^1)$, we set*

$$P^m_{f,0}(x,y,t) = \sum_{k=0}^{m} \Bigg[\sum_{\substack{n_1,n_2,n_3=0,1,\dots \\ n_1+n_2+2n_3=k}} \frac{\left(\frac{\sigma(X^{n_1},Y^{n_2},T^{n_3})}{(n_1+n_2+n_3)!} f\right)(0)}{n_1!n_2!n_3!} x^{n_1} y^{n_2} t^{n_3} \Bigg]. \tag{$P^m_{f,0}$}$$

Remark 3.1.2. Let $\mathrm{Sym}(X^{n_1}, Y^{n_2}, T^{n_3})$ be the sum of all k–iterated vector fields, each of them containing n_1 times the derivation X, n_2 times the derivation Y, and n_3 times the derivation T, then clearly $\mathrm{Sym}(X^{n_1}, Y^{n_2}, T^{n_3}) = \frac{\sigma(X^{n_1},Y^{n_2},T^{n_3})}{n_1!n_2!n_3!}$. It results

$$\begin{aligned} \mathrm{Sym}(X^{n_1}, Y^{n_2}) &= \mathrm{Sym}(X^{n_1-1}, Y^{n_2})\, X + \mathrm{Sym}(X^{n_1}, Y^{n_2-1})\, Y = \\ &= X\, \mathrm{Sym}(X^{n_1-1}, Y^{n_2}) + Y\, \mathrm{Sym}(X^{n_1}, Y^{n_2-1}). \end{aligned} \tag{S′}$$

Observe then that is natural to see the operator containing n_1 times the derivation X, n_2 times the derivation Y and n_3 times the derivation T, where $n_1 + n_2 + 2n_3 = k$, as a unique $(n_1 + n_2 + n_3)^{th}$–iterated derivation, more precisely as $\frac{n_1!n_2!n_3!}{(n_1+n_2+n_3)!} \cdot \mathrm{Sym}(X^{n_1}, Y^{n_2}, T^{n_3})$; so the following position appears quite natural. Set, for any non negative integers k, n_1, n_2, n_3 such that $n_1 + n_2 + 2n_3 = k$,

$$\frac{\partial^k}{\partial X^{n_1} \partial Y^{n_2} \partial T^{n_3}} = \frac{n_1!n_2!n_3!}{(n_1 + n_2 + n_3)!} \cdot \mathrm{Sym}(X^{n_1}, Y^{n_2}, T^{n_3}). \tag{3.1}$$

We call (3.1) a symmetrized $(n_1 + n_2 + n_3)$–derivation of order k containing n_1–times X, n_2–times Y and n_3–times T (symmetrized derivation for short, when we do not need to specify either the order k or the single vector fields). To indicate a symmetrized derivation of degree $d(\alpha)$ we simply write ∂^α, if it is not necessary to specify the vector fields involved. Observe that, according to (S′), the following property for symmetrized derivations holds. For any

$n_1, n_2 = 1, 2, \ldots$, it results

$$\begin{aligned}\frac{\partial^{n_1+n_2}}{\partial X^{n_1}\partial Y^{n_2}} &= \frac{n_1}{n_1+n_2}\cdot\frac{\partial^{n_1+n_2-1}}{\partial X^{n_1-1}\partial Y^{n_2}}X + \frac{n_2}{n_1+n_2}\cdot\frac{\partial^{n_1+n_2-1}}{\partial X^{n_1}\partial Y^{n_2-1}}Y = \\ &= \frac{n_1}{n_1+n_2}\cdot X\,\frac{\partial^{n_1+n_2-1}}{\partial X^{n_1-1}\partial Y^{n_2}} + \frac{n_2}{n_1+n_2}\cdot Y\,\frac{\partial^{n_1+n_2-1}}{\partial X^{n_1}\partial Y^{n_2-1}}. \qquad \text{(S'')}\end{aligned}$$

Observe that, since T belongs to the center of $\mathfrak{h}^1$, the differential operator $\mathrm{Sym}(X^{n_1}, Y^{n_2}, T^{n_3})$ satisfies the relation $\mathrm{Sym}(X^{n_1}, Y^{n_2}, T^{n_3}) = \binom{n_1+n_2+n_3}{n_3}\cdot \mathrm{Sym}(X^{n_1}, Y^{n_2})\ T^{n_3}$ or, equivalently, $\frac{\partial^k}{\partial X^{n_1}\partial Y^{n_2}\partial T^{n_3}} = \frac{\partial^{n_1+n_2}}{\partial X^{n_1}\partial Y^{n_2}}T^{n_3}$. Indeed, elements $\underbrace{X, \ldots, X}_{n_1}, \underbrace{Y, \ldots, Y}_{n_2}$ can be permuted in $(n_1+n_2)!$ ways, and for any of these ways we can insert n_3 fields T: fixed a way, we can insert first T in (n_1+n_2+1) different ways, the second in (n_1+n_2+2) different ways, and so on, n_3-th in $(n_1+n_2+n_3)$ ways; it follows that in each element of $\sigma\left(X^{n_1}, Y^{n_2}\right)$, after having inserted all the n_3 T, we obtain one element of $\sigma\left(X^{n_1}, Y^{n_2}, T^{n_3}\right)$, and since T commute, we have

$$(n_1+n_2+1)\cdot\ldots\cdot(n_1+n_2+n_3)\sigma\left(X^{n_1}, Y^{n_2}\right)T^{n_3} = \sigma\left(X^{n_1}, Y^{n_2}, T^{n_3}\right),$$

i.e.,

$$\sigma\left(X^{n_1}, Y^{n_2}, T^{n_3}\right) = \frac{(n_1+n_2+n_3)!}{(n_1+n_2)!}\sigma\left(X^{n_1}, Y^{n_2}\right)T^{n_3},$$

and, finally,

$$\frac{\partial^k}{\partial X^{n_1}\partial Y^{n_2}\partial T^{n_3}} = \frac{\partial^{n_1+n_2}}{\partial X^{n_1}\partial Y^{n_2}}T^{n_3}.$$

So, according to previous definition we can rewrite as follows

$$P^m_{f,0}(x,y,t) = \sum_{k=0}^{m}\left[\sum_{\substack{n_1,n_2,n_3=0,1,\ldots\\ n_1+n_2+2n_3=k}} \frac{\left(\frac{\partial^{n_1+n_2}}{\partial X^{n_1}\partial Y^{n_2}}T^{n_3}f\right)(0)}{n_1!n_2!n_3!}\,x^{n_1}y^{n_2}t^{n_3}\right] \qquad (3.2)$$

We stress that the set $\bigcup_{k=0}^{\infty}\mathfrak{S}_k$, where

$$\mathfrak{S}_k = \left\{\frac{\partial^{n_1+n_2}}{\partial X^{n_1}\partial Y^{n_2}}\,T^{n_3} : n_1, n_2, n_3 = 0, 1, \ldots,\ n_1+n_2+2n_3 = k\right\},$$

forms a basis for the vector space of all left invariant differential operators in $\mathbb{H}^1$.

Through definitions above then it is possible – see proposition below – to show that the linear mapping $D : \mathcal{P}_m \to \mathcal{D}_m$ defined through the position

$$D(P_m) = \sum_{k=0}^{m} \sum_{\substack{n_1,n_2,n_3=0,1,\ldots \\ n_1+n_2+2n_3=k}} \left(\frac{\partial^{n_1+n_2}}{\partial X^{n_1} \partial Y^{n_2}} T^{n_3} P_m(0) \right) \frac{\partial^{n_1+n_2}}{\partial X^{n_1} \partial Y^{n_2}} T^{n_3}, \ \forall P_m \in \mathcal{P}_m,$$

is an isomorphism having a diagonal matrix associated with respect to the two basis $\left\{x^{n_1} y^{n_2} t^{n_3}\right\}_{\substack{n_1,n_2,n_3=0,1,\ldots \\ n_1+n_2+2n_3=0,1,\ldots,m}}$ and $\bigcup_{k=0}^{m} \mathfrak{S}_k$. Roughly speaking the two bases are essentially dual to each other, i.e. any such a symmetrized derivation acts on monomials as if it were an Euclidean iterated derivative. We stress that this result is not at all trivial if one takes into account the very definition of symmetrized derivation.

In what follows we finally collect main property of the symmetrized Taylor polynomial.

Proposition 3.1.3. *For any $n_1, n_2, n_3, m_1, m_2, m_3 = 0, 1, \ldots$, it results*

$$\left[\frac{\partial^{n_1+n_2}}{\partial X^{n_1} \partial Y^{n_2}} T^{n_3} (x^{m_1} y^{m_2} t^{m_3}) \right]_{(x,y,t)=(0,0,0)} =$$

$$\begin{cases} m_1! m_2! m_3! & \text{if } n_i = m_i, \ i = 1,2,3 \\ 0 & \text{in any other case} \end{cases}.$$

Theorem 3.1.2. *Let m be a non negative integer and let $f \in C^m_{\mathbb{H}}(\mathbb{H}^1)$. Then $P^m_{f,0}$ is the Taylor polynomial of f at zero.*

Remark. With very minimal changes all previous considerations work in $\mathbb{H}^n$.

For the sake of clarity we explicitly write some symmetrized operators, according to previous definitions.

Example 3.1.3.

$$\sigma(X) = X;$$

$$\sigma(X, Y) = XY + YX;$$

$$\begin{aligned} \sigma\left(X^2, Y\right) &= XXY + XYX + XYX + YXX + YXX \\ &= 2X^2Y + 2XYX + 2YX^2; \end{aligned}$$

$$\begin{aligned}\sigma\left(X^2,Y\right) &= XYY+XYY+YXY\\ &\quad +YYX+YXY+YYX\\ &= 2XY^2+2YXY+2Y^2X;\end{aligned}$$

$$\begin{aligned}\sigma\left(X^3\right) &= XXX+XXX+XXX\\ &\quad +XXX+XXX+XXX\\ &= 6X^3;\end{aligned}$$

$$\begin{aligned}\sigma\left(X,T\right) &= XT+TX\\ &= 2XT.\end{aligned}$$

$$\begin{aligned}\frac{\partial^1}{\partial X} &= X;\\ \frac{\partial^2}{\partial X\partial Y} &= \frac{XY+YX}{2!};\\ \frac{\partial^3}{\partial X^2\partial Y} &= 2\frac{X^2Y+XYX+YX^2}{3!};\\ \frac{\partial^3}{\partial X\partial Y^2} &= 2\frac{XY^2+YXY+Y^2X}{3!};\\ \frac{\partial^3}{\partial X^3} &= 6\frac{X^3}{3!};\\ \frac{\partial^3}{\partial X\partial T} &= 2\frac{XT}{2!}.\end{aligned}$$

Although we give a detailed sketch of the proof of Whitney extension theorem in the setting of the first group $\mathbb{H}^1$, referring to the paper [64] for the general case of step two Carnot groups, we briefly recall Taylor polynomial in this last setting. Let $\mathbb{G}$ be a step 2 Carnot group, $\{X_1, \ldots, X_l\}$ a basis of V_1 and $\{T_1, \ldots, T_p\}$ a basis of V_2. Indeed, arguing as before, we can recover the Taylor polynomial starting from the Taylor expansion related to a family of $l+p$ left invariant vector fields; more precisely, observing that V_2 is contained in the center of $\mathfrak{g}$, we define the symmetrized derivation with respect to the vector fields $X_1, \ldots, X_l, T_1, \ldots, T_p$ according to the following definition.

Definition 3.1.11. *Set, for any non negative integers $k, n_1, \dots, n_l, s_1, \dots, s_p$, such that, $k = n_1 + \ldots + n_l + 2\,(s_1 + \ldots + s_p)$,*

$$\frac{\partial^{n_1+\cdots+n_l}}{\partial X_1^{n_1}\cdots\partial X_l^{n_l}} T_1^{s_1}\cdots T_p^{s_p} := \frac{n_1!\cdots n_l!}{(n_1+\cdots+n_l)!}\cdot \mathrm{Sym}(X_1^{n_1},\dots,X_l^{n_l})T_1^{s_1}\cdots T_p^{s_p}. \tag{3.3}$$

We call (3.3) *a* symmetrized k-derivation of order k *containing n_1-times X_1,. . .,n_l-times X_l and s_1-times T_1,. . . s_p-times T_p (symmetrized derivation for short, when we do not need to specify either the order k or the single vector fields).*

Remark 3.1.3. Then formulas become respectively,

$$\begin{aligned}\mathrm{Sym}(X_1^{n_1},\dots,X_l^{n_l}) &= \\ &= \mathrm{Sym}(X_1^{n_1-1},\dots,X_l^{n_l})X_1 + \cdots + \mathrm{Sym}(X_1^{n_1},\dots,X_l^{n_l-1})X_l, \end{aligned} \tag{S'$_2$}$$

and

$$\begin{aligned}\frac{\partial^{n_1+\cdots+n_l}}{\partial X^{n_1}\cdots\partial X^{n_l}} &= \frac{n_1}{n_1+\cdots+n_l}\cdot\frac{\partial^{n_1+\cdots+n_l-1}}{\partial X^{n_1-1}\cdots\partial X^{n_l}}\,X_1 + \\ &+\cdots\cdots\cdots\cdots\cdots\cdots\cdots+ \\ &+ \frac{n_l}{n_1+\cdots+n_l}\cdot\frac{\partial^{n_1+\cdots+n_l-1}}{\partial X^{n_1}\cdots\partial X^{n_l-1}}\,X_l;\end{aligned} \tag{S''$_2$}$$

as before, the first one is justified by the multinomial identity

$$\frac{(n_1+\cdots+n_l)!}{n_1!\cdots n_l!} = \frac{(n_1+\cdots+n_l-1)!}{(n_1-1)!\cdots n_l!} + \cdots\cdots + \frac{(n_1+\cdots+n_l-1)!}{n_1!\cdots(n_l-1)!},$$

while the second one follows immediately by the first one after some easy calculations.

In this case, the right formula for the Taylor polynomial for a given $f \in C^m_{\mathbb{H}}(\mathbb{G})$, comes to be

$$P^m_{f,0}(x) = \sum_{k=0}^{m}\Bigg[\sum_{d(I)=k}\frac{\Big(\frac{\partial^{n_1+\cdots+n_l}}{\partial X_1^{n_1}\cdots\partial X_l^{n_l}}T_1^{m_1}\cdots T_p^{m_p} f\Big)(0)}{n_1!\cdots n_l! m_1!\cdots m_p!}\,x^I\Bigg], \tag{$P^m_{f,0}$}$$

where, for any $k = 0, \dots, m$, $I = (n_1, \dots, n_l, m_1, \dots, m_p)$ denotes a general $(l+p)$–tuple of non negative integers, $d(I) = n_1 + \ldots + n_l + 2\,(m_1 + \ldots + m_p)$ and $x^I = x_1^{n_1}\cdots x_l^{n_l} t_1^{m_1}\cdots t_p^{m_p}$. So, as in the $\mathbb{H}^1$ case, it remains to check that $P^m_{f,0}$ is actually the m^{th} Taylor polynomial by verifying *ii)* of Theorem 3.1.1.

Analogously, if we set,

$$\mathfrak{S}_k = \Big\{ \frac{\partial^{n_1+\cdots+n_l}}{\partial X_1^{n_1}\cdots\partial X_l^{n_l}} T_1^{s_1}\cdots T_p^{s_p} :$$
$$n_1,\ldots,n_l,s_1,\ldots,s_p = 0,1,\ldots,\ n_1+\cdots+n_l+2\,(s_1+\cdots+s_p) = k\Big\}.$$

with $\mathfrak{S}_0 = \{\mathrm{id}\}$, it is possible to check that the set $\bigcup_{k=0}^m \mathfrak{S}_k$ is a basis of the vector space of all differential operators on $\mathbb{G}$ and that the Taylor polynomial satisfy *ii*) of Theorem 3.1.1. To this aim, it suffices to extend Proposition 3.1.3.

Proposition 3.1.4. *For any* $n_1,\ldots,n_l,s_1,\ldots,s_p,m_1,\ldots,m_l,r_1,\ldots,r_p$, *we have,*

$$\left[\frac{\partial^{n_1+\cdots+n_l}}{\partial X^{n_1}\cdots\partial X^{n_l}} T_1^{s_1}\cdots T_p^{s_p}(x_1^{m_1}\cdots x_l^{m_l}t_1^{r_1}\cdots t_p^{r_p})\right]_{(x_1,\ldots,x_l,t_1,\ldots,t_p)=(0,\ldots,0)} =$$
$$= \begin{cases} m_1!\cdots m_l!\,r_1!\cdots r_p! & \text{if } n_i = m_i,\ i=1,\ldots,l \text{ and } s_i = r_i,\ i=1,\ldots,p \\ 0 & \text{in any other case} \end{cases}.$$

We conclude this preliminary part with the following interesting proposition [64].

Proposition 3.1.5. *Let* $f \in C^m(\mathbb{H}^1)$, $m \in \mathbb{N}$, $m \geq 1$. *Then for each k-derivation* Z_{J_k}, $J_k = \{i_1,\ldots,i_s,j_1,\ldots,j_r\}$, $k = r+s$, *with* $d(Z_{J_k}) = s+2r \leqslant m$, *we have, in* $(x,y,t) = (0,0,0)$,

$$Z_{J_k}P^m_{f,0}(x,y,t) = P^{m-d(Z_{J_k})}_{Z_{J_k}f,0}(x,y,t),$$

for each $(x,y,t) \in \mathbb{H}^1$, *and by invariance, for each* $c \in \mathbb{H}^1$.

Proof. The proof is by induction on the degree $m \in \mathbb{N}$, because it is sufficient to check it for the fields X,Y,T. Let us verify the proposition for the cases $m = 2,3$, recalling that $X = \partial_1 + 2y\partial_t$, $Y = \partial_2 - 2x\partial_t$ and $T = -\frac{1}{4}(XY - YX)$. For $m = 2$ we have

$$P^2_{f,0}(x,y,t) = f(0) + Xf(0)x + Yf(0)y + \frac{(XY+YX)f(0)}{2}xy$$
$$+ \frac{X^2f(0)}{2}x^2 + \frac{Y^2f(0)}{2}y^2 + Tf(0)t,$$

so that

$$\begin{aligned}
XP^2_{f,0}(x,y,t) &= Xf(0) + \frac{(XY+YX)f(0)}{2}y + X^2f(0)x + 2yTf(0)\\
&= Xf(0) + \frac{XYf(0)}{2}y + \frac{YXf(0)}{2}y + X^2f(0)x\\
&- \frac{XYf(0)}{2}y + \frac{YXf(0)}{2}y\\
&= Xf(0) + X^2f(0)x + YXf(0)y\\
&= P^1_{Xf,0}(x,y,t),
\end{aligned}$$

and similarly,

$$\begin{aligned}
YP^2_{f,0}(x,y,t) &= Yf(0) + \frac{XYf(0)}{2}x + \frac{XYf(0)}{2}x + Y^2f(0)y - 2xTf(0)\\
&= Yf(0) + XYf(0)x + \frac{YXf(0)}{2}x + Y^2f(0)y\\
&+ \frac{XYf(0)}{2}x - \frac{YXf(0)}{2}x\\
&= Yf(0) + XYf(0)x + Y^2f(0)y\\
&= P^1_{Yf,0}(x,y,t);
\end{aligned}$$

we trivially also have $TP^2_{f,0}(x,y,t) = Tf(0) = P^0_{Tf,0}(x,y,t)$. Then, for $m = 3$ we have

$$\begin{aligned}
P^2_{f,0}(x,y,t) = f(0) + Xf(0)x + Yf(0)y + \frac{X^2f(0)}{2}x^2 + \frac{Y^2f(0)}{2}y^2 + Tf(0)t +\\
\frac{(XY+YX)f(0)}{2}xy + XTf(0)xt + YTf(0)yt + \frac{X^3f(0)}{3!}x^3\\
+ \frac{Y^3f(0)}{3!}y^3 + \frac{X^2f(0)}{3!}x^2y + \frac{XYXf(0)}{3!}x^2y + \frac{YX^2f(0)}{3!}x^2y\\
+ \frac{Y^2f(0)}{3!}y^2x + \frac{YXYf(0)}{3!}y^2x + \frac{XY^2f(0)}{3!}Y^2x,
\end{aligned}$$

and, setting $g = Xf$,

$$\begin{aligned}
XP^3_{f,0}(x,y,t) = {} & g(0) + Xg(0)x + 2Tf(0)y + \frac{XYf(0)}{2}y \\
& + \frac{YXf(0)}{2}y + XTf(0)t + 2XTf(0)xy + 2YTf(0)y^2 \\
& + \frac{X^2g(0)}{2}x^2 + \frac{X^2Yf(0)}{3}xy = \\
& \frac{XYXf(0)}{3}xy + \frac{YYXf(0)}{3}xy + \frac{Y^2g(0)}{3!}y^2 + \frac{YXYf(0)}{3!}y^2 \\
& + \frac{XY^2f(0)}{3!}y^2 + g(0) + Xg(0)x + \frac{YXf(0)}{2}y \\
& - \frac{XYf(0)}{2}y + \frac{XYf(0)}{2} + \frac{YXf(0)}{2}y + Tg(0)t\frac{X^2g(0)}{2}x^2 \\
& + \frac{XXYf(0)}{3}xy + \frac{XYXf(0)}{3}xy + \frac{YXXf(0)}{3}xy + 2XTf(0)xy \\
& + 2YTf(0)y^2 + \frac{YYXf(0)}{3!}y^2\frac{YXYf(0)}{3!}y^2 + \frac{XY^2f(0)}{3!}y^2,
\end{aligned}$$

just observing that $\frac{XYf(0)}{2} - \frac{XYf(0)}{2} + \frac{XYf(0)}{2} + \frac{YXf(0)}{2} = Yg(0)$; for the others it must then be checked that (using fields for short):

i) $\frac{YYX}{2} - \frac{YXY}{2} + \frac{YYX}{6} + \frac{YXY}{6} + \frac{XYY}{6} = \frac{YYX}{2}$;

ii) $\frac{XXY}{3} + \frac{XYX}{3} + \frac{YXX}{3} + \frac{XYX}{2} - \frac{XXY}{2} = \frac{XYX}{2} + \frac{YXX}{2}$.

For *i)* we have:

$$\begin{aligned}
& \frac{YYX}{6} + \frac{YXY}{6} + \frac{XYY}{6} = \frac{YXY}{2} \\
& \Longleftrightarrow YYX + YXY + XYY = 3YXY \\
& \Longleftrightarrow YYX - YXY + XYY - YXY = 0 \\
& \Longleftrightarrow Y(4T) + (-4T)Y = 4YT - 4YT = 0;
\end{aligned}$$

for *ii)* we have:

$$\begin{aligned}
& 2XXY + 2XYX + 2YXX + 3XYX - 3XXY = 3XYX + 3YXX \\
& \Longleftrightarrow -XXY + 2XYX - YXX = 0 \\
& \Longleftrightarrow XYX - XXY + XYX - YXX = 0 \\
& \Longleftrightarrow X(4T) + (-4T)X = 4XT - 4XT = 0.
\end{aligned}$$

In conclusion, previous checks, together with the obvious

$$TP^3_{f,0}(x,y,z) = P^1_{Tf,0}(x,y,z),$$

give basis for induction. Similarly too $YP^3_{f,0}(x,y,z) = P^2_{Yf,0}(x,y,z)$ holds. $\square$

3.2 Whitney's extension theorem

In order to prove Whitney's theorem we need some preliminary notation.

Definition 3.2.1 (Symmetrized m-jet). *Let $F \subseteq \mathbb{H}^1$, $m \in \mathbb{N}$, $\alpha = (n_1, n_2, n_3)$ with $n_1, n_2, n_3 \in \mathbb{N}$ and $d(\alpha) = n_1 + n_2 + 2n_3$. An m-jet on F is a collection $J^F = \{f^\alpha\}_{0 \leqslant d(\alpha) \leqslant m}$ of functions $f_\alpha : F \to \mathbb{R}$. An m-jet is of class $C^m(F)$ if $f_\alpha \in C^{m-d(\alpha)}(F)$ for all α and $0 \leq d(\alpha) \leq m$. Finally, if $f \in C^m(F)$, a symmetrized m-jet is the collection $J^F_{\mathbb{H},f} = \{\partial^\alpha f\}_{0 \leqslant d(\alpha) \leqslant m}$.*

Definition 3.2.2 (Taylor polynomial of an m-jet). *Let $F \subseteq \mathbb{H}^1$, $m \in \mathbb{N}$ and $J^F = \{f^\alpha\}_{0 \leqslant d(\alpha) \leqslant m}$ an m-jet on F. Set*

$$P^m_{J^F,c}(p) = \sum_{0 \leqslant d(\alpha) \leqslant m} \frac{f^\alpha(c)}{\alpha!} (c^{-1}p)^\alpha,$$

for all $c \in F$, $p \in \mathbb{H}^1$. We say that $P^m_{J^F,f}$ is the Taylor polynomial associated to the m-jet J^F. Analogous definition for the Taylor polynomial $P^m_{J^F_{\mathbb{H},f},c}$ associated to the m-jet $J^F_{\mathbb{H},f} = \{\partial^\alpha f\}_{0 \leqslant d(\alpha) \leqslant m}$ of a given function $f \in C^m(F)$.

We now give a sketch of the proof of Whitney's extension theorem in $\mathbb{H}^1$, referring to [64] for the general proof in the case of step two Carnot groups.

Theorem 3.2.1 (Whitney's extension theorem). *Let $F \subseteq \mathbb{H}^1$ be closed, $f : F \to \mathbb{R}$ and $m \in \mathbb{N}$. Then the following facts are equivalent:*

1) *there exists $J^F = \{f^\alpha\}_{0 \leqslant d(\alpha) \leqslant m}$ m-jet on F such that:*

   i) *$f(p) = f^\alpha(p)$ for each $p \in F$, $\alpha = (0,0,0)$;*

   ii) *for each $c \in F$ and for each compact $K \subseteq F$, the condition*

   $$f^\alpha(p) - \partial^\alpha P^m_{J^F,c}(p) = o\left((d(p,c))^{m-d(\alpha)} \right),$$

   holds for all α, $0 \leq d(\alpha) \leq m$ and uniformly for all $p, c \in K$.

2) *there exist* $\bar{f} : \mathbb{H}^1 \to \mathbb{R}$, $\bar{f} \in C^m_{\mathbb{H}}(\mathbb{H}^1)$, *such that the condition*

$$\partial^\alpha \bar{f}(p) = \partial^\alpha f(p),$$

holds for all $p \in F$ *and for all* α *with* $0 \leq d(\alpha) \leq m$.

In particular, if $f \in C^m(F)$, *then* 2) *holds.*

Proof. Clearly *2)* $\Longrightarrow$ *1)*, indeed just assume $J^F = J^F_{H,f}$ and $P^m_{J^F,c} = P^m_{c,f}$, so the thesis follows by properties of Taylor polynomial and the invariance of the fields.
Also, if $f \in C^m(F)$, obviously *i)* and *ii)* are true, and the thesis follows by *1)* $\Longrightarrow$ *2)*, that now we prove.
1) $\Longrightarrow$ *2)*. Let $F \subseteq \mathbb{H}^1$ be a closed set and, for all $p = (x, y, z)$, let

$$\|p\| = \left((x^2 + y^2)^2 + t^2\right)^{1/4},$$

the Korányi-Cygan norm in $\mathbb{H}^1$, whose induced distance defined by $\bar{d}(p, q) = \|q^{-1}p\|$, for each $p, q \in \mathbb{H}^1$, is equivalent to the Carnot-Carathéodory metric; then, according to Lemma 2.2.1 of Chapter 2, let $g \in C^\infty([0, +\infty[)$ such that $g(t) = 1$ for all $t \leq 1$, $g(t) = 0$ for all $t \geq l$, g decreasing, and let $\{\sigma_i\}_{i \in I}$ be defined as $\sigma_i(p) = g\left(\frac{\bar{d}(p,p_i)}{\alpha r_{p_i}}\right)$, with $\{p_i\}_{i \in I}$ the maximal family of points in $\mathbb{H}^1 \setminus F$ as in Lemma, and finally, let $\{\varphi_i\}_{i \in I}$ be the partition of unity, with $r_p = \bar{d}(p, F)$. Because the fields X and Y are homogeneous of degree 1, and T is homogeneous of degree 2, then ∂^α, with $0 \leq d(\alpha) \leq m$, is homogeneous of degree $d(\alpha)$: being $p \mapsto \|p\|$ a function of class $C^\infty(\mathbb{R}^3 \setminus \{(0, 0, 0)\}) \subseteq C^m_{\mathbb{H}}(\mathbb{H}^1 \setminus \{(0, 0, 0)\})$, and homogeneous in $\mathbb{H}^1$ of degree 1, it follows that $p \mapsto \partial^\alpha(\|p\|)$ is homogeneous of degree $1 - d(\alpha)$, and standard considerations guarantee that there exists a constant $C(\alpha)$ such that $|\partial\varphi_i(p)| \leqslant \frac{C(\alpha)}{r_p^{d(\alpha)}}$, for all $p \in \mathbb{H}^1 \setminus F$ and for all α such that $0 \leq d(\alpha) \leq m$.

So let *i)* and *ii)* be true, and define

$$\bar{f}(p) = \begin{cases} f(p) \quad if \quad p \in F \\ \sum\limits_{i \in I} \varphi_i(p) P^m_{J^F, c_i}(p) \quad if \quad p \in \mathbb{H}^1 \setminus F \end{cases},$$

with $c_i \in F$ such that $d(p_i, c_i) = d(p_i, F)$ for all $i \in I$; to achieve the thesis it suffices to verify that:

(F) for all $c \in F$, for all polynomial $P^m_{J^F,c}$ and for all compact set $K \subseteq \mathbb{H}^1$, the condition

$$\partial^\alpha \bar{f}(p) - \partial^\alpha P^m_{J^F,c}(p) = o\left(\left(d(p,c)^{m-d(\alpha)}\right)\right),$$

holds for all α such that $0 \le d(\alpha) \le m$ and uniformly for $p, c \in K$.

Note that also the following condition trivially holds because $\bar{f} \in C^\infty(\mathbb{R}^3 \setminus F)$:

($\mathbb{H}^1 \setminus F$) for all $c \in \mathbb{H}^1 \setminus F$, for all polynomial $P^m_{J^F_{\mathrm{H},\bar{f}},c}$ and for all compact set $K \subseteq \mathbb{H}^1 \setminus F$, the condition

$$\partial^\alpha \bar{f}(p) - \partial^\alpha P^m_{J^F_{\mathrm{H},\bar{f}},c}(p) = o\left(\left(d(p,c)^{m-d(\alpha)}\right)\right),$$

holds for all α such that $0 \le d(\alpha) \le m$ and uniformly for $p, c \in K$.

So, if *(F)* holds, it is

$$\begin{aligned}\partial^\alpha \bar{f}(p) - f^\alpha(p) &= \left(\partial^\alpha \bar{f}(p) - \partial^\alpha P^m_{J^F,c}(p)\right) + \left(\partial^\alpha P^m_{J^F,c}(p) - f^\alpha(p)\right)\\ &= o\left(\left(d(p,c)^{m-d(\alpha)}\right)\right),\end{aligned}$$

for all α such that $0 \le d(\alpha) \le m$ and uniformly in K: in particular $\bar{f}(p) = f(p)$ and *2)* is proved.

So let us verify *(F)*. Let $p \in \mathbb{H}^1$: if $p \in F$, $p \to c$, uniformly for $p, c \in K$, the condition *(F)* follows from *i)* and *ii)*. Let $p \in \mathbb{H}^1 \setminus F$: let us verify condition *(F)* for $\alpha = 0$, i.e. $\bar{f}(p) - P^m_{J^F,c}(p) = o\left((d(p,c)^m)\right)$. To this aim, note that $\bar{f}(p) - P^m_{J^F,c}(p) = \sum_{i \in I} \varphi_i(p)\left(P^m_{J^F,c_i}(p) - P^m_{J^F,c}(p)\right)$: let verify that, for a given $q \in F$, then $\partial^\alpha P^m_{J^F,q}(p) - \partial^\alpha P^m_{J^F,c}(p) = o\left((d(p,c) + d(c,q))^{m-d(\alpha)}\right)$. Indeed, if $h(p) = P^m_{J^F,q}(p) - P^m_{J^F,c}(p)$, for all $p \in \mathbb{H}^1$ and $q, c \in F$ so, writing $h(p) = \sum_{0 \leqslant d(\alpha) \leqslant m} \frac{\partial^\alpha h(c)}{\alpha!}(c^{-1}p)^\alpha$ we have

$$\begin{aligned}\partial^\alpha h(x) &= \partial^\alpha \left(P^m_{J^F,q}(p) - P^m_{J^F,c}(p)\right)\Big|_{p=c}\\ &= \partial^\alpha P^m_{J^F,q}(c) - f^\alpha(c)\\ &= o\left(d(p,c)^{m-d(\alpha)}\right),\end{aligned}$$

and then,

$$\begin{aligned}
\left|P^m_{J^F,q}(p) - P^m_{J^F,c}(p)\right| &= \left|\sum_{0\leqslant d(\alpha)\leqslant m} \frac{\partial^\alpha h(c)}{\alpha!}(c^{-1}p)^\alpha\right| \\
&\leq \sum_{0\leqslant d(\alpha)\leqslant m} o(1)\frac{d(c,q)}{\alpha!}^{m-d(\alpha)}\left|c^{-1}p\right|^{d(\alpha)} \\
&\leqslant o(1)\sum_{0\leqslant d(\alpha)\leqslant m} \frac{m!}{\alpha!}d(c,q)^{m-d(\alpha)}d(p,c)^{d(\alpha)} \\
&\leqslant o(1)\sum_{0\leqslant |\alpha|\leqslant m} \frac{m!}{\alpha!}d(c,q)^{m-|\alpha|}d(p,c)^{|\alpha|} \\
&\leqslant o(1)\left((d(p,c)+d(c,q))^m\right),
\end{aligned}$$

i.e., $P^m_{J^F,q}(p) - P^m_{J^F,c}(p) = o\left((d(p,c)+d(c,q))^m\right)$; the general case follows from Proposition 3.1.5. By Lemma 2.2.1 of Chapter 2, there exists an absolute constant $\bar{C}$ such that $d(q,c) \leq \bar{C}d(p,c)$, so we obtain, for $\alpha = 0$, the required condition $\bar{f}(p) - P^m_{J^F,c}(p) = o\left((d(p,c)^m)\right)$.

Let now $d(\alpha) = 1$: then we have $\partial^\alpha \bar{f}(p) - \partial^\alpha P^m_{J^F,c}(p) = o\left(d(p,c)^{m-1}\right)$; indeed

$$\partial^\alpha \bar{f}(p) = \sum_{i\in I}\varphi_i(p)\partial^\alpha P^m_{J^F,c_i}(p) + \sum_{i\in I}\partial^\alpha\varphi_i(p)P^m_{J^F,c_i}(p),$$

so that

$$\begin{aligned}
\partial^\alpha \bar{f}(p) - \partial^\alpha P^m_{J^F,c}(p) &= \sum_{i\in I}\varphi_i(p)\left(\partial^\alpha P^m_{J^F,c_i}(p) - \partial^\alpha P^m_{J^F,c}(p)\right) \\
&\quad + \sum_{i\in I}\partial^\alpha\varphi_i(p)\left(P^m_{J^F,c_i}(p) - P^m_{J^F,q}(p)\right) \\
&= o\left(d(p,c)^{m-1}\right).
\end{aligned}$$

Analogous considerations hold for $2 \leq d(\alpha) \leq m$, completing the proof. $\square$

Appendix A

First insights into Menger convexity

A.1 Betweenness in metric spaces

Convexity is a concept usually developed in linear spaces. However since the seventies of the 20th century, convexity is also study in metric spaces in general. It is a well known fact that the theory of convex sets in linear spaces uses the concept of convex function, which is very useful in applications. In the literature, therefore, there are many generalizations of the notion of convex sets, which want to generalize what happens in linear spaces if we consider for example spaces without an algebraic structure such the distance spaces and in particular the metric spaces. Then it is natural to expect that the analogue of linear convexity for a set allow us to define also an analogue of convex function in a more general framework.

We present various types of convexity of a space according to different ideas. Let us start clarifying few necessary definitions and try to keep them as general as possible.

Definition A.1.1 (Betweenness). *Let (X, ρ) be a distance space. If $x, y, z \in X$ are three pairwise distinct points, z lies between x and y if we have*

$$\rho(x, y) = \rho(x, z) + \rho(z, y).$$

This relation usually is called "betweenness relation" and can also be represented for convenience by writing xzy.

The notion of betweenness was introduced by K. Menger in the context of distance geometry [16], where distance is the only primitive notion and all others relations are defined explicitly in terms of distance [65].

In the general context of distance spaces the betweenness relation is symmetric, in the sense that if z lies between x and y, then z lies between y and x, as specified by the next proposition together with another property.

Proposition A.1.1. *Let (X, ρ) be a distance space. Then the relation of betweenness has the following properties:*

1) *if xzy, then yzx for all $x, y, z \in X$;*
2) *if xzy, then neither xyz nor zxy for all $x, y, z \in X$;*

Proof.

1) Assuming xzy holds, by definition

$$\rho(x, y) = \rho(x, z) + \rho(z, y)$$

holds, and by symmetry of distance we have

$$\rho(y, x) = \rho(z, x) + \rho(y, z),$$

that is yzx;

2) Assuming xzy holds, by definition

$$\rho(x, y) = \rho(x, z) + \rho(z, y)$$

and $x \neq z \neq y$ hold; by xyz we have also

$$\rho(x, z) = \rho(x, y) + \rho(y, z)$$

and $x \neq y \neq z$, so addition of these two equalities gives $\rho(z, y) = 0$, which contradicts $z \neq y$. Similarly also zxy cannot holds. □ Instead the following proposition expresses a transitivity property for betweenness relation but valid in the specific case of metric spaces.

Proposition A.1.2. *Let (X, d) be a metric space. Then we have the following equivalence:*

$$xzy \text{ and } xyt$$

$$\Updownarrow$$

$$xzt \text{ and } zxy$$

for all $x, y, z, t \in X$.

Proof. Assuming xzy and xyt hold, because each triple consists of pairwise distinct points, the points x, z, y, t are pairwise distinct except for, perhaps the pair z, t. But it is $z \neq t$, for in the contrary case xzy and xyz both hold contradicting property 2) of proposition A.1.1. Then by definition we have

$$d(x,z) + d(z,y) = d(x,y),$$

and

$$d(x,y) + d(y,t) = d(x,t),$$

so by addition and triangular inequality

$$\begin{aligned} d(x,z) + d(z,y) + d(y,t) &= d(x,t) \\ &\leq d(x,z) + d(z,t) \end{aligned}$$

holds, that together with triangular inequality applied to points z, y, t shows that

$$d(z,y) + d(y,t) = d(z,t).$$

This last result substituted above gives

$$d(x,z) + d(z,t) = d(x,t),$$

and since each two points are distinct, finally we have xzt and zyt. The converse can be proved in similar manner. $\square$

Remark A.1.1. We note that the notion of betweenness relation of definition A.1.1 has been introduced for general distance spaces involving three pairwise distinct points x, y, z which must verify a type triangular equality, in the sense of the triangular inequality valid in the specific case of metric spaces. The notion of betweenness therefore does not a priori require the existence of some inequality relating to the points of space but can be introduced in a very general context.

A.1.1 Metric segments and convex sets

Another notion of fundamental importance in the study of metric convexity, namely that of "metric segment" is now presented.

Definition A.1.2 (Metric segment). *Let (X, ρ) be a distance space. If $x, y \in X$ the metric segment (or briefly the ρ-segment) connecting x and y is the set*

$$[x, y]_\rho = \{z \in X : \rho(x, z) + \rho(z, y) = \rho(x, y)\}.$$

Remark A.1.2. We note that in the definition A.1.2 of metric segment the notion of betweenness relation of definition A.1.1 can be used, in the sense that if $x, y \in X$ are two distinct points one can define the metric segment connecting x and y as the set $[x, y]_\rho = \{z \in X : xzy\}$, where z will be different from x and y because if xzy holds then x, y, z are three pairwise distinct points. However in this way the metric segment never contains extreme points x and y. Instead by definition A.1.2 always we have $\{x, y\} \subseteq [x, y]_\rho$.

In arbitrary distance space it may happen that a metric segment of fixed extremes is of little significance, in the sense that it can contain a few points, or on the contrary, it can contain many points, this fact depending on the specific space. Let us note that even is the inclusion $\{x, y\} \subseteq [x, y]_\rho$ is always true and there exist distance spaces (X, ρ) in which $[x, y]_\rho \setminus \{x, y\} \neq \emptyset$ holds for all pairs of points x, y, also there exist spaces corresponding even to a more extreme situation: there is no ρ-segment such that $[x, y]_\rho \setminus \{x, y\} \neq \emptyset$.

Example A.1.1. *Consider the distance space $(\mathbb{R}, \rho)$ of Example 1.1.6. Fixed $x, y \in \mathbb{R}$, the condition $\rho(x, z) + \rho(z, y) = \rho(x, y)$ for $z \in \mathbb{R}$ becomes the equation $(x - y)^2 = (x - z)^2 + (z - y)^2$, which by explicit resolution gives only the two points $z = x$, $z = y$. So in this space a ρ-segment $[x, y]_\rho$ only contains its endpoints, i.e. we have $[x, y]_\rho = \{x, y\}$.*

Example A.1.2. *Consider the distance space $(\mathbb{R}, d)$ with d Euclidean metric. Fixed $x, y \in \mathbb{R}$, the condition $\rho(x, z) + \rho(z, y) = \rho(x, y)$ for $z \in \mathbb{R}$ becomes the equation $|x - y|^2 = |x - z|^2 + |z - y|^2$, which by explicit resolution gives $z \in [x, y]$. So, for all x, y in this space, we have that $[x, y]_d = [x, y]$.*

Example A.1.3. *Consider the distance space $(\mathbb{R}^2, \rho)$ of Example 1.1.8 for $n = 2$. We have $\rho(x, y) = (x_1 - y_1)^2 + (x_2 - y_2)^2$ for $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$, and it is easy to check that there exist points $x, y \in \mathbb{R}^2$ such that the condition $\rho(x, z) + \rho(z, y) = \rho(x, y)$ occurs for points $z = (z_1, z_2) \notin \{x, y\}$. In fact, by direct calculation, this condition is equivalent to $(z_1 - \frac{x_1+y_1}{2})^2 + (z_2 - \frac{x_2+y_2}{2})^2 = \frac{(x_1+y_1)^2}{4} + \frac{(x_2+y_2)^2}{4} - (x_1 y_1 + x_2 y_2)$; choosing $x \neq y$ with $y = (0, 0)$, then we* *obtain in the cartesian plane $z_1 - z_2$ a circumference, centered in $(\frac{x_1}{2}, \frac{x_2}{2})$ and with radius $\frac{\sqrt{x_1^2+x_2^2}}{2}$.*

Obviously an important question is to establish how an idea of convexity in the metric sense corresponds to the standard definition of convexity in the usual sense. The most appropriate field for establishing these relationships is therefore certainly that of normed spaces. If $(V, \|\cdot\|)$ is a normed space, the metric on V is defined in the classical way as $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in V$, and two different notions of convexity can be in parallel introduced using two different definitions of segment, i.e. the definition A.1.2 of metric segment and that classic of line segment: if $\mathbf{x}, \mathbf{y} \in V$ then the line segment joining $\mathbf{x}$ and $\mathbf{y}$ is the set

$$[\mathbf{x}, \mathbf{y}] = \{(1 - \lambda)\mathbf{x} + \lambda\mathbf{y} : 0 \leqslant \lambda \leqslant 1\}.$$

Several immediate consequences can be derived from the two definitions of segment given above, first of all that expressed by the following proposition.

Proposition A.1.3. *Let $(V, \|\cdot\|)$ a normed space equipped with the standard metric defined as $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in V$. Then it is $[\mathbf{x}, \mathbf{y}] \subseteq [\mathbf{x}, \mathbf{y}]_d$ for all $\mathbf{x}, \mathbf{y} \in V$.*

Proof. Fixed $\mathbf{x}, \mathbf{y} \in V$, let $z \in [\mathbf{x}, \mathbf{y}]$ be, so that by definition of $[\mathbf{x}, \mathbf{y}]$ we have

$$\mathbf{z} = (1 - \lambda)\mathbf{x} + \lambda\mathbf{y}$$

for some λ with $0 \leqslant \lambda \leqslant 1$. Then we have

$$\begin{aligned}
d(\mathbf{x}, \mathbf{z}) &= \|\mathbf{x} - \mathbf{z}\| \\
&= \|\mathbf{x} - ((1 - \lambda)\mathbf{x} + \lambda\mathbf{y})\| \\
&= \|\lambda(\mathbf{x} - \mathbf{y})\| \\
&= |\lambda| \, \|\mathbf{x} - \mathbf{y}\| \\
&= \lambda d(\mathbf{x}, \mathbf{y}),
\end{aligned}$$

and

$$\begin{aligned} d(\mathbf{z},\mathbf{y}) &= \|\mathbf{z}-\mathbf{y}\| \\ &= \|(1-\lambda)\mathbf{x}+\lambda\mathbf{y}-\mathbf{y})\| \\ &= \|(1-\lambda)(\mathbf{x}-\mathbf{y})\| \\ &= |1-\lambda|\,\|\mathbf{x}-\mathbf{y}\| \\ &= (1-\lambda)d(\mathbf{x},\mathbf{y}), \end{aligned}$$

so that $d(\mathbf{x},\mathbf{z})+d(\mathbf{z},\mathbf{y}) = d(\mathbf{x},\mathbf{y})$ holds. Finally by definition of $[\mathbf{x},\mathbf{y}]_d$ we conclude that $\mathbf{z} \in [\mathbf{x},\mathbf{y}]_d$. $\square$

About the notion of algebraic convexity for the subsets of a normed space $(V, \|\cdot\|)$, in the case of the algebraic convexity, according to the classical definition a subset $A \subseteq V$ is said be convex (or linearly convex) if for all $\mathbf{x},\mathbf{y} \in A$ the segment $[\mathbf{x},\mathbf{y}]$ is contained in A. Generalizing in the case of a distance space it is possible to give in a natural way the following definition of "metrically convex" set.

Definition A.1.3 (Metric convexity)**.** *Let (X,ρ) be a distance space. A subset $A \subseteq X$ is metrically convex (or ρ-convex) if $[x,y]_\rho \subseteq A$ for all $x,y \in A$.*

The empty set is ρ-convex and also V is ρ-convex. There is another description of ρ-convexity, namely: the subset $A \subseteq X$ is ρ-convex if, for any three points $x,y \in A$, $z \in V$ the equality $\rho(x,y) = \rho(x,z)+\rho(z,y)$ implies that $z \in A$. Obviously, this definition of ρ-convexity is equivalent to the preceding one.

In the case of normed space $(V, \|\cdot\|)$ by Proposition A.1.3 it is immediate to verify that if $\mathcal{C}(V)$ and $\mathcal{C}_d(V)$ denote the families of, respectively, convex and d-convex subsets of V then $\mathcal{C}_d(V) \subseteq \mathcal{C}(V)$.

Remark A.1.3. Definitions imply that each d-convex set in a normed spaces is also convex, but the converse is not true in general. Furthermore, a linear segment is a linear convex set, but again the converse can not be true.

Example A.1.4. *Consider the normed space $(\mathbb{R}^2, \|\cdot\|)$ with the norm defined as $\|\mathbf{x}\| = \max\{|x_1|, |x_2|\}$ for all $\mathbf{x} = (x_1,x_2) \in \mathbb{R}^2$. The unit ball B in this space is a square described by the inequalities $|x_1| \leqslant 1$, $|x_2| \leqslant 1$, and it is linearly convex, but not d-convex. Indeed, the points $\mathbf{a} = (1,1)$ and $\mathbf{b} = (-1,1)$ belong to unit ball, but not the point $\mathbf{c} = (0,2)$. At the same time we have $d(\mathbf{a},\mathbf{b}) = 2$ and $d(\mathbf{a},\mathbf{c}) = d(\mathbf{c},\mathbf{b}) = 1$, i.e. $\mathbf{c} \in [\mathbf{a},\mathbf{b}]_d$.*

Example A.1.5. *Let us recall that a convex set $A \subseteq \mathbb{R}^n$ is called a convex body if it is closed and has interior points. Now consider the normed space $(\mathbb{R}^3, \|\cdot\|)$ with the norm defined as $\|\mathbf{x}\| = \max\{|x_1|, |x_2|, |x_3|\}$ for all $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$. The unit ball B in this space is a cube described by the inequalities $|x_1| \leqslant 1$, $|x_2| \leqslant 1$ and $|x_3| \leqslant 1$. We note that the only d-convex body in this space is given by $\mathbb{R}^3$ itself. Since each body contains a subset which is homothetic to the unit ball B, it suffices to show that a d-convex set A containing B is necessarily equal to $\mathbb{R}^3$. Thus assume that a d-convex set A contains B, and set*

$$\mathbf{a}_k = \begin{cases} (k,1,0) & for \quad k \quad even \\ (k,0,1) & for \quad k \quad odd \end{cases}$$

and

$$\mathbf{b}_k = \begin{cases} (k,-1,0) & for \quad k \quad even \\ (k,0,-1) & for \quad k \quad odd. \end{cases}$$

Then it is easy to show that $d(\mathbf{a}_k, \mathbf{b}_k) = 2$ and $d(\mathbf{a}_k, \mathbf{a}_{k+1}) = d(\mathbf{a}_k, \mathbf{b}_{k+1}) = d(\mathbf{b}_k, \mathbf{a}_{k+1}) = d(\mathbf{b}_k, \mathbf{b}_{k+1}) = 1$, and therefore $\mathbf{a}_{k+1}, \mathbf{b}_{k+1} \in [\mathbf{a}_k, \mathbf{b}_k]_d$. Since $\mathbf{a}_1, \mathbf{b}_1 \in B \subseteq A$ and A is d-convex, the points $\mathbf{a}_k, \mathbf{b}_k$ belong to A. Furthermore, since any d-convex set is linearly convex, this implies that the positive x_1-semi axis is contained in A. Reasoning in the same way for other semi axis we have that every coordinate axis is contained in A, thus, by convexity arguments, $A = \mathbb{R}^3$. Furthermore, if for $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ the vector $\mathbf{a} - \mathbf{b}$ is not parallel to a spatial diagonal of the cube B, then the d-segment $[\mathbf{a}, \mathbf{b}]_d$ has interior points, i.e., it is a body. Since such a body is bounded, it cannot coincide with $\mathbb{R}^3$, hence in $\mathbb{R}^3$ there are d-segments which are not d-convex.

A.2 Menger and Takahashi's convexity

In this section we recall another definition of metric convexity, adopted by K. Menger in his main work [16].

Definition A.2.1 (Menger's convexity). *Let (X, d) be a metric space. A subset $A \subseteq X$ is metrically convex in the sense of Menger if and only if it contains for each pairs of its distinct points at least one between-point, i.e. for all $x, y \in A$ with $x \neq y$ then there exists $z \in A$ such that xzy holds.*

It should be noted that this original definition of Menger concerning metric convexity for the sets in general metric spaces when applied to a normed space $(V, \|\cdot\|)$ imply weaker notions than definition A.1.3. For instance one can derives that a set $A \subseteq V$ is metrically convex in the sense of Menger if $([\mathbf{x}, \mathbf{y}]_d \backslash \{\mathbf{x}, \mathbf{y}\}) \cap A \neq \emptyset$ holds for every pair of distinct points $\mathbf{x}, \mathbf{y} \in A$.

Still according to Menger, in a metric space also the definition of segment is given in a more general way using the notion of congruence. For two metric spaces (X_1, d_1) and (X_2, d_2), if $x_1, y_1 \in X_1$ and $x_2, y_2 \in X_2$, then x_1, y_1 are congruent to x_2, y_2 if and only if $d_1(x_1, y_1) = d_2(x_2, y_2)$. For the sets, if A and B are subsets of the same or different metric spaces, then they are congruent if and only if there exists a mapping f of A onto B such that each point-pair of A is mapped onto a congruent point-pair of B, i.e., A and B are congruent if and only if f is an isometric map of A onto B. Then obviously a such f is biuniform, in the sense that for $x, y \in A$ we have $d_1(x, y) = d_2(f(x), f(y))$ and so $f(x) = f(y)$ if and only if $x = y$. Furthermore f is a symmetric, reflexive and transitive relation and it is easily to see that it is a continuous map, hence if two subsets of metric spaces are congruent they are also homeomorphic.

Using this notion of congruence, the next definition therefore can introduce the segments in a metric space according to Menger.

Definition A.2.2 (Menger's segment). *Let (X, d) be a metric space. If $S \subseteq X$ and $x, y \in S$, then S is a segment in the sense of Menger with end-points x, y if and only if it is congruent with a line segment of length $d(x, y)$ of the Euclidean metric space $\mathbb{R}$.*

Remark A.2.1. It should be noted that, more explicitly, the definition A.2.2 requires the existence of an isometry $f : [0, d(x, y)] \to S$ with $f(0) = x$ and $f(d(x, y)) = y$.

Menger, a pioneer in the axiomatic study of distance spaces, starting from above definitions was the first to discover and to prove the following remarkable fact: each two points of a closed and convex subset of a compact metric space are joined by a metric segment belonging to the subset, and using the transfinite induction he extended this same result also for complete metric space. Here we want show an elegant proof based on the classical Caristi's fixed point theorem [66] that we recall:

Theorem A.2.1 (Caristi's theorem). *Let (X, d) be a complete metric space and let $\phi : X \to \mathbb{R}$ be a lower semicontinuous. function which is bounded below. If $T : X \to X$ is an arbitrary mapping which satisfies the condition*

$$d(x, T(x)) \leqslant \varphi(x) - \varphi(T(x))$$

for all $x \in X$, then T has a fixed point.

This proof of Menger's theorem the is basically the same as the original except for the proof of the following lemmas which utilize Caristi's theorem instead of a lengthy transfinite induction.

Lemma A.2.1. *Let (X, d) be a complete metric space, $x, y \in X$ with $x \neq y$, λ such that $0 < \lambda < d(x, y)$ and $S = S(x, y, \lambda) = \{z \in [x, y]_d : d(x, z) \leqslant \lambda\} \cup \{x\}$. Then there exists a point $z_\lambda \in X$ such that*

i) $z_\lambda \in S(x, y, \lambda)$;

ii) $u \in [x, y]_d$ *and* $xz_\lambda y$ *imply* $d(x, u) > \lambda$.

Proof. If there exists $z' \in S$ with $d(x, z') < \lambda$ such that $xz'y$ implies $u \notin S$, we take $z_\lambda = z'$. If for each $z \in S$ with $d(x, z) < \lambda$ there exists y_z such that xzy_z holds, we define the map $G : S \to S$ by taking $G(z) = y_z$ if $d(x, z) < \lambda$ and $G(z) = z$ otherwise. Now define $\phi : S \to \mathbb{R}^+$ by taking $\phi(z) = \lambda - d(x, z)$. Then clearly ϕ is continuous and for $z \in S$ we have

$$\begin{aligned} d(z, G(z)) &= d(x, G(z)) - d(x, z) \\ &= \lambda - d(x, z) - (\lambda - d(x, G(z))) \\ &= \varphi(z) - \varphi(G(z)). \end{aligned}$$

Since S is closed, hence complete, by Caristi's theorem A.2.1 we have $G(z') = z'$ for some point $z' \in S$. This implies $d(x, z') = \lambda$ and so $z' = z_\lambda$ satisfies conditions *i)* and *ii)*. $\square$

Lemma A.2.2. *Let (X, d) be a complete and convex metric space in the sense of Menger, $x, y \in X$ with $x \neq y$ and λ such that $0 < \lambda < d(x, y)$. Then there exists $z' \in X$ such that $xz'y$ and $d(x, z') = \lambda$.*

Proof. By Lemma A.2.1 there exists a point $z_\lambda \in X$ such that $z_\lambda \in S(x, y, \lambda)$ and if $u \in [x, y]_d$, $xz_\lambda y$ then $d(x, u) > \lambda$. Let $\lambda' = d(x, y) - \lambda$ and again apply Lemma A.2.1 to obtain $y_{\lambda'} \in X$ such that $y_{\lambda'} \in S(y, z_\lambda, \lambda')$ and $u \in [y, z_\lambda]_d$, $yy_{\lambda'}u$ imply $d(y, u) > \lambda'$. If $z_\lambda = y_{\lambda'}$, then, since $d(x, y) = d(x, z_\lambda) + d(z_\lambda, y)$, it follows that $d(x, z_\lambda) = \lambda$. If $z_\lambda \neq y_{\lambda'}$, since X is convex in the sense of Menger, there exists $w \in X$ such that $z_\lambda w y_{\lambda'}$. By assumption the relations $xz_\lambda y$, $z_\lambda y_{\lambda'} y$ and $z_\lambda w y_{\lambda'}$ hold. It follows by Proposition A.1.2 that xwy, $xz_\lambda w$, ywz_λ and $yy_{\lambda'}w$ also hold. Now xwy and $xz_\lambda w$ imply $d(x, w) > \lambda$, while ywz_λ and

$yy_{\lambda'}w$ imply $d(y,w) > \lambda$. Therefore we obtain $d(x,y) = d(x,w) + d(w,y) > \lambda + \lambda' = d(x,y)$, a contradiction. $\square$

Now we can give the proof of Menger's theorem in the following form.

Theorem A.2.2 (Menger's theorem). *Let (X,d) be a complete and convex in the sense of Menger metric space. Then any two distinct points $x, y \in X$ are joined by a segment in the sense of Menger, i.e. there exists an isometry $f : [0, d(x,y)] \to X$ with $f(0) = x$ and $f(d(x,y)) = y$.*

Proof. Let $x_0, x_1 \in X$ with $x_0 \neq x_1$. By Lemma A.2.2 there exists the "midpoint" of the pair (x_0, x_1), i.e. the point $x_{1/2} \in X$ such that $d(x_0, x_{1/2}) = d(x_{1/2}, x_1) = \frac{1}{2}d(x_0, x_1)$. Let $l = d(x_0, x_1)$ and define the mapping F by taking

$$F(0) = x_0, \quad F(\frac{l}{2}) = x_{1/2}, \quad F(l) = x_1.$$

Applying the Lemma A.2.2 again there exist points $x_{1/4}, x_{3/4}$ which are respective midpoints of the pairs $(x_0, x_{1/2})$ and $(x_{1/2}, x_1)$. Setting

$$F(\frac{l}{4}) = x_{1/4}, \quad F(\frac{3l}{4}) = x_{3/4},$$

by Proposition A.1.2 we have that F is an isometry on the set $\left\{0, \frac{l}{4}, \frac{l}{2}, \frac{3l}{4}, l\right\}$. By induction we obtain the set of points $\{x_{p/2^n}\} \subseteq X$ with $1 \leq p \leq 2^n - 1$, $n = 1, 2 \ldots$, and taking $F(\frac{pl}{2^n}) = x_{l/2^n}$ the mapping F is an isometry. Since the set $\left\{\frac{pl}{2^n}\right\}$ is a dense subset of $[0, l]$ and since X is complete, it is possible to obtain an isometry $f : [0, d(x_0, x_1)] \to X$ extending F in the obvious way to the entire interval $[0, l]$, so we have a segment in the sense of Menger in X joining the points x_0 and x_1, completing the proof. $\square$

Now we show another idea of convexity in a metric space, introduced by Takahashi [67]. This type of convexity is described in an abstract form and in a sense is a natural generalization of convexity in normed spaces and Euclidean spaces in particular.

Definition A.2.3 (Takahashi's convex structure). *Let (X,d) be a metric space and $I = [0,1]$. A continuous function $W : X \times X \times I \in X$ is said to be a convex structure on X if for each $x, y \in X$ and all $t \in I$, then the condition*

$$d(u, W(x,y,t)) \leqslant (1-t)d(u,x) + td(u,y) \tag{A.1}$$

holds for all $u \in X$.

A metric space (X, d) with a convex structure W is called a convex metric space in the sense of Takahashi, and is denoted by (X, W, d). A subset K of X is called convex if $W(x, y; t) \in K$ whenever $x, y \in C$ and $t \in I$.

What makes Takahashi's notion of convexity solid is the invariance under taking intersections and convexity of closed balls. We have the following easy propositions.

Proposition A.2.1. *Let (X, W, d) be a Takahashi's convex metric space. Then the intersection of an arbitrary family of Takahashi's convex subsets of X is a Takahashi's convex subset of X.*

Proposition A.2.2. *Let (X, W, d) be a Takahashi's convex metric space. Then open balls $B(x, r)$ and the closed balls $\bar{B}(x, r)$ in X are Takahashi's convex subset of X.*

Proof. For $y, z \in B(x, r)$ and $t \in I$, there exists $W(y, z, y) \in X$. Since X is a Takahashi's convex metric space, $d(x, W(y, z, t) \leq td(x, y) + (1 - t)d(x, z) < tr + (1 - t)r = r$. Therefore $W(y, z, t) \in B(x, r)$. Similarly, $\bar{B}(x, r)$ is a Takahashi's convex subset of X.

Proposition A.2.3. *Let (X, W, d) be a Takahashi's convex metric space. Then for $x, y \in X$ and $t \in I$, it is*

$$d(x, y) = d(x, W(x, y, t)) + d(W(x, y, t), y).$$

Proof. Since X is a Takahashi's convex metric space, we obtain

$$\begin{aligned} d(x, y) &\leq d(x, W(x, y, t)) + d(W(x, y, t), y) \\ &\leq td(x, x) + (1 - t)d(x, y) + td(x, y) + (1 - t)d(y, y) \\ &= td(x, y) + (1 - t)d(x, y) = d(x, y) \end{aligned}$$

for $x, y \in X$ and $t \in I$. Therefore, $d(x, y) = d(x, W(x, y, t)) + d(W(x, y, t), y)$ for $x, y \in X$ and $t \in I$.

Proposition A.2.4. *Let (X, W, d) be a Takahashi's convex metric space. Then for any $x, y \in X$ and any $t \in I$ we have $d(x, W(x, y, t)) = td(x, y)$ and $d(y, W(x, y, t)) = (1 - t)d(x, y)$.*

Proof. For simplicity, let a, b and c stand for $d(x, W(x, y, t))$, $d(y, W(x, y, t))$ and $d(x, y)$ respectively. By (A.1) we get $a \leq tc$ and $b \leq (1-t)c$. But $c \leq a + b$ by the triangle inequality, so $c \leq a + b \leq (1-t)c + tc = c$. This means $a + b = c$. If $a < tc$ then we would have $a + b < c$ which is a contradiction, therefore, we must have $a = tc$ and consequently $b = (1-t)c$.

The necessity for the condition (A.1) on W to be a convex structure on a metric space (X, d) is natural. To see this, assume that $(X, \|\cdot\|)$ is a normed space. Then the mapping $W : X \times X \times I \to X$ given by

$$W(\mathbf{x}, \mathbf{y}, t) = (1-t)\mathbf{x} + t\mathbf{y}, \quad \mathbf{x}, \mathbf{y} \in X, \quad t \in I, \tag{A.2}$$

defines a convex structure on X. Indeed, if d is the metric induced by the norm, then we have

$$\begin{aligned} d(\mathbf{u}, W(\mathbf{x}, \mathbf{y}, t)) &= \|\mathbf{u} - ((1-t)\mathbf{x} + t\mathbf{y}\| \\ &\leq (1-t)\|\mathbf{u} - \mathbf{x}\| + t\|\mathbf{u} - \mathbf{y}\| \\ &= (1-t)d(\mathbf{u}, \mathbf{x}) + td(\mathbf{u}, \mathbf{y}), \end{aligned}$$

for all $\mathbf{u} \in X$, $t \in I$. The picture gets clearer in the linear space $\mathbb{R}^2$ with the Euclidean metric and the convex structure given by (A.2). In this case, given two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ and a $t \in I$, $\mathbf{z} = W(\mathbf{x}, \mathbf{y}, t)$ is a point that lies on the line segment joining $\mathbf{x}$ and $\mathbf{y}$. Moreover, Proposition A.2.4 implies that if $d(\mathbf{x}, \mathbf{y}) = L$ then $d(\mathbf{x}, \mathbf{z}) = tL$ and $d(\mathbf{z}, \mathbf{y}) = (1-t)L$ and we arrive at an interesting exercise of elementary trigonometry to show that $d(\mathbf{u}, \mathbf{z}) \leq (1-t)d(\mathbf{u}, \mathbf{x}) + td(\mathbf{u}, \mathbf{y})$ for any point $\mathbf{u}$ in the plane, that which can be solved applying the Pythagorean theorem to suitable right triangles and using the triangular inequality $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{u}) + d(\mathbf{u}, \mathbf{y})$.

We conclude this section by noting that also another definition of convex metric space can be given, according to R. R. Khalil, which is not abstract like that of Takahashi and that does not involve metric segments like that of Menger, but the intersection of closed balls [68].

Definition A.2.4 (Khalil's convexity)**.** *Let (X, d) be a metric space. A subset $A \subseteq X$ is metrically convex in the sense of Khalil if and only if for each pairs of its distinct points $x, y \in X$, the condition*

$$\bar{B}(x, \lambda) \cap \bar{B}(y, r - \lambda) \neq \emptyset, \quad r = d(x, y), \quad \lambda \in [0, r]$$

holds.

It is important to underline that there are some equivalences between the three different ideas of convexity that have been introduced, as the following propositions show. Indeed, the convexity of Menger and that of Khalil are equivalent in complete metric spaces, but in general only Khalil'convexity implies Menger'convexity [69]; on the other hand, Takahashi's convexity always implies Khalil's convexity.

Proposition A.2.5. *Let (X, d) be a metric space. If X is convex in the sense of Khalil, then X is convex in the sense of Menger.*

Proof. Let $x, y \in X$, $x \neq y$ and $r = d(x, y)$. By Definition A.2.4 there exists $z \in X$ such that $z \in \bar{B}(x, \lambda) \cap \bar{B}(y, r - \lambda)$ for any $\lambda \in [0, r]$. Then we have $d(x, z) \leq \lambda$, $d(z, y) \leq r - \lambda$ and, by triangular inequality,

$$\begin{aligned} d(x, y) &\leq d(x, z) + d(z, y) \\ &\leq \lambda + r - r \\ &\leq \lambda \\ &\leq d(x, y); \end{aligned}$$

we obtain $d(x, z) + d(z, y) = d(x, y)$, i.e. $z \in [x, y]_d$. Assuming $\lambda > 0$, if $z = x$, we have $d(x, y) \leq r - \lambda \leq d(x, y) - \lambda$, that is a contradiction; assuming $\lambda < d(x, y)$, if $z = y$, we have $d(x, y) \leq \lambda$, a contradiction again. Finally, by Definition A.2.1, X is convex in the sense of Menger, indeed we have $z \neq x, y$ such that $z \in [x, y]_d$. $\square$

Proposition A.2.6. *Let (X, d) be a complete metric space. If X is convex in the sense of Menger, then X is convex in the sense of Khalil.*

Proof. Let $x, y \in X$, $x \neq y$, $r = d(x, y)$ and $\lambda \in [0, r]$. By Theorem A.2.2, x and y are joined by a segment in the sense of Menger, i.e. there exists an isometry $f : [0, d(x, y)] \to X$ with $f(0) = x$ and $f(d(x, y)) = y$. This means that for any $t \in [0, 1]$, there exists a point $z \in X$ such that $d(x, z) = td(x, y)$ and $d(z, y) = (1 - t)d(x, y)$, so assuming $t = \frac{\lambda}{r}$, we have $d(x, z) = td(x, y) = tr = \lambda$ and $d(z, y) = (1 - t)d(x, y) = (1 - t)r = r - \lambda$, i.e. $z \in \bar{B}(x, r) \cap \bar{B}(y, \lambda - r)$, concluding that X is convex in the sense of Khalil. $\square$

Proposition A.2.7. *Let (X, d) be a metric space. If X is convex in the sense of Takahashi, then X is convex in the sense of Menger.*

Proof. Let $x, y \in X$ and $t \in I = [0,1]$. Considering $z = W(x,y,t)$, by Proposition A.2.4 we have $d(x,z) = d(x, W(x,y,t)) = td(x,y)$ and $d(z,y) = d(y, W(x,y,t)) = (1-t)d(x,y)$, so that $d(x,z) + d(z,y) = d(x,y)$, i.e. $z \in [x,y]_d$; assuming $t \neq 0, 1$ then we have $z \neq x, y$, so X is convex in the sense of Menger. $\square$

Finally we observe that it is not known whether convexity in the sense of Menger or Khalil implies convexity in the sense of Takahashi, under suitable hypothesis.

A.3 Convex function

A very interesting possible definition of convex function in a metric sense can be given in relation to the notion of d-convex set [70]. For these convex functions, defined in a general metric space, it is then immediate to develop some of the typical properties of convex functions in the ordinary sense, i.e. in the context of normed spaces [71].

Definition A.3.1 (d-convex function)**.** *Let (X, d) be a metric space. A function $f : X \to \mathbb{R}$ is said to be d-convex if for any points $x, y, z \in X$, with $x \neq y$ and $z \in [x,y]_d$, the inequality*

$$f(z) \leqslant \frac{d(z,y)}{d(x,y)} f(x) + \frac{d(x,z)}{d(x,y)} f(y) \tag{A.3}$$

holds.

The naturalness of definition (A.3.1) is justified by the next results. Let us first recall that a normed space $(V, \|\cdot\|)$ is said to be strictly convex is said to be strictly convex (or rotund) if the mid-point of any line segment joining two different points on the unit sphere S of V does not lie on it, i.e., the implication holds $\|\mathbf{x}\| = 1, \|\mathbf{y}\| = 1, \mathbf{x} \neq \mathbf{y} \implies \left\|\frac{\mathbf{x}+\mathbf{y}}{2}\right\| < 1$, holds. The notation of strict convexity can be described in various equivalent ways according to the following result, .

Theorem A.3.1. *Let $(V, \|\cdot\|)$ be a normed space. Then the following properties are equivalent:*

a) V is strictly convex;

b) *If* $\mathbf{x}, \mathbf{y} \in S$ *and* $\mathbf{x} \neq \mathbf{y}$*, then* $\|\lambda \mathbf{x} + (1-\lambda)\mathbf{y}\| < 1$ *for every* $\lambda \in (0,1)$*, i.e. the unit sphere* S *does not contain any segment;*

c) *If* $\mathbf{x}, \mathbf{y} \in S$ *and* $\mathbf{x} \neq \mathbf{y}$*, then* $\|\lambda \mathbf{x} + (1-\lambda)\mathbf{y}\| < 1$ *for some* $\lambda \in (0,1)$*;*

d) *If for any* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$*, all different,* $\|\mathbf{x}-\mathbf{y}\| = \|\mathbf{x}-\mathbf{z}\| + \|\mathbf{z}-\mathbf{y}\|$ *then there exists* $\lambda \in (0,1)$ *such that* $\mathbf{z} = \lambda \mathbf{x} + (1-\lambda)\mathbf{y}$*;*

e) *If* $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$ *for all* $x \neq \mathbf{0}$*,* $\mathbf{y} \neq \mathbf{0}$ *then* $\mathbf{x} = c\mathbf{y}$*, for some* $c > 0$*;*

f) *If for* $\mathbf{x}, \mathbf{y} \in S$*,* $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$ *then* $\mathbf{x} = \mathbf{y}$*;*

g) *The function* $h : V \to]0, +\infty[$ *, defined by* $h(\mathbf{x}) = \|\mathbf{x}\|^2$ *is strictly convex.*

Proof. *a)* $\Longrightarrow$ *b)* Take $\mathbf{x}, \mathbf{y} \in S$ and $\mathbf{x} \neq \mathbf{y}$. For $0 < \lambda < \frac{1}{2}$ we have

$$\begin{aligned} \mathbf{z} &= \lambda \mathbf{x} + (1-\lambda)\mathbf{y} \\ &= 2\lambda \frac{\mathbf{x}+\mathbf{y}}{2} + (1-2\lambda)\mathbf{y} \end{aligned}$$

and then

$$\begin{aligned} \|\mathbf{z}\| &\leqslant 2\lambda \left\| \frac{\mathbf{x}+\mathbf{y}}{2} \right\| + (1-2\lambda)\|\mathbf{y}\| \\ &< 2\lambda + (1-2\lambda) \\ &= 1. \end{aligned}$$

When $0 < \lambda < \frac{1}{2}$ we have

$$\begin{aligned} \|\mathbf{z}\| &\leqslant (2\lambda - 1)\|\mathbf{x}\| + (2-2\lambda)\left\| \frac{\mathbf{x}+\mathbf{y}}{2} \right\| \\ &< 2\lambda - 1 + 2 - 2\lambda \\ &= 1. \end{aligned}$$

Hence for all $\lambda \in (0,1)$ we obtain $\|\lambda \mathbf{x} + (1-\lambda)\mathbf{y}\| < 1$;

b) $\Longrightarrow$ *a)* Just take $\lambda = \frac{1}{2}$;

a) $\Longrightarrow$ *c)* Obvious;

c) $\Longrightarrow$ *a)* Take $\mathbf{x}, \mathbf{y} \in S$ and $\mathbf{x} \neq \mathbf{y}$. If $0 < \lambda_0 < \frac{1}{2}$, setting $\alpha = \frac{\frac{1}{2}-\lambda_0}{1-\lambda_0}$ then

$$\frac{\mathbf{x}+\mathbf{y}}{2} = \alpha \mathbf{x} + (1-\alpha)(\lambda_0 \mathbf{x} + (1-\lambda_0)\mathbf{y})$$

and so

$$\begin{aligned}\left\|\frac{\mathbf{x}+\mathbf{y}}{2}\right\| &\leqslant \alpha\,\|\mathbf{x}\| + (1-\alpha)\,\|\lambda_0\mathbf{x}+(1-\lambda_0)\mathbf{y}\| \\ &< 1.\end{aligned}$$

If $\frac{1}{2} < \lambda_0 < 1$, setting $\beta = \frac{1}{2\lambda_0}$, then

$$\frac{\mathbf{x}+\mathbf{y}}{2} = \beta(\lambda_0\mathbf{x}+(1-\lambda_0)\mathbf{y}) + (1-\beta)\mathbf{y}$$

an so

$$\begin{aligned}\left\|\frac{\mathbf{x}+\mathbf{y}}{2}\right\| &= \beta\,\|\lambda_0\mathbf{x}+(1-\lambda_0)\mathbf{y}\| + (1-\beta)\,\|\mathbf{y}\| \\ &< 1.\end{aligned}$$

a) $\Longrightarrow$ *d)* Take $\mathbf{x},\mathbf{y},\mathbf{z} \in V$, all different and such that $\|\mathbf{x}-\mathbf{y}\| = \|\mathbf{x}-\mathbf{z}\| + \|\mathbf{z}-\mathbf{y}\|$. Then $\|\mathbf{x}-\mathbf{z}\| \neq \mathbf{0}$, $\|\mathbf{z}-\mathbf{y}\| \neq \mathbf{0}$ ans suppose that $\|\mathbf{x}-\mathbf{z}\| \leq \|\mathbf{z}-\mathbf{y}\|$. We have $\left\|\frac{1}{2}\frac{\mathbf{x}-\mathbf{z}}{\|\mathbf{x}-\mathbf{z}\|} + \frac{1}{2}\frac{\mathbf{z}-\mathbf{y}}{\|\mathbf{z}-\mathbf{y}\|}\right\| =$

$$\begin{aligned}
&= \left\|\frac{1}{2}\frac{\mathbf{x}-\mathbf{z}}{\|\mathbf{x}-\mathbf{z}\|} + \frac{1}{2}\frac{\mathbf{z}-\mathbf{y}}{\|\mathbf{z}-\mathbf{y}\|} - \frac{1}{2}\frac{\mathbf{z}-\mathbf{y}}{\|\mathbf{x}-\mathbf{z}\|} + \frac{1}{2}\frac{\mathbf{z}-\mathbf{y}}{\|\mathbf{x}-\mathbf{z}\|}\right\| \\
&\geqslant \left\|\frac{1}{2}\frac{\mathbf{x}-\mathbf{z}}{\|\mathbf{x}-\mathbf{z}\|} + \frac{1}{2}\frac{\mathbf{z}-\mathbf{y}}{\|\mathbf{x}-\mathbf{z}\|}\right\| - \left\|\frac{1}{2}\frac{\mathbf{z}-\mathbf{y}}{\|\mathbf{x}-\mathbf{z}\|} + \frac{1}{2}\frac{\mathbf{z}-\mathbf{y}}{\|\mathbf{z}-\mathbf{y}\|}\right\| \\
&\geqslant \left\|\frac{1}{2}\frac{\mathbf{x}-\mathbf{z}}{\|\mathbf{x}-\mathbf{z}\|} + \frac{1}{2}\frac{\mathbf{z}-\mathbf{y}}{\|\mathbf{x}-\mathbf{z}\|}\right\| - \left\|\frac{1}{2}\frac{\mathbf{z}-\mathbf{y}}{\|\mathbf{x}-\mathbf{z}\|} + \frac{1}{2}\frac{\mathbf{z}-\mathbf{y}}{\|\mathbf{z}-\mathbf{y}\|}\right\| \\
&= \left\|\frac{1}{2}\frac{\mathbf{x}-\mathbf{z}+\mathbf{z}-\mathbf{y}}{\|\mathbf{x}-\mathbf{z}\|}\right\| - \left\|\frac{1}{2}\frac{\|\mathbf{z}-\mathbf{y}\|-\|\mathbf{x}-\mathbf{z}\|}{\|\mathbf{x}-\mathbf{z}\|\,\|\mathbf{z}-\mathbf{y}\|}(\mathbf{z}-\mathbf{y})\right\| \\
&= \frac{1}{2}\frac{\|\mathbf{x}-\mathbf{y}\|-\|\mathbf{z}-\mathbf{y}\|+\|\mathbf{x}-\mathbf{z}\|}{\|\mathbf{x}-\mathbf{z}\|} \\
&= \frac{1}{2}\frac{\|\mathbf{x}-\mathbf{z}\|+\|\mathbf{z}-\mathbf{y}\|-\|\mathbf{z}-\mathbf{y}\|+\|\mathbf{x}-\mathbf{z}\|}{\|\mathbf{x}-\mathbf{z}\|} \\
&= 1.
\end{aligned}$$

Thus, in view of the assumption *b)*, we have

$$\frac{\mathbf{x}-\mathbf{z}}{\|\mathbf{x}-\mathbf{z}\|} = \frac{\mathbf{z}-\mathbf{y}}{\|\mathbf{z}-\mathbf{y}\|}$$

or

$$\left(\frac{1}{\|\mathbf{x}-\mathbf{z}\|} + \frac{1}{\|\mathbf{z}-\mathbf{y}\|}\right)\mathbf{z} = \frac{1}{\|\mathbf{x}-\mathbf{z}\|}\mathbf{x} + \frac{1}{\|\mathbf{z}-\mathbf{y}\|}\mathbf{y}.$$

If we take $\lambda = \frac{1}{\|\mathbf{x}-\mathbf{z}\|} / \left(\frac{1}{\|\mathbf{x}-\mathbf{z}\|} + \frac{1}{\|\mathbf{z}-\mathbf{y}\|} \right)$ we get $\mathbf{z} = \lambda\mathbf{x} + (1-\lambda)\mathbf{y})$ and *d)* holds;

d) $\Longrightarrow$ *a)* Take $\mathbf{x}, \mathbf{y} \in S$, $\mathbf{x} \neq \mathbf{y}$ and assume $\left\| \frac{\mathbf{x}+\mathbf{y}}{2} \right\| = 1$. Thus $\|\mathbf{x} - (-\mathbf{y})\| = \|\mathbf{x}\| + \|-\mathbf{y}\|$ and then by *d)* applied for $\mathbf{z} = \mathbf{0}$, there exists $\lambda \in (0,1)$ such that $\lambda\mathbf{x} + (1-\lambda)(-\mathbf{y}) = \mathbf{0}$. Hence

$$\begin{aligned} 0 &= \|\lambda\mathbf{x} + (1-\lambda)(-\mathbf{y})\| \\ &\geqslant |\lambda| \, \|\mathbf{x}\| - (1-\lambda) \, \|-\mathbf{y}\| \\ &= |2\lambda - 1| \end{aligned}$$

gives $\lambda = \frac{1}{2}$ and so $\frac{1}{2}\mathbf{x} + \frac{1}{2}(-\mathbf{y}) = 0$ or $\mathbf{x} = \mathbf{y}$, which is a contradiction with $\mathbf{x} \neq \mathbf{y}$;

e) $\Longrightarrow$ *b)* Take $\mathbf{x}, \mathbf{y} \in S$ and $\mathbf{x} \neq \mathbf{y}$. The equality $\lambda\mathbf{x} = c(1-\lambda)\mathbf{y}$, $c > 0$, $0 < \lambda < 1$ can not take place because if so then $\lambda \, \|\mathbf{x}\| = c(1-\lambda) \, \|\mathbf{y}\|$ and $\lambda = c(1-\lambda)$which means that $\mathbf{x} = \mathbf{y}$ and we get a contradiction with assumption. If $\lambda\mathbf{x} \neq c(1-\lambda)\mathbf{y}$ and V satisfies *e)*, then

$$\begin{aligned} \|\lambda\mathbf{x} + (1-\lambda)\mathbf{y}\| &< \|\lambda\mathbf{x}\| + \|(1-\lambda)\mathbf{y}\| \\ &= \lambda \, \|\mathbf{x}\| + (1-\lambda) \, \|\mathbf{y}\| \\ &= 1, \end{aligned}$$

i.e. *b)* holds;

b) $\Longrightarrow$ *e)* Let V satisfy *b)*, $\mathbf{x} \neq \mathbf{0}$, $\mathbf{y} \neq \mathbf{0}$, $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$, but $\mathbf{x} \neq c\mathbf{y}$ for any $c > 0$. Then $\mathbf{x}' = \frac{\mathbf{x}}{\|\mathbf{x}\|}$, $\mathbf{y}' = \frac{\mathbf{y}}{\|\mathbf{y}\|}$, satisfy $\mathbf{x}' \neq \mathbf{y}'$ and $\|\mathbf{x}'\| = \|\mathbf{y}'\| = 1$. So we have $\|\lambda\mathbf{x}' + (1-\lambda)\mathbf{y}'\| < 1$ for any $\lambda \in (0,1)$ and for $\lambda = \frac{\|\mathbf{x}\|}{\|\mathbf{x}\|+\|\mathbf{y}\|}$ we get

$$\begin{aligned} \left\| \frac{\|\mathbf{x}\|}{\|\mathbf{x}\| + \|\mathbf{y}\|}\mathbf{x}' + \frac{\|\mathbf{y}\|}{\|\mathbf{x}\| + \|\mathbf{y}\|}\mathbf{y}' \right\| &= \left\| \frac{\|\mathbf{x}\|}{\|\mathbf{x}\| + \|\mathbf{y}\|} \frac{\mathbf{x}}{\|\mathbf{x}\|} + \frac{\|\mathbf{y}\|}{\|\mathbf{x}\| + \|\mathbf{y}\|} \frac{\mathbf{y}}{\|\mathbf{y}\|} \right\| \\ &= \frac{\|\mathbf{x} + \mathbf{y}\|}{\|\mathbf{x}\| + \|\mathbf{y}\|} \\ &< 1, \end{aligned}$$

so that $\|\mathbf{x} + \mathbf{y}\| < \|\mathbf{x}\| + \|\mathbf{y}\|$, i.e. a contradiction;

a) $\Longrightarrow$ *f)* Assume that $\mathbf{x}, \mathbf{y} \in S$, $\mathbf{x} \neq \mathbf{y}$, $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$ and strict convexity of V. Then $\left\| \frac{\mathbf{x}+\mathbf{y}}{2} \right\| < 1$ therefore $\|\mathbf{x} + \mathbf{y}\| < 2 = \|\mathbf{x}\| + \|\mathbf{y}\|$, i.e. a contradiction;

f) $\Longrightarrow$ *a)* Take $\mathbf{x}, \mathbf{y} \in S$ and $\mathbf{x} \neq \mathbf{y}$. By *f)* we have $\|\mathbf{x}+\mathbf{y}\| < \|\mathbf{x}\| + \|\mathbf{y}\|$ and thus $\left\|\frac{\mathbf{x}+\mathbf{y}}{2}\right\| < 1$, i.e. V is strictly convex;

b) $\Longrightarrow$ *g)* Observe that h is a convex function. In fact

$$\begin{aligned}
h(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) &= \|\lambda\mathbf{x} + (1-\lambda)\mathbf{y}\|^2 \\
&\leqslant (\lambda\|\mathbf{x}\| + (1-\lambda)\|\mathbf{y}\|)^2 \\
&= \lambda^2\|\mathbf{x}\|^2 + 2\lambda(1-\lambda)\|\mathbf{x}\|\|\mathbf{y}\| + (1-\lambda)^2\|\mathbf{y}\|^2 \\
&\leqslant \lambda^2\|\mathbf{x}\|^2 + \lambda(1-\lambda)\left(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\right) + (1-\lambda)^2\|\mathbf{y}\|^2 \\
&= \lambda\|\mathbf{x}\|^2 + (1-\lambda)\|\mathbf{y}\|^2 \\
&= \lambda h(\mathbf{x}) + (1-\lambda)h(\mathbf{y}).
\end{aligned}$$

Suppose now that h is not strictly convex, which means there exists $\lambda_0 \in (0,1)$ such that we have equality

$$h(\lambda_0\mathbf{x} + (1-\lambda_0)\mathbf{y}) = \lambda_0 h(\mathbf{x}) + (1-\lambda_0)h(\mathbf{y}),$$

i.e.

$$\|\lambda_0\mathbf{x} + (1-\lambda_0)\mathbf{y}\|^2 = \lambda_0\|\mathbf{x}\|^2 + (1-\lambda_0)\|\mathbf{y}\|^2.$$

Then, if we take $\mathbf{x}, \mathbf{y} \in S$ we obtain $\|\lambda_0\mathbf{x} + (1-\lambda_0)\mathbf{y}\| = 1$ and we have a contradiction with the strict convexity of V;

g) $\Longrightarrow$ *b)* Let h be a strictly convex function and $\mathbf{x}, \mathbf{y} \in V$, $\mathbf{x} \neq \mathbf{y}$, $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$. Suppose that there exists $\lambda(0,1)$ such that $\|\lambda\mathbf{x} + (1-\lambda)\mathbf{y}\| = 1$. Then

$$\begin{aligned}
h(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) &= \|\lambda\mathbf{x} + (1-\lambda)\mathbf{y}\|^2 \\
&= 1 \\
&= \lambda\|\mathbf{x}\|^2 + (1-\lambda)\|\mathbf{y}\|^2 \\
&= \lambda h(\mathbf{x}) + (1-\lambda)h(\mathbf{y}),
\end{aligned}$$

i.e. a contradiction. $\square$

Now we can show how in the case of normed spaces a metrically convex function is convex in the usual sense, with an peculiarity if the space is strictly convex.

Theorem A.3.2. *Let $(V, \|\cdot\|)$ a normed space equipped with the standard metric d. Then every d-convex function $f : V \to \mathbb{R}$ is convex in the usual sense, and these concepts coincide if and only if the space V is strictly convex.*

Proof. Let $f : V \to \mathbb{R}$ be a d-convex function. Take $\mathbf{x}, \mathbf{y} \in V$, $\mathbf{x} \neq \mathbf{y}$ and $\mathbf{z} = \lambda\mathbf{x} + (1-\lambda)\mathbf{y}$, $\lambda \in [0,1]$. We have

$$\begin{aligned} d(\mathbf{x},\mathbf{z}) &= \|\mathbf{x}-\mathbf{z}\| \\ &= \|\mathbf{x}-(\lambda\mathbf{x}+(1-\lambda)\mathbf{y})\| \\ &= (1-\lambda)\,\|\mathbf{x}-\mathbf{y}\| \\ &= d(\mathbf{x},\mathbf{y}), \end{aligned}$$

and

$$\begin{aligned} d(\mathbf{z},\mathbf{y}) &= \|\mathbf{z}-\mathbf{y}\| \\ &= \|\lambda\mathbf{x}+(1-\lambda)\mathbf{y}-\mathbf{y}\| \\ &= \lambda\,\|\mathbf{x}-\mathbf{y}\| \\ &= d(\mathbf{x},\mathbf{y}), \end{aligned}$$

so $d(\mathbf{x},\mathbf{z}) + d(\mathbf{z},\mathbf{y}) = d(\mathbf{x},\mathbf{y})$, i.e. $\mathbf{z} \in [\mathbf{x},\mathbf{y}]_d$, and also $\lambda = \frac{d(\mathbf{z},\mathbf{y})}{d(\mathbf{x},\mathbf{y})}$, $1-\lambda = \frac{d(\mathbf{x},\mathbf{z})}{d(\mathbf{x},\mathbf{y})}$. By d-convexity of f then we obtain $f(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) \leqslant \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y})$, i.e. f is convex in the usual sense. Now let V be strictly convex and $f : V \to \mathbb{R} \cup \{-\infty\}$ a convex function in the usual sense. Take $\mathbf{x}, \mathbf{y} \in V$, $\mathbf{x} \neq \mathbf{y}$ and $\mathbf{z} \in [\mathbf{x},\mathbf{y}]_d$, i.e. $\mathbf{z} \in V$ such that $\|\mathbf{x}-\mathbf{z}\| + \|\mathbf{z}-\mathbf{y}\| = \|\mathbf{x}-\mathbf{y}\|$. If $\mathbf{z} = \mathbf{x}$, by $\mathbf{x} \neq \mathbf{y}$, we have $\mathbf{z} \neq \mathbf{y}$, and $\frac{d(\mathbf{x},\mathbf{z})}{d(\mathbf{x},\mathbf{y})} = 0$, $\frac{d(\mathbf{z},\mathbf{y})}{d(\mathbf{x},\mathbf{y})} = 1$, so (A.3) trivially occurs as $f(\mathbf{z}) \leq f(\mathbf{z})$; similarly if $\mathbf{z} = \mathbf{y}$. Assuming $\mathbf{z} \neq \mathbf{x}$, $\mathbf{z} \neq \mathbf{y}$, by *d)* of Theorem A.3.1, there exists $\lambda \in (0,1)$ such that $\mathbf{z} = \lambda\mathbf{x} + (1-\lambda)\mathbf{y}$; then we have $\lambda = \frac{d(\mathbf{z},\mathbf{y})}{d(\mathbf{x},\mathbf{y})}$, $1-\lambda = \frac{d(\mathbf{x},\mathbf{z})}{d(\mathbf{x},\mathbf{y})}$, and, by usual convexity of f,

$$f(\mathbf{z}) \leqslant \frac{d(\mathbf{z},\mathbf{y})}{d(\mathbf{x},\mathbf{y})} f(\mathbf{x}) + \frac{d(\mathbf{x},\mathbf{z})}{d(\mathbf{x},\mathbf{y})} f(\mathbf{y}),$$

holds, i.e. f is a d-convex function. □

Note that if (X,d) is a convex metric space in the sense of Menger, and $(\mathbb{R}, d_e)$ is the usual Euclidean metric space, defining on the product $X \times \mathbb{R}$ the metric $\tilde{d}((x_1,r_1),(x_2,r_2)) = d((x_1,x_2)) + d_e((r_1,r_2))$, for all $(x_1,r_1),(x_2,r_2) \in X \times \mathbb{R}$, then the metric space $(X \times \mathbb{R}, \tilde{d})$ is convex in the sense of Menger, and the following proposition about the epigraph of a d-convex function can be easily checked:

Proposition A.3.1. *Let (X,d) be a convex metric space in the sense of Menger and $f : X \to \mathbb{R}$ a function. Then f is a d-convex function if and only if the the epigraph* *of f, i.e. the set epi$(f) = \{(x,r) \in X \times \mathbb{R} : f(x) \leqslant \alpha\}$, $\alpha \in \mathbb{R}$, is a d-convex set.*

Also the following theorems can be easily proved, using essentially the definition A.3.1, showing how this notion of metric convexity has the good properties of usual convex functions defined in normed spaces.

Theorem A.3.3. *Let (X,d) be a metric space. Then for any d-convex function $f : X \to \mathbb{R}$ and any $\alpha \in \mathbb{R}$ the level sets $\{z \in X : f(z) \leq \alpha\}$ and $\{z \in X : f(z) < \alpha\}$ are d-convex sets.*

Theorem A.3.4. *Let (X,d) be a metric space. Then for any d-convex function $f : X \to \mathbb{R}$ and any non-decreasing function $g : \mathbb{R} \to \mathbb{R}$, the function $h(x) = g(f(x))$ is a d-convex function.*

Theorem A.3.5. *Let (X,d) be a metric space and let $\{f_i\}_{i\in\mathbb{N}}$ be a sequence of d-convex function $f_i : X \to \mathbb{R}$. If the numerical sequence $f_i(z) : X \to \mathbb{R} \cup \{-\infty\}$ converges for any point $z \in X$, then the function $f(x) = \lim\limits_{i\to+\infty} f_i(x)$ is a d-convex function.*

Theorem A.3.6. *The sum of d-convex functions is a d-convex function, and the pointwise supremum of an arbitrary family of d-convex functions is a d-convex function.*

A.3.1 Metric convex hull

It is simple to show how someone of properties of sets which are convex in the algebraic sense also remain verified by sets which are convex according to this metric sense. Next proposition concerns the intersection of ρ-convex sets, useful for introducing the notion of ρ-convex hull which generalizes that in the case of the usual convexity.

Proposition A.3.2. *Let (X,ρ) be a distance space. Then the intersection of an arbitrary family of ρ-convex sets of X is a ρ-convex set of X.*

Proof. Let $\{A_\alpha\}_{\alpha\in I}$ denote a family of ρ-convex sets of X, where α runs over an arbitrary index set I. We set $A = \bigcap\limits_{\alpha\in I} A_\alpha$. If $x,y \in A$, then for all $\alpha \in I$ the relation $x,y \in A_\alpha$ holds, and by the ρ-convexity of A_α we have $[x,y]_\rho \subseteq A_\alpha$. Since this inclusion holds for each $\alpha \in I$, we have $[x,y]_\rho \subseteq A$, thus the set A is ρ-convex. $\square$

The ρ-convex hull of a subsets F of a distance space (X, ρ) is naturally the smallest ρ-convex set containing F. The Proposition A.3.2 shows that the ρ-convex hull exists for an arbitrary F, and then it is the intersection of the all ρ-convex sets containing F. In the sequel we indicate the ρ-convex hull of a set F with $\mathrm{conv}_\rho(F)$.

If we now consider the special case of $\mathbb{R}^n$, it is well know that the classical Carathéodory's theorem implies the following. If F is an arbitrary set of $\mathbb{R}^n$, denoting by $I(F)$ the union of the segments $[\mathbf{a}, \mathbf{b}]$ with endpoints $\mathbf{a}, \mathbf{b} \in F$, we call "process of segment joining" the transition from set F to the set $I(F)$. Iterating, we can consider the set $I(I(F))$ obtained by the two-fold application of this process, and so on. We remark that if $\mathrm{conv}(F)$ is the usual convex hull of F, the sets $I(F)$, $I(I(F))$,..., constructed by successive application of segment joining, are contained in $\mathrm{conv}(F)$, and Carathéodory's theorem says that, starting with F, a finite number of iterations of this process yields $\mathrm{conv}(F)$. The process of segment joining can be introduced for d-segments in $\mathbb{R}^n$ too, and analogously, we write $I_d(F)$ for the set obtained from $F \subseteq \mathbb{R}^n$ by this process. Hence one is motivated to ask whether from an arbitrary set $F \subseteq \mathbb{R}^n$ the d-convex hull $\mathrm{conv}_d(F)$ is obtainable by a finite number of corresponding iterations. It is easy to see that the question has a negative answer. Indeed if $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]_d$, where $\mathbf{a}, \mathbf{b} \in F$, then $d(\mathbf{a}, \mathbf{x}) + d(\mathbf{x}, \mathbf{b}) = d(\mathbf{a}, \mathbf{b})$ holds and hence at least one of the distances $d(\mathbf{a}, \mathbf{x})$, $d(\mathbf{x}, \mathbf{b})$ is not larger than $\frac{1}{2}d(\mathbf{a}, \mathbf{b})$. For instance, let $d(\mathbf{a}, \mathbf{x}) \leq d(\mathbf{a}, \mathbf{b}) \leq \mathrm{diam}(F)$, where $\mathrm{diam}(F)$ denotes the diameter of the set F, i.e., the upper bound of the distances $d(\mathbf{x}, \mathbf{y})$ with $\mathbf{x}, \mathbf{y} \in F$. Analogously, if $\mathbf{y} \in [\mathbf{c}, \mathbf{d}]_d$ for $\mathbf{c}, \mathbf{d} \in F$, then one can write $d(\mathbf{c}, \mathbf{y}) \leq \frac{1}{2}d(\mathbf{c}, \mathbf{d}) \leq \mathrm{diam}(F)$. Hence we obtain $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{a}) + d(\mathbf{a}, \mathbf{c}) + d(\mathbf{c}, \mathbf{y}) \leq 2\mathrm{diam}(F)$. Now if the set F has diameter h, then the set obtained from F by a single process of d-segment joining has a diameter not larger than $2h$; the next step of d-segment joining yields a set of diameter not larger than $4h$, and so on. Hence a finite number of iterations of the process of d-segment joining yields again a bounded set. So, starting with the unit ball B from Example A.1.5, a finite number of such iterations cannot give the d-convex hull $\mathrm{conv}_d(B) = \mathbb{R}^3$ of B.

Anyway, countably many iterations of the process of d-segment joining yield the d-convex hull, according to the following proposition.

Theorem A.3.7. *Let $F_0 \subseteq \mathbb{R}^n$ be an arbitrary set. If F_i denote the set obtained from F_{i-1} by the process of d-segment joining, with $i = 1, 2, \ldots$, then $\mathrm{conv}_d(F) = \bigcup_{i=1}^{\infty} F_i$.*

Proof. By $F_i \subseteq \mathrm{conv}_d(F)$ for each $i = 1, 2, \dots$, we have $\bigcup_{i=1}^{\infty} F_i \subseteq \mathrm{conv}_d(F_0)$. To prove the converse inclusion, it suffices to verify that $\bigcup_{i=1}^{\infty} F_i$ is a d-convex set, since $\mathrm{conv}_d(F_0)$ is the smallest d-convex set containing F_0. For any $\mathbf{x}, \mathbf{y} \in \bigcup_{i=1}^{\infty} F_i$ there are some index i and j such that $\mathbf{x} \in F_i$, $\mathbf{y} \in F_j$, and by $F_0 \subseteq F_l \subseteq F_2 \subseteq \dots$ we have that $\mathbf{x}, \mathbf{y} \in F_k$, with $k = \max\{i, j\}$. By the introduced construction we conclude that $[\mathbf{x}, \mathbf{y}]_d \subseteq F_{k+1} \subseteq \bigcup_{i=1}^{\infty} F_i$, hence $\bigcup_{i=1}^{\infty} F_i$ is a d-convex set. □

A.3.2 Metric convex function and Jensen's inequality

Let (X, d) be a convex metric space in the sense of Menger, $B \subseteq X$, $\mathcal{C}_B = \{C \subseteq X : C\, d\text{-convex}\, C \supseteq B\}$ and let $\mathrm{conv}(B) = \bigcap_{C \in \mathcal{C}_B} C$ denote the metric convex hull of B in the sequel. Now we introduce, according to next definition, some useful sets, defined by recursion starting from the set B, characterizing in detail the metric convex hull of the set. Afterwards we show the validity of a Jensen's inequality for a d-convex function defined in a d-convex subset of a metric space, i.e. generalizing condition (A.3) to any number of points.

Definition A.3.2. *Let (X, d) be a convex metric space in the sense of Menger, $m, n \in \mathbb{N}$, $m \geq 1$, $n \geq 0$, $y_1, \dots, y_m \in X$ and $B \subseteq X$. Then following sets are defined inductively:*

i) $\mathrm{co}^{(0)}(\{y_1, \dots, y_m\}) = \{y_1, \dots, y_m\}$,

$\mathrm{co}^{(n+1)}(\{y_1, \dots, y_m\}) = \{z \in X : \exists p_1, \dots, p_{n+2} \in \{y_1, \dots, y_m\},$

$\exists z', z'' \in \mathrm{co}^{(n)}(\{p_1, \dots, p_{n+2}\}), z \in [z', z'']_d\}$;

ii) $\mathrm{co}^{(0)}(B) = B$,

$\mathrm{co}^{(n+1)}(B) = \{z \in X : \exists x_1, \dots, x_{n+2} \in B, \exists z', z'' \in \mathrm{co}^{(n)}(\{x_1, \dots,$

$x_{n+2}\}), z \in [z', z'']_d\}$;

iii) $\widetilde{\mathrm{co}}^{(0)}(B) = B$,

$$\widetilde{\mathrm{co}}^{(n+1)}(B) = \bigcup_{x,y \in \widetilde{\mathrm{co}}^{(n)}(B)} [x, y]_d.$$

These sets have good properties, and in particular, their union is equal to the d-convex hull of B, as shown in the following lemma and proposition.

Lemma A.3.1. *Let (X,d) be a convex metric space in the sense of Menger, $m \in \mathbb{N}$, $m \geq 1$, $y_1, \ldots, y_m \in X$ and $A, B \subseteq X$. Then*

j) $A \subseteq B \Longrightarrow \operatorname{co}^{(n)}(A) \subseteq \operatorname{co}^{(n)}(B)$, $\widetilde{\operatorname{co}}^{(n)}(A) \subseteq \widetilde{\operatorname{co}}^{(n)}(B)$, *for all* $n \in \mathbb{N}$;

jj) $\operatorname{co}^{(n)}(\{y_1, \ldots, y_m\}) \subseteq \operatorname{co}^{(n+1)}(\{y_1, \ldots, y_m\})$, $\operatorname{co}^{(n)}(B) \subseteq \operatorname{co}^{(n+1)}(B)$, $\widetilde{\operatorname{co}}^{(n)}(B) \subseteq \widetilde{\operatorname{co}}^{(n+1)}(B)$, *for all* $n \in \mathbb{N}$;

jjj) $\operatorname{co}^{(n)}(B) \subseteq \widetilde{\operatorname{co}}^{(n)}(B)$, *for all* $n \in \mathbb{N}$.

Proof. *j)* obvious by definitions and induction;

jj) $\operatorname{co}^{(0)}(\{y_1, \ldots, y_m\}) \subseteq \operatorname{co}^{(1)}(\{y_1, \ldots, y_m\})$ is obvious, i.e. the thesis for $n = 1$; let $\operatorname{co}^{(n)}(\{y_1, \ldots, y_m\}) \subseteq \operatorname{co}^{(n+1)}(\{y_1, \ldots, y_m\})$ be true for $n \in \mathbb{N}$, $n \geq 1$. If $z \in \operatorname{co}^{(n+1)}(\{y_1, \ldots, y_m\})$, then there exist: $p_1, \ldots, p_{n+2} \in \{y_1, \ldots, y_m\}$, $z', z'' \in \operatorname{co}^{(n)}(\{p_1, \ldots, p_{n+2}\})$ and $z \in [z', z'']_d$; setting $p_{n+3} = p_{n+2}$, there exist: $p_1, \ldots, p_{n+3} \in \{y_1, \ldots, y_m\}$, $z', z'' \in \operatorname{co}^{(n)}(\{p_1, \ldots, p_{n+3}\}) \subseteq \operatorname{co}^{(n+1)}(\{p_1, \ldots, p_{n+3}\})$, and $z \in [z', z'']$, i.e., $z \in \operatorname{co}^{(n+2)}(\{y_1, \ldots, y_m\})$, so the thesis is proved by induction. If $B \subseteq X$, then $\operatorname{co}^{(0)}(B) \subseteq \operatorname{co}^{(1)}(B)$ is obvious; let $\operatorname{co}^{(n)}(B) \subseteq \operatorname{co}^{(n+1)}(B)$ be true for $n \in \mathbb{N}$, $n \geq 1$. If $z \in \operatorname{co}^{(n+1)}(B)$, then there exist: $x_1, \ldots, x_{n+2} \in B$, $z', z'' \in \operatorname{co}^{(n)}(\{x_1, \ldots, x_{n+2}\})$ and $z \in [z', z'']_d$; setting $x_{n+3} = x_{n+2}$, there exist: $x_1, \ldots, x_{n+3} \in B$, $z', z'' \in \operatorname{co}^{(n)}(\{x_1, \ldots, x_{n+3}\}) \subseteq \operatorname{co}^{(n+1)}(\{x_1, \ldots, x_{n+3}\})$ and $z \in [z', z'']_d$, i.e. $z \in \operatorname{co}^{(n+2)}(B)$, so the thesis is proved by induction. Finally, $\widetilde{\operatorname{co}}^{(0)}(B) \subseteq \widetilde{\operatorname{co}}^{(1)}(B)$ is obvious; let $\widetilde{\operatorname{co}}^{(n)}(B) \subseteq \widetilde{\operatorname{co}}^{(n+1)}(B)$ be true for $n \in \mathbb{N}$, $n \geq 1$. If $z \in \widetilde{\operatorname{co}}^{(n+1)}(B)$, then there exist: $x, y \in \widetilde{\operatorname{co}}^{(n)}(B)$, $z \in [x, y]_d$, and again the thesis is proved by induction;

jjj) $\operatorname{co}^{(0)}(B) \subseteq \widetilde{\operatorname{co}}^{(0)}(B)$ is obvious; let $\operatorname{co}^{(n)}(B) \subseteq \widetilde{\operatorname{co}}^{(n)}(B)$ be true for $n \in \mathbb{N}$, $n \geq 1$. If $z \in \operatorname{co}^{(n+1)}(B)$, then there exist: $x_1, \ldots, x_{n+2} \in B$, $z', z'' \in \operatorname{co}^{(n)}(\{x_1, \ldots, x_{n+2}\})$, $z \in [z', z'']_d$; by $\operatorname{co}^{(n)}(\{x_1, \ldots, x_{n+2}\}) \subseteq \operatorname{co}^{(n)}(B) \subseteq \widetilde{\operatorname{co}}^{(n)}(B)$, we have $z', z'' \in \widetilde{\operatorname{co}}^{(n)}(B)$ and $z \in \widetilde{\operatorname{co}}^{(n+1)}(B)$, i.e. the thesis is proved by induction. □

Proposition A.3.3. *Let (X,d) be a convex metric space in the sense of Menger and $B \subseteq X$. Then*

$$\operatorname{conv}(B) = \bigcup_{n \in \mathbb{N}} \operatorname{co}^{(n)}(B) = \bigcup_{n \in \mathbb{N}} \widetilde{\operatorname{co}}^{(n)}(B).$$

Proof. Let us prove that $\operatorname{conv}(B) \subseteq \bigcup_{n \in \mathbb{N}} \operatorname{co}^{(n)}(B)$, noting that we just need to prove d-convexity of $\bigcup_{n \in \mathbb{N}} \operatorname{co}^{(n)}(B)$. Let $x, y \in \bigcup_{n \in \mathbb{N}} \operatorname{co}^{(n)}(B)$, so there exists $\bar{n} \in \mathbb{N}$ such that $x, y \in \operatorname{co}^{(\bar{n})}(B)$; if $\bar{n} = 0$ we have $[x, y]_d \subseteq \operatorname{co}^{(1)}(B)$; if $\bar{n} \geq 1$, then there exist: $x_1, \ldots, x_{\bar{n}+1} \in B$, $y_1, \ldots, y_{\bar{n}+1} \in B$, $z', z'' \in$

$\mathrm{co}^{(\bar{n}-1)}(\{x_1,\ldots,x_{\bar{n}+1}\})$, $w', w'' \in \mathrm{co}^{(\bar{n}-1)}(\{y_1,\ldots,y_{\bar{n}+1}\})$, with $x \in [z', z'']_d$ and $y \in [w', w'']_d$; observing that

$$\begin{aligned}&\mathrm{co}^{(\bar{n})}(\{x_1,\ldots,x_{\bar{n}+1},y_1,\ldots,y_{\bar{n}+1}\}) =\\ &\{z \in X : \exists p_1,\ldots,p_{\bar{n}+1} \in \{x_1,\ldots,x_{\bar{n}+1},y_1,\ldots,y_{\bar{n}+1}\},\\ &\exists z', z'' \in \mathrm{co}^{(\bar{n}-1)}(\{p_1,\ldots,p_{\bar{n}+1}\}), z \in [z', z'']_d\},\end{aligned}$$

and

$$\begin{aligned}&\mathrm{co}^{(\bar{n})}(\{x_1,\ldots,x_{\bar{n}+1},y_1,\ldots,y_{\bar{n}+1}\}) =\\ &\{z \in X : \exists p_1,\ldots,p_{\bar{n}+1} \in \{x_1,\ldots,x_{\bar{n}+1},y_1,\ldots,y_{\bar{n}+1}\},\\ &\exists w', w'' \in \mathrm{co}^{(\bar{n}-1)}(\{p_1,\ldots,p_{\bar{n}+1}\}), z \in [w', w'']_d\},\end{aligned}$$

then we have

$$\begin{aligned}x, y \in &\mathrm{co}^{(\bar{n})}(\{x_1,\ldots,x_{\bar{n}+1},y_1,\ldots,y_{\bar{n}+1}\}) \subseteq\\ &\mathrm{co}^{(2\bar{n})}(\{x_1,\ldots,x_{\bar{n}+1},y_1,\ldots,y_{\bar{n}+1}\}),\end{aligned}$$

so by

$$\begin{aligned}&\mathrm{co}^{(2\bar{n}+1)}(B) = \{z \in X : \exists x_1,\ldots,x_{\bar{n}+1},y_1,\ldots,y_{\bar{n}+1} \in B,\\ &\exists x, y \in \mathrm{co}^{(2\bar{n})}(\{x_1,\ldots,x_{\bar{n}+1},y_1,\ldots,y_{\bar{n}+1}\})\},\end{aligned}$$

it follows that $z \in \mathrm{co}^{(2\bar{n}+1)}(B)$, i.e. $[x,y]_d \subseteq \mathrm{co}^{(2\bar{n}+1)}(B) \subseteq \bigcup_{n\in\mathbb{N}} \mathrm{co}^{(n)}(B)$, proving d-convexity of $\bigcup_{n\in\mathbb{N}} \mathrm{co}^{(n)}(B)$. The inclusion $\bigcup_{n\in\mathbb{N}} \mathrm{co}^{(n)}(B) \subseteq \bigcup_{n\in\mathbb{N}} \widetilde{\mathrm{co}}^{(n)}(B)$ immediately follows by *jjj)* of Lemma A.3.1. Finally, let $\mathcal{C}(B)$ the family of d-convex sets $C \subset X$ such that $C \supseteq B$, and let us prove that $C \supseteq \bigcup_{n\in\mathbb{N}} \widetilde{\mathrm{co}}^{(n)}(B)$, for any $C \in \mathcal{C}(B)$. Indeed, fixed $C \in \mathcal{C}(B)$, by $\widetilde{\mathrm{co}}^{(0)}(B) = B \subseteq C$, we have $\widetilde{\mathrm{co}}^{(1)}(B) \subseteq C$, and assuming $\widetilde{\mathrm{co}}^{(n)}(B) \subseteq C$, by d-convexity, $\widetilde{\mathrm{co}}^{(n+1)}(B) \subseteq C$ holds, so that the inclusion $\bigcup_{n\in\mathbb{N}} \widetilde{\mathrm{co}}^{(n)}(B) \subseteq C$ is true for all $C \in \mathcal{C}(B)$. It follows that $\bigcup_{n\in\mathbb{N}} \widetilde{\mathrm{co}}^{(n)}(B) \subseteq \bigcap_{\mathcal{C}}(B)C = \mathrm{conv}(B)$, concluding the proof. □

Finally, we prove the validity of the Jensen's inequality for a d-convex function.

Proposition A.3.4 (Jensen's inequality). *Let (X,d) be a convex metric space in the sense of Menger, $K \subseteq X$ a d-convex set and $f : K \to \mathbb{R}$ a d-convex function. Then, for all $m \in \mathbb{N}$, $m \geq 1$, for all $\{y_1,\ldots,y_m\} \subseteq K$, for all $n \in \mathbb{N}$, $n \geq 1$, for all $x \in \widetilde{\mathrm{co}}^{(n)}(\{y_1,\ldots,y_m\})$, there exist an integer $k \in \mathbb{N}$, $1 \leq k \leq 2^n$, and there exist*

$y_{i_1}, \ldots, y_{i_k} \in \{y_1, \ldots, y_m\}$ *and* $\lambda_1, \ldots, \lambda_k \in [0,1]$ *such that*

$$f(x) \leqslant \sum_{j=1}^{k} \lambda_j f(y_{i_j}). \tag{A.4}$$

Proof. If $n = 1$, let $i_1, i_2 \in \{1, \ldots, m\}$ and $x \in [y_{i_1}, y_{i_2}]_d$, then we have $f(x) \leqslant \lambda_1 f(y_1) + \lambda_2 f(y_2)$, with $\lambda_1 = \frac{d(x,y_2)}{d(y_1,y_2)}$ and $\lambda_2 = \frac{d(y_1,x)}{d(y_1,y_2)}$. Now let (A.4) be true for $n \in \mathbb{N}$, $n \geq 2$; let $x \in \widetilde{\text{co}}^{(n+1)}(\{y_1, \ldots, y_m\})$, so there exist $z_1, z_2 \in \widetilde{\text{co}}^{(n)}(\{y_1, \ldots, y_m\})$ such that $x \in [z_1, z_2]_d$. Then there exist $\bar{\lambda}_1, \bar{\lambda}_2 \in [0,1]$ such that, by induction,

$$\begin{aligned} f(x) &\leqslant \bar{\lambda}_1 f(z_1) + \bar{\lambda}_2 f(z_2) \\ &\leqslant \bar{\lambda}_1 \sum_{h=1}^{h_1} \tilde{\lambda}_{i_h} f(y_{i_h}) + \bar{\lambda}_2 \sum_{k=1}^{k_1} \tilde{\lambda}_{i_k} f(y_{i_k}), \end{aligned}$$

with $h_1, k_1 \leq 2^{n-1}$, and setting $\lambda_{i_h} = \bar{\lambda}_1 \tilde{\lambda}_{i_h}$, $\lambda_{i_k} = \bar{\lambda}_2 \tilde{\lambda}_{i_k}$, the thesis holds. $\square$

We conclude proving an extension theorem for d-convex functions defined in a subset B of a metric space (X, d), in the sense that the condition (A.3.1) holds whenever, for $z \in [x, y]_d$, with $x, y \in B$, then $z \in B$.

Theorem A.3.8. *Let (X, d) be a convex metric space in the sense of Menger, $B \subseteq X$, and $f : X \to \mathbb{R}$ a d-convex function. Then there exists a d-convex function $\bar{f}$: $conv(B) \to \mathbb{R}$ such that $\bar{f}(z) = f(z)$, for all $z \in B$.*

Proof. Let $x \in \text{conv}(B)$. If it where $x \in B$, by Proposition A.3.4, then we could find $n, k \in \mathbb{N}$, with $k \geq 1$, $k \leq 2^n$, and $x_1, \ldots, x_k \in B$, $\lambda_1, \ldots, \lambda_k \in [0,1]$, $\sum_{i=1}^{k} \lambda_i = 1$, such that $f(x) \leqslant \sum_{i=1}^{k} \lambda_i f(x_i)$. So we define a function $\bar{f}$: $\text{conv}(B) \to [-\infty, +\infty[$ as follows: $\bar{f}(x) = \inf \left\{ \sum_{i=1}^{n} \lambda_i f(x_i) \right\}$, with λ_i, x_i as above. Clearly $\bar{f}(x) = f(x)$ if $x \in B$. It remains to check that $\bar{f}$ is d-convex. Let $x_1, x_2 \in \text{conv}(B)$ and $x \in [x_1, x_2]_d$; then there exists $\bar{n} \in \mathbb{N}$ such that $x_1, x_2 \in \widetilde{\text{co}}^{(\bar{n})}(B)$, so that $x \in \widetilde{\text{co}}^{(\bar{n}+1)}(B)$; also, for $x_1, x_2 \in \widetilde{\text{co}}^{(\bar{n})}(B)$, there exist $n_1, n_2 \in \mathbb{N}$, $x_1^1, \ldots, x_{n_1}^1, x_1^2, \ldots, x_{n_2}^2 \in B$, $\lambda_1^1, \ldots, \lambda_{n_1}^1, \lambda_1^2, \ldots, \lambda_{n_2}^2 \in [0,1]$ with sum 1, and there exist $\lambda, \mu \in [0,1]$, $\lambda + \mu = 1$, such that from the sum

$$\underbrace{\sum_{i=1}^{n_1} \lambda \lambda_i^1 f(x_i^1)}_{\mathrm{I}} + \underbrace{\sum_{i=1}^{n_2} \mu \lambda_i^2 f(x_i^2)}_{\mathrm{II}},$$

it follows $\bar{f}(x) \leqslant \mathrm{I} + \mathrm{II}$, and consequently the thesis. $\square$

Bibliography

[1] M. Fréchet. “Sur quelques points du calcul fonctionnel”. In: *Rendiconti del Circolo Matematico di Palermo (1884-1940)* 22.1 (1906), pp. 1–72.

[2] E. W. Chittenden. “On the metrization problem and related problems in the theory of abstract sets”. In: *Bulletin of the American Mathematical Society* 33.1 (1927), pp. 13–34.

[3] A. H. Frink. “Distance functions and the metrization problem”. In: *Bulletin of the American Mathematical Society* 43.2 (1937), pp. 133–143.

[4] V. Schroeder. “Quasi-metric and metric spaces”. In: *Conformal Geometry and Dynamics of the American Mathematical Society* 10.18 (2006), pp. 355–360.

[5] K. Chrząszcz, J. Jachymski, and F. Turoboś. “Two refinements of Frink’s metrization theorem and fixed point results for Lipschitzian mappings on quasimetric spaces”. In: *Aequationes mathematicae* 93.1 (2019), pp. 277–297.

[6] R. Fagin, R. Kumar, and D. Sivakumar. “Comparing Top k lists”. In: *SIAM Journal on discrete mathematics* 17.1 (2003), pp. 134–160.

[7] R. R. Coifman and G. L. Weiss. *Analyse harmonique non-commutative sur certains espaces homogénes: étude de certaines intégrales singuliéres*. Lecture notes in mathematics 242. Berlin, New York: Springer-Verlag, 1971. ISBN: 9783540057031.

[8] P. Urysohn. “Über die Mächtigkeit der zusammenhängenden Mengen”. In: *Math. Ann.* 94 (1925), pp. 262–295.

[9] H. Lebesgue. “Sur le problème de Dirichlet”. In: *Rend. del Circ. Mat. di Palermo* 24 (1907), pp. 371–402.

[10] H. Tietze. “Über Funktionen, die auf einer abgeschlossenen Menge stetig sind”. In: *J. für die reinenund angew. Math.* 145 (1915), pp. 95–102.

[11] H. Whitney. “Analitic Extensions of Differentiable Functions defined in Closed Sets”. In: *Trans. American Mathematical Society* 36.1 (1934), pp. 63–89.

[12] C. Fefferman. “A sharp form of Whitney’s extension theorem”. In: *Annals of Mathematics* 161 (2005), pp. 509–577.

[13] E. M. Stein. *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, 1970.

[14] H. Federer. *Geometric Measure Theory*. Classics in Mathematics. Springer, 1996. ISBN: 9783642620102.

[15] G. Arena, A. O. Caruso, and A. Causa. "Taylor formula on step two Carnot groups". In: *Revista Matemática Iberoamericana* 26 (2010), pp. 239–259.

[16] K. Menger. "Untersuchungen über allgemeine Metrik". In: *Mathematische Annalen* 100.1 (1928), pp. 75–163.

[17] L. M. Blumenthal. *Theory and applications of distance geometry*. 2d ed. Bronx, N.Y: Chelsea Pub. Co, 1970. ISBN: 9780828402422.

[18] S. Willard. *General Topology*. Addison - Wesley Publishing Company, 1970. ISBN: 9780486434797.

[19] D. Mitrea et al. *Groupoid Metrization Theory: With Applications to Analysis on Quasi-metric Spaces and Functional Analysis*. Boston: Birkhäuser Boston, 2013. ISBN: 9780817683962.

[20] T. Suzuki. "Basic inequality on ab-metric space and its applications". In: *Journal of inequalities and applications* 2017.1 (2017), pp. 1–11.

[21] W. H. Schikhof. *Ultrametric calculus: An introduction to p-adic analysis*. Vol. 4. Cambridge University Press, 2010. ISBN: 0521032873.

[22] S. Semmes. "Some basic aspects of analysis on metric and ultrametric spaces". In: *arXiv preprint arXiv:1306.2421* (2013).

[23] N. J. Kalton, N. T. Peck, and J. W. Roberts. *An F-space sampler*. Vol. 89. Cambridge University Press Archive, 1984. ISBN: 9780511662447.

[24] G. Köthe. *Topological vector spaces*. Translated by Garling, D. J. H. Springer - Verlag, 1969. ISBN: 9780387045092.

[25] W. Kirk and N. Shahzad. *Fixed Point Theory in Distance Spaces*. Springer, Cham, 2014. ISBN: 9783319109275.

[26] H. Aimar, B. Iaffei, and L. Nitti. "On the Macìas - Segovia metrization of quasi-metric spaces". In: *Revista de la Unión Matemática Argentina* 41.2 (1998), pp. 67–76.

[27] R. Engelking. *General Topology*. Heldermann Verlag Berlin, 1989. ISBN: 9783885380061.

[28] M. Paluszyński and K. Stempak. "On quasi-metric and metric spaces". In: *Proceedings of the American Mathematical Society* 137 (2009), pp. 4307–4312.

[29] T. Van An, L. Q. Tuyen, and N. Van Dung. "Stone-type theorem on b-metric spaces and applications". In: *Topology and its Applications* 185 (2015), pp. 50–64.

[30] P. Kumam, N. V. Dung, and V. T. L. Hang. "Some equivalences between cone-metric spaces and-metric spaces". In: *Abstract and Applied Analysis*. Vol. 2013. Hindawi. 2013.

[31] T. V. An and N. V. Dung. "Remarks on Frink's metrization technique and applications". In: *arXiv preprint arXiv:1507.01724* (2015).

[32] R. A. Macìas and C. Segovia. "Lipschitz Functions on Spaces of Homogeneous Type". In: *Advances in Mathematics* 33 (1979), pp. 257–270.

[33] I. Genebashvili et al. *Weight theory for integral transforms on spaces of homogeneous type*. Pitman Monographs and Surveys in Pure and Applied Mathematics. CRC Press, 1998. ISBN: 0582302951.

[34] D. Yang, D. Yang, and G. Hu. *The Hardy Space H1 with Non-doubling Measures and Their Applications*. Vol. 2084. Lecture Notes in Mathematics. Cham: Springer International Publishing, 2013. ISBN: 9783319008257.

[35] T. Hytönen. "A framework for non-homogeneous analysis on metric spaces, and the RBMO space of Tolsa". In: *Publicacions Matemàtiques* 54 (2010), pp. 485–504.

[36] J. Soria and P. Tradacete. "The least doubling constant of a metric measure space". In: *arXiv preprint arXiv:1806.06758* (2018).

[37] J. Heinonen. *Lectures on Analysis on Metric Spaces*. Universitext. Springer New York, 2001. ISBN: 9781461301318.

[38] L. Ambrosio and P. Tilli. *Selected Topics on Analysis in Metric Spaces*. Scuola Normale Superiore. Pisa, 2000.

[39] J. Luukkainen and E. Saksman. "Every complete doubling metric space carries a doubling measure". In: *Proceedings of the American Mathematical Society* 126.2 (1998), pp. 531–534.

[40] A. Käenmäki, T. Rajala, and V. Suomala. "Existence of doubling measures via generalised nested cubes". In: *Proceedings of the American Mathematical Society* 140.9 (2012), pp. 3275–3281.

[41] E. Saksman. "Remarks on the nonexistence of doubling measures". In: *Annales-Academiae Scientiarum Fennicae Series A1 Mathematica*. Vol. 24. Academia scientiarum Fennica. 1999, pp. 155–164.

[42] F. M. Christ. "A T(b) theorem with remarks on analytic capacity and the Cauchy integral". en. In: *Colloquium Mathematicum* 60.2 (1990), pp. 601–628. ISSN: 0010-1354, 1730-6302.

[43] F. M. Christ. *Lectures on singular integral operators*. Regional conference series in mathematics no. 77. Providence, R.I: Published for the Conference Board of the Mathematical Sciences by the American Mathematical Society, 1990. ISBN: 9780821807286.

[44] H. Whitney. "Analitic Extensions of Differentiable Functions defined in Closed Sets". In: *Trans. American Mathematical Society* 36.1 (1934), pp. 63–89.

[45] N. G. Meyers and J. Serrin. "$H = W$". In: *Proc. Nat. Acad. Sci. U.S.A.* 51 (1964), pp. 1055–1056.

[46] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. Studies in Advanced Mathematics. CRC PRESS, 1999.

[47] E. J. Mickle. "On the extension of a transformation". In: *Bulletin of the American Mathematical Society* 55.2 (1949), pp. 160–164.

[48] V. Caselles, J. M. Morel, and C. Sbert. "An axiomatic approach to image interpolation". In: *IEEE Transactions on image processing* 7.3 (1998), pp. 376–386.

[49] J. M. Steele. "Certifying smoothness of discrete functions and measuring legitimacy of images". In: *Journal of Complexity* 5.3 (1989), pp. 261–270.

[50] B. Dacorogna. *Direct methods in the calculus of variations*. Vol. 78. Springer Science & Business Media, 2007.

[51] F. Memoli, G. Sapiro, and P. Thompson. *Brain and Surface Warping via Minimizing Lipschitz Extensions (PREPRINT)*. Tech. rep. MINNESOTA UNIV MINNEAPOLIS INST FOR MATHEMATICS and ITS APPLICATIONS, 2006.

[52] M. Kirszbraun. "Über die zusammenziehende und Lipschitzsche Transformationen". In: *Fundamenta Mathematicae* 22.1 (1934), pp. 77–108.

[53] F. A. Valentine. "On the extension of a vector function so as to preserve a Lipschitz condition". In: *Bulletin of the American Mathematical Society* 49.2 (1943), pp. 100–108.

[54] F. A. Valentine. "Contractions in non-Euclidean spaces". In: *Bulletin of the American Mathematical Society* 50.10 (1944), pp. 710–713.

[55] F. A. Valentine. "A Lipschitz condition preserving extension for a vector function". In: *American Journal of Mathematics* 67.1 (1945), pp. 83–93.

[56] R. A. Macías and C. Segovia. "A decomposition into atoms of distributions on spaces of homogeneous type". In: *Advances in Mathematics* 33.3 (1979), pp. 271–309.

[57] M. P. Do Carmo. *Riemannian Geometry*. Mathematics: Theory & Applications. Birkhäuser, 1993.

[58] J. M. Lee. *Introduction to Smooth Manifolds*. Vol. 218. Graduate Texts in Mathematics. Springer, 2003.

[59] V. S. Varadarajan. *Lie Groups, Lie Algebras, and Their Representations*. Vol. 102. Graduate Texts in Mathematics. Springer-Verlag New York, Inc., 1984.

[60] B. Franchi, R. Serapioni, and F. Serra Cassano. “On the Structure of Finite Perimeter Sets in Step 2 Carnot Groups”. In: *The Journal of Geometric Analysis* 13.3 (2003), pp. 421–466.

[61] A. Bellaïche and J. J. Risler. *Sub-Riemannian Geometry*. Vol. 144. Progress in Mathematics. Birkhäuser, 1996.

[62] G. B. Folland and E. M. Stein. *Hardy Spaces On Homogeneus groups*. Vol. 28. Mathematical Notes. Princeton University Press and University of Tokyo Press, 1982.

[63] B. Franchi, R. Serapioni, and F. Serra Cassano. “Rectifiability and perimeter in the Heisenberg Group”. In: *Math. Ann.* 321 (2001), pp. 479–531.

[64] Caruso A. O. and P. Palmisano. “Taylor Formula on step two - Carnot Groups and applications”. In: *preprint* (2022).

[65] L. M. Blumenthal. *Theory and applications of distance geometry*. Vol. 242. Chelsea Publishing Company, Incorporated, 1970.

[66] W. A. Kirk. “Caristi’s fixed point theorem and metric convexity”. In: *Colloquium Mathematicum*. Vol. 36. 1. Institute of Mathematics Polish Academy of Sciences. 1976, pp. 81–86.

[67] W. Takahashi. “A convexity in metric space and nonexpansive mappings, I”. In: *Kodai Mathematical Seminar Reports*. Vol. 22. 2. Department of Mathematics, Tokyo Institute of Technology. 1970, pp. 142–149.

[68] R. Khalil. “Best approximation in metric spaces”. In: *Proceedings of the American mathematical society* 103.2 (1988), pp. 579–586.

[69] T. D. Narang and S. Tejpal. “Equivalence of various types of convexities in metric spaces”. In: *Bull. Gauhati University Math. Association* 9 (2006), pp. 54–61.

[70] V. P. Soltan and P. S. Soltan. “d-Convex functions”. In: *Doklady Akademii Nauk* 249.3 (1979), pp. 555–558.

[71] V. Boltyanski, H. Martini, and P. S. Soltan. *Excursions into combinatorial geometry*. Springer Science & Business Media, 2012.