Investigations in Hadamard Spaces

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
"Doctor rerum naturalium"
der Georg-August-Universität Göttingen

im Promotionsprogramm Mathematical Sciences
der Georg-August University School of Science (GAUSS)

vorgelegt von
Arian Bërdëllima
aus Albanien

Göttingen, 2020

Betreuungsausschuss:

Prof. Dr. D. Russell Luke
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Max Wardetzky
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Mitglieder der Prüfungskommission:

Referentin:

Prof. Dr. D. Russell Luke
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Korreferent:

Prof. Dr. Max Wardetzky
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Thorsten Hohage
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Gerlind Plonka-Hoch
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Stephan Huckemann
Institut für Mathematische Stochastik
Georg-August-Universität Göttingen

Prof. Dr. Thomas Schick
Mathematisches Institut
Georg-August-Universität Göttingen

Tag der mündichten Prüfung: 27.08.2020

Acknowledgements

I would like to thank my two advisors Prof. Dr. D. Russell Luke and Prof. Dr. Max Wardetzky for being my mentors and my teachers spending countless hours with me on this work. I am grateful to Russell for allowing me to join his research group in the Fall of 2016 and for introducing me to the field of convex analysis and continuous optimization. I am very thankful to Max for sparking my interest in Geometry while attending his lectures in Geometric Processing. This ultimately led to the geometrical flavour of my work. I would like to express my sincere gratitude to all professors whose lectures I followed during the time of my PhD studies. They have equipped me with a strong set of analytical skills and have made me a better mathematician. I cannot leave without mentioning Deutscher Akademischer Austauschdienst (DAAD) to whom I am very thankful for making all this possible by awarding me a fellowship to pursue my PhD studies in Germany. My past and present colleagues deserve a great deal of appreciation. In particular Matt and Yura for many insightful discussions we have had during coffee breaks and for their help in proofreading my manuscript. Thanks also to Anna-Lena for sharing her thesis template with me and to Florian Lauster for the scientific collaboration. I am also grateful to my many friends and relatives.

Last but not least I thank my wonderful family; my mother Nadire, my father Shyqyri and my sisters Lorela and Anisa for being very supportive during these years. I dedicate this work to them!

Naim Frashëri (1846–1900)

Albanian Renaissance

“Punë punë natë e ditë që të shohim pakëz dritë.”

“Work work day and night so we see a little light.”

Contents

List of Figures

CHAPTER 1

Introduction

This thesis investigates the interplay between geometry and convex analysis in Hadamard spaces. Formally a Hadamard space is a complete CAT(0) space, i.e. a metric space of nonpositive curvature in the sense of Alexandrov. The class of CAT(0) spaces is part of the larger set of Alexandrov spaces that were introduced by A. D. Alexandrov [3] in 1951. They are defined by axioms similar to Euclidean geometry, but certain equalities are turned into inequalities. Depending on the sign of the inequality one can get Alexandrov spaces with curvature bounded above or curvature bounded below. Even though the definition of these two classes of spaces are similar their properties and applications are quite different. The work on spaces with curvature bounded above started in the late 1950s with extensive studies carried out during 1960s and 1970s. Mikhail Gromov [53] stressed the importance of Alexandrov's definition of curvature for what it might mean for a metric space to have a curvature bounded above by a real number κ. For these spaces Gromov coined the acronym CAT(κ) from the names E. J. Cartan (1869–1951), A. D. Alexandrov (1912–1999) and V. A. Toponogov (1930–2004) in recognition of their pioneering work. Meanwhile, Hadamard spaces are named after the French mathematician J. Hadamard (1865–1963). The terms a complete CAT(0) space and Hadamard space will often be used interchangibly. For an extensive treatment of CAT(κ) spaces, e.g. see [32, Bridson and Haefliger] or for a self contained material regarding CAT(0) spaces, e.g. see [2, Petrunin et al.].

The work on convex and functional analysis in CAT(0) spaces started in the late 1960s with the work of Reshetnyak [98] and gained a real momentum in the 1990s and early 2000s with the work of Kirk and Panyanak [72], Jost [61], Kohlenbach [73], Reich [95], [97], Reich et al. [96], Sturm [108], [109], and Lopez et al. [8] to mention a few. In the last decade there has been an increasing interest in optimization methods with applications to CAT(0) spaces, e.g. see Bačak [43] and references therein for a general treatment or Owen [93], Owen and Provan [94], Owen et al. [6] for applications to the space of phylogenetic trees, or Ardila [7] for applications of CAT(0) geometry in robotics.

Motivated by numerous applications of CAT(0) geometry, our work builds upon the results in convex analysis and Alexandrov geometry of many previous authors. Our investigations answer several questions in the theory of CAT(0) spaces some of which were posed as open problems in a recent review article by Bačak [12]. In a nutshell our thesis develops along the following lines:

(i) Weak topologies in Hadamard spaces.

(ii) Convex hulls of compact sets.

(iii) Mean tree problem in phylogenetic tree spaces.

(iv) Mosco convergence in Hadamard spaces.

(v) Firmly nonexpansive operators and their applications in Hadamard spaces.

Each of these topics develops into its own chapter. These chapters, though self contained, are connected to each other, e.g. Chapter 5 can be considered as a direct application of the theory developed in Chapter 4. We start with Chapter 2 which lays out the basic concepts needed for the later work. In this chapter our starting point of view is that of a general metric space which then is endowed with a convex structure in the sense of Takahashi [111]. For short we call them convex metric spaces or convex spaces; then we restrict ourselves to a proper subset of convex spaces which admit a convex structure that is jointly convex. It turns out that the set of jointly convex metric spaces coincides with the set of the so called Busemann spaces. The latter have a notion of nonpositive curvature which is weaker than that of Alexandrov. Afterwards we look at jointly convex metric spaces that admit a convex structure that is strongly convex. We show that this is a proper subset of jointly convex spaces and coincides with the set of CAT(0) spaces or equivalently Hadamard spaces if we let them be complete. This top-down approach is depicted in Figure 1.1:

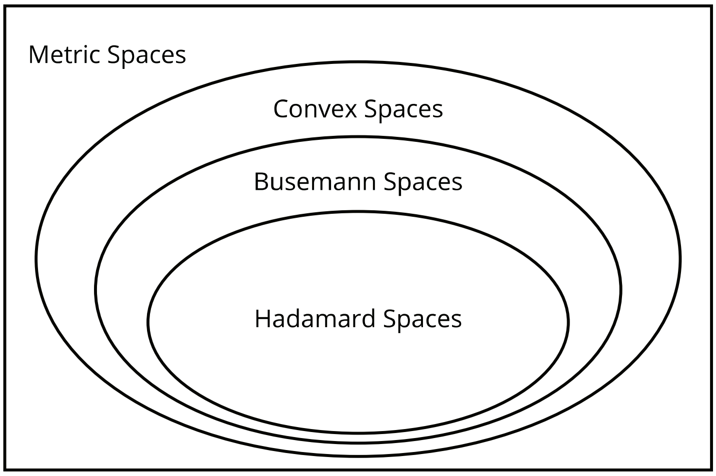


Figure 1.1: Hadamard spaces as a proper subset of Busemann spaces.

In Chapter 2 we also introduce the basic geometrical elements for defining CAT(κ) spaces. The original notion of curvature of Alexandrov in terms of comparison triangles can be understood also as a generalization of the notion of sectional curvature in Riemannian manifolds. In fact a Riemannian manifold has sectional curvature bounded above if and only if it has a curvature bounded above in the sense of Alexandrov [32, Theorem 1A.6]. Moreover Hilbert spaces are known to be the only Banach spaces that are Hadamard [32, Proposition II.1.14]. Therefore Hadamard spaces can be regarded as a generalization of Hilbert spaces and Riemannian manifolds of nonpositive sectional curvature. This can be viewed as a bottom-up approach, see Figure 1.2:

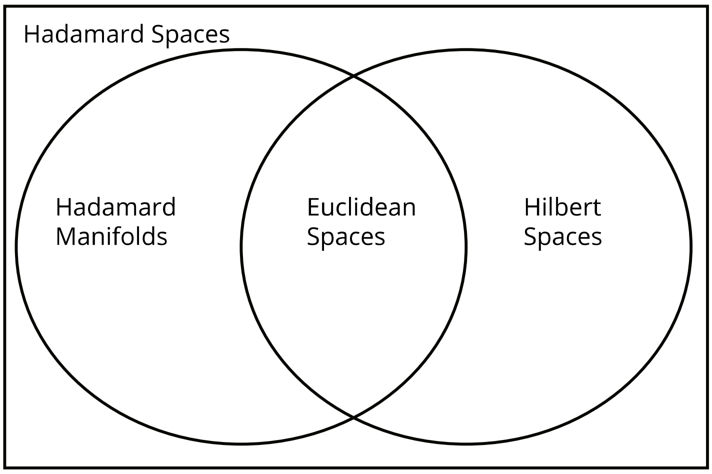


Figure 1.2: Hadamard space as a more general space.

Chapter 3 investigates the problem of identifying a topology which corresponds to the notion of Δ-convergence in Hadamard spaces. Lim [80] introduced Δ-convergence as a concept of weak convergence in a general metric space (X, d). This concept was adopted later by Kirk and Panyanak [72] in the setting of CAT(0) spaces. Jost [61] introduced a notion of weak convergence in CAT(0) spaces, which was rediscovered by Espínola and Fernández-León [50], who also proved that it is equivalent to Δ-convergence. In his work Bačak refers to it as simply the weak convergence. However we save this name for an equivalent notion of convergence, that along geodesic segments. This latter notion has the advantage that a weakly converging sequence need not be bounded and coincides with Δ-convergence on bounded sets. Motivated by a suggestion of Bačak we construct a topology τ_w, which is weaker than the usual metric topology. It holds that the weak convergence yields convergence in τ_w. Moreover if the underlying Hadamard space is *weakly proper* then convergence in τ_w implies weak convergence and so τ_w would be the correct topology for weak convergence. Although the construction of the weak topology τ_w turns out to be rather simple, its construction has important consequences. We will show how this topology can be used in order to obtain certain compactness results that are similar to those known from the linear setting of Banach spaces. In particular, we prove that in a separable weakly proper Hadamard space satisfying a certain regularity condition weak compactness and weak sequential compactness are equivalent (Theorem 3.22). Moreover a bounded closed convex set C in a separable Hadamard space is weakly compact whether or not the Hadamard space is weakly proper (Theorem 3.23). We also suggest a notion of dual spaces corresponding to this topology. Another contribution in Chapter 3, which has a particular importance in optimization theory, is the existence of a steepest descent direction in a locally compact space (Theorem 3.40) and in a general Hadamard space Theorem 3.41. Next, we offer a comparison of our notion to previous notions of weak topologies in Hadamard spaces. In particular, we will compare to the work of Monod [90], which offers a notion of weak topology that is weaker than ours (at least in the case of weakly proper spaces), but which does not yield convergence along geodesics and is thus too weak for our purpose (Proposition 3.44). Likewise, we will compare the work of Kakavandi [64], who offers a topology that is stronger than ours and for which it is unknown whether compactness results hold in general (Theorem 3.43). Finally, we introduce the notion of a geodesically monotone operator, which to our knowledge is new in the literature. In particular we show that a Hadamard space equipped with Monod's weak topology is Hausdorff whenever projections onto geodesic segments are geodesically monotone operators (Theorem 3.46). We show by example that in general in Hadamard spaces projections are not geodesically monotone. This is in contrast with Hilbert spaces where projections onto geodesic segments are always monotone.

In Chapter 4 we look at the problem of the closure of convex hulls of compact sets in Hadamard spaces. It is known that in a Banach space X given a compact set $K \subset X$ the closure of its convex hull $\operatorname{cl}\operatorname{co} K$ is compact [5, Theorem 5.35]. It is not known whether such a result carries over to Hadamard spaces. The problem remains widely open even for the simplest case of a set of only three points. This was pointed out first by Gromov in [54]. Motivated by the problem of the mean tree in phylogenetics, where it is can be shown that it lies in the closure of the convex hull of the given set of trees, we study convex hulls of finite sets in Hadamard spaces. We start first in the setting of a locally compact space. It is easy to show that in a locally compact space the closure of the convex hull of a compact set is compact. This fact together with the so called finite set property is equivalent to local compactness (Theorem 4.6). Moreover we introduce the notion of regular space, which to our knowledge is a new concept, and it happens to be an equivalent definition for locally compact spaces (Theorem 4.4). As a direct application of the theory developed in Chapter 3 we obtain in Chapter 4 that in a separable Hadamard space the closure of the convex hull of a bounded weakly compact set is weakly compact (Theorem 4.9) and that in particular the closure of the convex hull of a compact set is weakly compact (Corollary 4.10). In Chapter 4 we also introduce and develop the concept of threading of a set. In particular the operation of threading is well behaved with respect to compact sets, i.e. threading of a compact set is always compact (Lemma 4.18). We show that convex hull of a set can be expressed as the union of threadings of all degrees of this set (Theorem 4.15). In particular for the class of Hadamard spaces of finite type we obtain that the convex hull of a compact set is always compact (Proposition 4.20). Moreover an important application of threading is with respect to the so called Fréchet mean of a given finite set of points. We derive a constructibility theorem which essentially states that the Fréchet mean of a finite set in a Hadamard space of finite type is contructible in at most a finite number of steps and that this number of steps is expressible in terms of the threading degree of the given set (Theorem 4.25).

Chapter 5 can be regarded as an application of the theory of threadings developed in Chapter 4. Our work here is motivated by works of Owen [93], [94] and the seminal paper of Billera, Holmes and Vogtmann [26] about phylogenetic tree spaces. We derive certain results about orthant path spaces (Theorem 5.6) and isometry results (Theorem 5.7, Proposition 5.8). In our view the most important contribution from this chapter is Theorem 5.13 which essentially states that convex hull of a compact set in the tree space of trees with four leaves is a compact set. Moreover the Fréchet mean can be calculated in at most a finite number of steps (Corollary 5.14).

In Chapter 6 our main focal point is Mosco convergence in Hadamard spaces. Here motivated by the works of Bačak [43], [13] we establish sufficient conditions for a sequence of closed convex functions to Mosco converge to some closed convex function. Finding sufficient conditions was posed as an open problem in [13]. In this chapter we introduce the notions of pointwise aymptotic boundedness and uniform asymptotic boundedness for a given sequence of real valued functions, which to our knowledge is new in the literature. We find that if the sequence of functions has asymptotic bounded slope on the Hadamard space and if the corresponding sequence of proximal mappings converges pointwise to the proximal mapping of the limiting function then the sequence of functions Mosco converges to the limiting function (Theorem 6.20). Bačak's work and ours is related to an earlier work of Attouch [11] about Mosco convergence in smooth Banach spaces. However our Theorem 6.20 is not exactly in the flavour of Attouch. To obtain a result which is close to Attouch's Theorem [11, Theorem 3.26] we investigate a normalization condition given by Attouch [11, Theorem 3.26] and derive Theorem 6.23.

Chapter 7 is the last chapter of our thesis. The purpose of Chapter 7 is to lay the foundations for the extension of certain fixed point methods to Hadamard spaces. We follow the framework established in [84] which is built on only two fundamental elements in the Euclidean setting:

(i) pointwise almost α-averaging [84, Definition 2.2];

(ii) and metric regularity [60, Definition 2.1.b].

Almost averaged mappings are, in general, set-valued. In Hadamard spaces, there are several difficulties that arise: first, there is no straight-forward generalization of the averaging property since addition is not defined on Hadamard spaces; and second, multivaluedness, which comes with allowing mappings to be expansive. The issue of multivaluedness introduces technical overhead, but does not, at this early stage, seem to present any conceptual difficulties. We therefore restrict our attention to an appropriate generalization of single-valued, pointwise α-averaged mappings. The main contribution is establishing a calculus for these mappings in Hadamard spaces, showing in particular how the property is preserved under compositions (Theorem 7.9) and convex combinations (Theorem 7.13). This is of central importance to splitting algorithms that are built by such convex combinations and compositions. We then apply this theory in the study of cyclic projections (Theorem 7.16), averaged projections (Theorem 7.20), projected proximal mapping and projected gradient flow. Moreover in Chapter 7 we investigate also metric regularity and generalize a recent theorem of Luke et al. about metric regularity in Euclidean spaces (Theorem 7.24). This result shows the interplay between metric regularity, quasi α-firmly nonexpansiveness and local linear convergence of a given operator in a Hadamard space.

Chapter 2

Basic Concepts

2.1. Convex Metric Spaces

2.1.1 Convex structures

In 1928 Menger [88] started the investigation concerning convexity in metric spaces. This direction of research was continued by many authors (see Blumenthal [27] and references therein). The terms 'metrically convex' and 'convex metric space' appeared first in Blumenthal [27] and were used later by many authors, Lalek and Nitka [79], Borsuk [29], Busemann [42], Kohlenbach [73] to mention a few. In this section we talk about convex metric spaces and discuss some of their fundamental characteristics. Throughout this part (X, d) will denote a metric space. A point $z \in X$ is said to be *between* x, y if $z \neq x, z \neq y$ and $d(x, y) = d(x, z) + d(z, y)$. A metric space (X, d) is said to be *convex* if for every pair $x, y \in X$ with $x \neq y$ there is a point $z \in X$ between x, y. In 1970 Takahashi [112] introduced another notion of convexity into metric spaces, studied properties of such spaces and proved several fixed point theorems for nonexpansive mappings. A mapping $W : X \times X \times [0, 1] \to X$ is said to be a *convex structure* on X if for all $x, y \in X$ and all $t \in [0, 1]$ the following inequality holds

$$d(z, W(x, y, t)) \leqslant td(z, x) + (1 - t)d(z, y), \quad \forall z \in X. \tag{2.1}$$

Such kind of metric spaces with a convex structure W seem to be often called w-*convex metric spaces* (see Shimizu [103], Shimizu and Takahashi [104]). However in this work we simply refer to w-convexity as convexity because the metric spaces we deal with, will always admit a convex structure W. A midpoint of the pair x, y is a point m such that

$$d(x, m) = d(y, m) = \frac{1}{2}d(x, y). \tag{2.2}$$

From inequality (2.1) elementary calculations show (see Proposition 3, [112])

$$d(x, y) = d(x, W(x, y, t)) + d(W(x, y, t), y), \quad \forall x, y \in X, \forall t \in [0, 1]. \tag{2.3}$$

In particular in (2.1) if $t = 1/2$ the point $W(x, y, 1/2)$ satisfies

$$d(z, W(x, y, 1/2)) \leqslant \frac{1}{2}d(z, x) + \frac{1}{2}d(z, y)$$

for all $z \in X$. When $z = x$ and $z = y$ we get respectively $d(x, W(x, y, 1/2)) \leqslant d(x, y)/2$ and $d(y, W(x, y, 1/2)) \leqslant d(x, y)/2$ which together with identity (2.3) implies

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

Therefore in a convex metric space any pair of points x, y has a midpoint m (compare with convexity in Krakus [75], [76]). From this obervation it is also immediate that a w-convex space is convex in the sense of Menger. In general for any dyadic number $t = k/2^l$ for $k = 1, 2, ..., 2^l$ it holds

$$d(x, W(x, y, t)) = td(x, y) \quad \text{and} \quad d(y, W(x, y, t)) = (1 - t)d(x, y). \tag{2.4}$$

It is known that in a complete convex metric space (X, d) every pair of distinct points are joined by a *segment* with endpoints x, y (see [27], p.41). This result was first proved by Menger ([88], p.89 see also Aronszajn [9]). By a *segment* we mean a subset in X containing x and y that is isometric to a real interval of length $l = d(x, y)$. If $\Phi : [0, l] \to H$ is an isometry by completeness of the space and an argument of approximation it follows from 2.4 that $\Phi(t) = W(x, y, t/l)$ for all $t \in [0, l]$.

In a metric space (X, d) with a convex structure W a subset $S \subseteq X$ is convex whenever $W(x, y, t) \in S$ for all $x, y \in S$ and for all $t \in [0, 1]$. It follows by definition that the intersection of an arbitrarily collection of convex sets is convex. The intersection of all the convex sets containing a given subset $S \subseteq X$ is called the *convex hull* of S and it is denoted by $\operatorname{co} S$. A more general class of sets in a convex metric space are the so called *star-shaped* sets. A set S is said to be star-shaped if there exists some element $x_0 \in S$ such that $W(x, x_0, t) \in S$ for all $x \in S$ and all $t \in [0, 1]$. Evidently the set of all star shaped sets of X contains as a proper subset the collection of all convex sets of X (see Azam and Beg [23]).

For a given subset $S \subseteq X$ the *metric projection* onto S is a set-valued mapping on X defined as

$$P_S x := \{y \in S \mid d(x, y) \leqslant d(x, z), \forall z \in S\}. \tag{2.5}$$

Clearly $P_S x = \{x\}$ for any $x \in S$. When S is convex then $P_S x$ is convex for if $y_1, y_2 \in P_S x$ then

$$\begin{aligned} d(x, W(y_1, y_2, t)) &\leqslant (1 - t)d(x, y_1) + td(x, y_2) \\ &\leqslant (1 - t)d(x, z) + td(x, z) = d(x, z), \;\; \forall z \in S, \forall t \in [0, 1] \end{aligned}$$

implies $W(y_1, y_2, t) \in P_S x$.

Given two metric spaces (X_1, d_1) and (X_2, d_2) we let $X := X_1 \times X_2$ be the *product space* equipped with the canonical metric $d(x, y)^2 = d_1(x_1, y_1)^2 + d_2(x_2, y_2)^2$ where $x := (x_1, x_2)$

and $y := (y_1, y_2)$. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be an (extended) real valued function defined on a metric space X. The set

$$\{(x, \nu) \in X \times \mathbb{R} \mid \nu \geqslant f(x)\} \tag{2.6}$$

is called the *epigraph* of f and we denote it by $\operatorname{epi} f$. Evidently the epigraph $\operatorname{epi} f$ is a subset of the product space $X \times \mathbb{R}$. The *effective domain* of f is defined to be the set $\operatorname{dom} f := \{x \in X \mid f(x) < +\infty\}$. A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be *lower-semicontinuous (lsc)* if $f(x) \leqslant \liminf_{y \to x} f(y)$ where convergence $y \to x$ is in the usual metric topology. Analogous definition of lsc can be given for any other topology which the metric space (X, d) might be equipped with. For a given real number $\alpha \in \mathbb{R}$ the set

$$\{x \in \operatorname{dom} f \mid f(x) \leqslant \alpha\} \tag{2.7}$$

is called the *level set* of f and we denote it by $\operatorname{lev}_\alpha f$. It is known that a function f is lsc iff its level sets are closed sets for any α (see Theorem 7.1 [99], p.51). We say a function f is *closed* whenever its level sets $\operatorname{lev}_\alpha$ are closed sets for any $\alpha \in \mathbb{R}$. Motivated by this equivalence often the term lsc function and closed function will be used interchangeably. A function $f : X \to \mathbb{R}$ defined on a convex metric space X is said to be a *convex function* if

$$f(W(x, y, t)) \leqslant t f(x) + (1 - t) f(y), \quad \forall x, y \in X, \forall t \in [0, 1]. \tag{2.8}$$

It follows by definition that a function f is convex iff its epigraph $\operatorname{epi} f$ is convex as a subset of $X \times \mathbb{R}$. Moreover by (2.8) αf is a convex function whenever f is a convex function and $\alpha \geqslant 0$. For any two convex functions f_1, f_2 the sum $f = f_1 + f_2$ is also a convex function. A function $f : X \to \mathbb{R}$ is said to be *strongly convex with parameter* $\mu > 0$ whenever

$$f(W(x, y, t)) \leqslant t f(x) + (1 - t) f(y) - \frac{\mu}{2}(1 - t) t d(x, y)^2, \quad \forall x, y \in X, \forall t \in [0, 1]. \tag{2.9}$$

Obviously a strongly convex function is convex. Moreover strong convexity of functions is preserved under positive scalar multiplication and pointwise addition. In particular if f_1, f_2 are strongly convex functions with parameters μ_1, μ_2 respectively then $f := f_1 + f_2$ is a strongly convex function with parameter $\mu_1 + \mu_2$.

Proposition 2.1. *A metric space (X, d) is trivial if it consists of a single element. There does not exist a nontrivial convex metric space with a strongly convex metric function.*

Proof. Suppose that (X, d) is a metric space with a strongly convex metric function. This means that there is some $\mu > 0$ such that

$$d(z, W(x, y, t)) \leqslant t d(z, x) + (1 - t) d(z, y) - \frac{\mu}{2}(1 - t) t d(x, y)^2, \quad \forall z \in X.$$

For $z = x$ and $z = y$ we get respectively

$$d(x, W(x, y, t)) \leqslant (1 - t) d(x, y) - \frac{\mu}{2}(1 - t) t d(x, y)^2 \tag{2.10}$$

$$d(y, W(x, y, t)) \leqslant t d(y, x) - \frac{\mu}{2}(1 - t) t d(x, y)^2. \tag{2.11}$$

Adding (2.10),(2.11) and using identity (2.3) yield $d(x, y) = d(x, W(x, y, t)) + d(y, W(x, y, t)) \leqslant d(x, y) - \mu(1 - t) t d(x, y)^2$ which is impossible unless $x = y$. This completes the proof. □

2.1.2 Joint convex structures

A mapping $W : X \times X \times [0,1] \to X$ is said to be a *joint convex structure* on a metric space X if for all $x_0, x_1, y_0, y_1 \in X$ and $t \in [0,1]$

$$d(W(x_0, x_1, t), W(y_0, y_1, t)) \leqslant td(x_0, y_0) + (1-t)d(x_1, y_1). \tag{2.12}$$

A metric space admitting such a structure is said to be a *joint convex metric space*. It is clear that a joint convex structure is a convex structure. In particular in a complete joint convex metric space any two points are joined by a segment. A metric space (X, d) is a *Busemann space* if for every $x, y, z \in X$ we have

$$d(m_1, m_2) \leqslant \frac{1}{2} d(x, y) \tag{2.13}$$

where m_1 is a midpoint of x, z and m_2 a midpoint of y, z.

Theorem 2.2. *A complete metric space (X, d) is a Busemann space if and only if it admits a joint convex structure W. In particular a Busemann space is a convex metric space. Moreover for any $x, y \in X$ the segment joining x with y is unique.*

Proof. Let (X, d) be a Busemann space and let $W : X \times X \times [0,1] \to X$ be a mapping defined as $W(x, y, 1/2) := \{m \in X : d(m, x) = d(m, y) = d(x, y)/2\}$ for any $x, y \in X$. Let $x_0, x_1, y_0, y_1 \in X$ and $m \in W(x_0, y_1, 1/2)$. Then we have

$$d(m_1, m_2) \leqslant d(m_1, m) + d(m, m_2) \leqslant \frac{1}{2} d(x_0, y_0) + \frac{1}{2} d(x_1, y_1)$$

for any $m_1 \in W(x_0, x_1, 1/2), m_2 \in W(y_0, y_1, 1/2)$ where the second inequality follows from assumption that X is a Busemann space. In general following midpoint of the midpoint argument one gets

$$d(m_1(t), m_2(t)) \leqslant (1-t)d(x_0, y_0) + td(x_1, y_1)$$

for any $m_1(t) \in W(x_0, x_1, t), m_2(t) \in W(y_0, y_1, t)$ for all t of the form $t = k/2^l$ where $k = 1, 2, ..., 2^l$ and $l \in \mathbb{N}$. Since the dyadic numbers are dense in the reals then completeness of X and an argument of approximation yields that the last inequality holds for all $t \in [0,1]$. Therefore (X, d) admits a joint convex structure W. The reverse implication is obvious. From the same inequality it follows that midpoints are unique. Indeed if $\{m_1, m_2\} \subseteq W(x, y, 1/2)$ then using $x_0 = y_0 = x, x_1 = y_1 = y$

$$0 \leqslant d(m_1, m_2) \leqslant \frac{1}{2} d(x, x) + \frac{1}{2} d(y, y) = 0$$

implies $m_1 = m_2$. Now let $[0, l] \subseteq \mathbb{R}$ be a closed interval of length $l = d(x, y)$. Let $\Phi, \Psi : [0, l] \to X$ be two isometries such that $\Phi(0) = \Psi(0) = x$ and $\Phi(l) = \Psi(l) = y$. Uniqueness of midpoints implies in particular that $\Phi(l/2) = \Psi(l/2) = W(x, y, 1/2)$. An iterative application of midpoint property yields $\Phi(l/4) = \Psi(l/4) = W(x, y, 1/4)$ and $\Phi(3l/4) = \Psi(3l/4) = W(x, y, 3/4)$ and so on. In general we get $\Phi(t) = \Psi(t) = W(x, y, t/l)$ for all $t \in D \cap [0, l]$ where D is the set of dyadic numbers in $\mathbb{R}$. Since $D \cap [0, l]$ is dense in $[0, l]$ then any $t \in [0, l]$ can be successively approximated by a sequence $(t_n)_{n \in \mathbb{N}} \subseteq D \cap [0, l]$. Continuity of the isometries implies $\Phi(t) = \lim_n \Phi(t_n) = \lim_n \Psi(t_n) = \Psi(t)$ therefore $\Phi([0, l]) = \Psi([0, l])$. This completes the proof. □

2.1.3 Strongly convex structures

The mapping $W : X \times X \times [0,1] \to X$ is said to be a *strongly convex structure* on X if for all $x, y \in X$ and $t \in [0,1]$ the following condition holds[1]

$$d(z, W(x,y,t))^2 \leqslant td(z,x)^2 + (1-t)d(z,y)^2 - (1-t)td(x,y)^2, \ \ \forall z \in X. \tag{2.14}$$

A metric space (X,d) with a strongly convex structure W is said to be a *strongly convex metric space*. The following remarks are immediate:

1. A strongly convex structure is a convex structure. In particular a strongly convex metric space is convex;
2. If $(X_1,d_1),(X_2,d_2)$ are strongly convex then so is the product space $X_1 \times X_2$ when equipped with the canonical metric $d(\cdot,\cdot)^2 = d_1(\cdot,\cdot)^2 + d_2(\cdot,\cdot)^2$.

Proposition 2.3. *In a strongly convex metric space the midpoints are unique. If additionally the space is complete then any two points $x, y \in X$ determine a unique segment $[x,y]$.*

Proof. Let (X,d) be strongly convex then by 1 (X,d) is convex. If m_1, m_2 are midpoints of x, y then $\{m_1, m_2\} \subseteq W(x,y,1/2)$. Characterization inequality (2.14) implies

$$d(z,m_i)^2 \leqslant \frac{1}{2}d(z,x)^2 + \frac{1}{2}d(z,y)^2 - \frac{1}{4}d(x,y)^2, \ \ \forall z \in X, i = 1,2.$$

When $z = m_j$ for $j \neq i$ and $j = 1,2$ we obtain

$$d(m_j,m_i)^2 \leqslant \frac{1}{2}d(m_j,x)^2 + \frac{1}{2}d(m_j,y)^2 - \frac{1}{4}d(x,y)^2, \ \ \forall z \in X, i = 1,2.$$

Applying for a second time inequality (2.14) to the terms $d(m_j,x)^2$ and $d(m_j,y)^2$ respectively implies

$$d(m_j,m_i)^2 \leqslant \frac{1}{4}(d(y,x)^2 + d(x,y)^2 - d(x,y)^2) - \frac{1}{4}d(x,y)^2 = \frac{1}{4}d(y,x)^2 - \frac{1}{4}d(x,y)^2 = 0.$$

Therefore $m_1 = m_2$ and so the midpoint for any pair $x, y \in X$ is unique (compare with strong convexity in [76]). Now let $[0,l] \subseteq \mathbb{R}$ be a closed interval of length $l = d(x,y)$. Let $\Phi, \Psi : [0,l] \to X$ be two isometries such that $\Phi(0) = \Psi(0) = x$ and $\Phi(l) = \Psi(l) = y$. Same arguments as in Theorem 2.2 imply $\Phi([0,l]) = \Psi([0,l])$. □

Proposition 2.4. *Let (X,d) be a complete strongly convex metric space and $S \subseteq X$ a closed convex set. Then $P_S x$ is nonempty and unique (i.e. consists of a single element) for every $x \in X$. Moreover the following condition is satisfied*

$$d(x,P_S x)^2 + d(P_S x, z)^2 \leqslant d(x,z)^2, \quad \forall x \in X \setminus S, \forall z \in S \setminus P_S x. \tag{2.15}$$

[1] If we insert in (2.14) a general parameter $\mu/2$ like in (2.9) than it turns out that $\mu \in (0,2]$ and only when $\mu = 2$ the space is guaranteed to be convex, making thus the definition consistent.

Proof. By definition $y \in P_S x$ whenever $d(x,y) \leqslant d(x,z)$ for all $z \in S$ equivalently $y \in \arg\min_{z\in S} d(x,z)$. On the other hand minimizing $d(x,\cdot)$ is the same a minimizing $d(x,\cdot)^2$. We claim $d(x,\cdot)^2$ has at most one minimizer. Indeed let $y_1, y_2 \in \arg\min_{z\in S} d(x,z)^2$ and $y_1 \neq y_2$. For $t \in [0,1], W(y_1,y_2,t) \in S$ since S is convex. Assumption (X,d) is strongly convex implies

$$\begin{aligned} d(x, W(y_1,y_2,t))^2 &\leqslant (1-t)d(x,y_1)^2 + td(x,y_2)^2 - (1-t)td(y_1,y_2)^2 \\ &< (1-t)d(x,y_1)^2 + td(x,y_2)^2 = d(x,y_1)^2 \end{aligned}$$

contradicting that y_1 is a minimizer. Thus $P_S x$ consists of at most one element. Now let $(y_n)_{n\in\mathbb{N}}$ be a minimizing sequence i.e. $\lim_n d(x,y_n)^2 = \inf_{z\in S} d(x,z)^2$. On the other hand by identity (2.1) for any $t \in (0,1)$ we get

$$\lim_{m,n} d(x, W(y_m,y_n,t)) \leqslant \inf_{z\in S} d(x,z)$$

hence $(W(y_m,y_n,t))_{m,n}$ is a minimizing sequence too. By strong convexity it follows then

$$(1-t)td(y_m,y_n)^2 \leqslant (1-t)d(x,y_m)^2 + td(x,y_n)^2 - d(x, W(y_m,y_n,t))^2$$

and in the limit $\lim_{m,n} d(y_m,y_n) = 0$. Therefore $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Assumption (X,d) is complete implies that $\lim_n y_n = y$ for some $y \in X$. By continuity of the metric $y \in S$ and satisfies $d(x,y) = \inf_{z\in S} d(x,z)$. To prove identity (2.15) let $x \in X \setminus S$ and $z \in S \setminus P_S x$. Consider $W(P_S x, z, t)$ for $t \in (0,1)$ then $W(P_S x, z, t) \in S$ since S is convex. Because W is a strongly convex structure then

$$d(x, W(P_S x, z, t))^2 \leqslant (1-t)d(x, P_S x)^2 + td(x,z)^2 - (1-t)td(P_S x, z)^2$$

together with $d(x, P_S x) \leqslant d(x, W(P_S x, z, t))$ implies $td(x,P_S x)^2 + (1-t)td(P_S x, z)^2 \leqslant td(x,z)^2$. Dividing by t and taking limit at $t \downarrow 0$ implies (2.15). □

2.2. Geodesics, Angles and Length of a Curve

2.2.1 Geodesics

In this section we follow standard terminology in metric geometry theory (see [41], [32]). Let (X,d) be a metric space. A *geodesic* joining $x \in X$ to $y \in X$ is a map $\gamma : [0,1] \to X$ such that $\gamma(0) = x, \gamma(1) = y$ and $d(\gamma(t), \gamma(t')) = |t - t'|d(x,y)$ for all $t, t' \in [0,1]$. The image of γ in X is known as the *geodesic segment* between x and y. With some abuse of language geodesic and geodesic segment will be used interchangeably. The metric space (X,d) is said to be a *geodesic metric space* if any two points in X are joined by a geodesic. If this geodesic is unique for every pair of points we say that (X,d) is a *uniquely geodesic metric space*. By [32, Remark 1, p. 4] a complete metric space is a geodesic space if and only if every pair of points has a midpoint. This in turn implies the following immediate result which we present without proof.

Proposition 2.5. *A complete convex metric space is necessarily a geodesic space.*

Examples of geodesic spaces include normed vector spaces, Riemannian manifolds, and polyhedral complexes among others.

2.2.2 Angles

A *comparison triangle* in $\mathbb{E}^2$ for three points $p, q, r \in X$ is a triangle in the Euclidean plane $\mathbb{E}^2$ with vertices $\overline{p}, \overline{q}, \overline{r}$ such that $d(p,q) = \|\overline{p} - \overline{q}\|$, $d(q,r) = \|\overline{q} - \overline{r}\|$ and $d(p,r) = \|\overline{p} - \overline{r}\|$. Such a triangle is unique up to an isometry and we denote it by $\overline{\Delta}$. The interior angle of $\overline{\Delta}$ at $\overline{p}$ is called the *comparison angle* between q and r at p and it is denoted by $\overline{\angle}_p(q,r)$. Now assume additionally that (X,d) is a geodesic space. Let $\gamma : [0,1] \to X$ and $\eta : [0,1] \to X$ be two geodesics starting at p i.e. $\gamma(0) = \eta(0) = p$. The *Alexandrov angle* between γ and η is the number $\angle_p(\gamma, \eta) \in [0, \pi]$ defined as

$$\angle_p(\gamma, \eta) := \limsup_{t,t' \to 0} \overline{\angle}_p(\gamma(t), \eta(t')). \tag{2.16}$$

Alternatively the angle between two geodesics can be expressed in terms of the metric

$$\cos(\overline{\angle}_p(\gamma(t), \eta(t'))) = \frac{1}{tt'}(t^2 + t'^2 - d(\gamma(t), \eta(t'))^2). \tag{2.17}$$

Note that in the Euclidean space $\mathbb{E}^n$ the Alexandrov angle coincides with the usual Euclidean angle. An alternative definition is given by Alexandrov [3], [4]. The *strong upper angle* between two geodesics $\gamma : [0,1] \to X$ and $\eta : [0,1] \to X$ starting from the same point p is the number $\angle_p(\gamma, \eta) \in [0, \pi]$ such that

$$\angle_p(\gamma, \eta) := \lim_{t \to 0} \sup_{t' \in (0,1]} \overline{\angle}_p(\gamma(t), \eta(t')). \tag{2.18}$$

It was shown by Alexandrov that (2.18) is equivalent to (2.16). By Proposition 1.14 [32] the Alexandrov angle satisfies the triangle inequality

$$\angle_p(\gamma_1, \gamma_2) \leqslant \angle_p(\gamma_1, \gamma_3) + \angle_p(\gamma_3, \gamma_2) \tag{2.19}$$

whenever $\gamma_1, \gamma_2, \gamma_3 : [0,1] \to X$ are geodesics starting from the same point p.

2.2.3 Length of a Curve

A *curve* or a *path* in X is a continuous map c from a compact interval $[a,b] \subset \mathbb{R}$ to X. The length $l(c)$ of a curve $c : [a,b] \to X$ is defined as

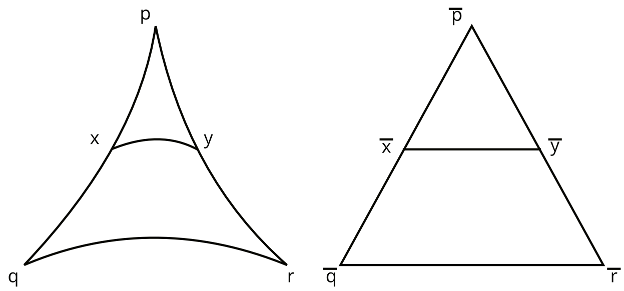
$$l(c) := \sup_{a = t_0 \leqslant t_1 \leqslant \ldots \leqslant t_n = b} \sum_{k=0}^{n-1} d(c(t_i), c(t_{i+1})) \tag{2.20}$$

where the supremum is taken over all possible partitions of $[a,b]$ (no bound on n). The length $l(c)$ of a curve c is nonnegative, possibly unbounded. The curve c is said to be *rectifiable* if its length is finite. If (X,d) is a metric space, not necessarily geodesic, then the set X can be equipped with a metric d' called the *induced length metric* defined as

$$d'(x,y) := \inf\{l(c) | c : [0,1] \to X, c(0) = x, c(1) = y\}. \tag{2.21}$$

If d coincides with d' we say (X,d) is a *length space*. Clearly any geodesic metric space is a length space. In general the converse is not true (see [43]).

Figure 2.1: Geodesic triangle (left) and its comparison triangle (right).



2.3. CAT(κ) Spaces

2.3.1 The CAT(κ) inequality

A CAT(κ) space is a metric space of curvature bounded above by some real number κ. It can be viewed as a generalization of a CAT(0) space. In this thesis we focus our study on CAT(κ) spaces for which $\kappa \leqslant 0$. To make the definition of a CAT(κ) space precise we need the notion of a *model space*. Following standard literature [41], [54], [32] for a given real number κ and a fixed integer $n \geqslant 2$ denote by M_κ^n the following metric spaces

- if $\kappa = 0$ then $M_\kappa^n = \mathbb{E}^n$;
- if $\kappa > 0$ then M_κ^n is obtained from the sphere $\mathbb{S}^n$ by multiplying the distance function by the factor $1/\sqrt{\kappa}$;
- if $\kappa < 0$ then M_κ^n is obtained from the hyperbolic space $\mathbb{H}^n$ by multiplying the distance function by the factor $1/\sqrt{-\kappa}$.

By virtue of [32, Proposition 2.11] M_κ^n is a uniquely geodesic metric space i.e. any two points are connected by a unique geodesic whenever $\kappa \leqslant 0$. These metric spaces, including the case $\kappa > 0$, are known as the model spaces. A model space is a length space and therefore its intrinsic metric is given by its induced length metric $d'(\cdot,\cdot)$. Let (X,d) be a geodesic metric space and Δ a geodesic triangle in X determined by the points $p,q,r \in X$. A geodesic triangle $\overline{\Delta}$ determined by points $\overline{p},\overline{q},\overline{r} \in M_\kappa^n$ is a comparison triangle for Δ whenever $d(p,q) = d'(\overline{p},\overline{q}), d(q,r) = d'(\overline{q},\overline{r})$ and $d(p,r) = d'(\overline{p},\overline{r})$. A point $\overline{x} \in [\overline{p},\overline{q}]$ is said to be a comparison point for $x \in [p,q]$ whenever $d(p,x) = d'(\overline{p},\overline{q})$, similarly for the other two geodesic segments of the triangle. Then, Δ is said to satisfy CAT(κ) inequality if for all $x,y \in \Delta$ and all comparison points $\overline{x},\overline{y} \in \overline{\Delta}$,

$$d(x,y) \leqslant d'(\overline{x},\overline{y}). \tag{2.22}$$

The metric space (X,d) is then called a CAT(κ) space if it is a geodesic metric space and all its geodesic triangles satisfy CAT(κ) inequality. In other words triangles in a CAT(κ) space are at least as thin as the geodesic triangles in the model space M_κ^n (see 2.1 for the case of a CAT(0) space).

Theorem 2.6 ([32, Theorem 1.12]). *The followings are true:*

1. *If X is a CAT(κ) space, then it is a CAT(κ') space for every $\kappa' \geqslant \kappa$.*

2. *If X is a CAT(κ') space for every $\kappa' > \kappa$, then it is a CAT(κ) space.*

2.3.2 Examples of CAT(κ) spaces

1. Any pre-Hilbert space, a not necessarily complete *inner product space*, (see [116], p. 39-40), is a CAT(0) space. This follows immediately from the comparison triangle definition. Moreover by Proposition II.1.14 [32] if a vector space is CAT(κ) for some $\kappa \in \mathbb{R}$ then it has to be a pre-Hilbert space. Consequently there does not exist a vector space which is not pre-Hilbert and it is of nonpositive curvature.

2. A convex subset of a Euclidean space $\mathbb{E}^n$ is a CAT(0) space whenever it is endowed with the induced metric.

3. The complement of the planar set $\{(x, y) \in \mathbb{R}^2 | x > 0, y > 0\}$ when equipped with its induced length metric is a CAT(0) space. In general the complement of any polygon in M^n_κ is a CAT(κ) space.

4. Phylogenetic tree space is a CAT(0) space (see [26]). In general an M_κ polyhedral complex, with Shapes(K) finite and satisfying the *link condition* is a CAT(κ) space (for details see [32], p. 206).

5. If (X_1, d_1) and (X_2, d_2) are CAT(0) spaces then so is the metric space $(X_1 \times X_2, d)$ where $d(\cdot, \cdot)^2 := d_1(\cdot, \cdot)^2 + d_2(\cdot, \cdot)^2$.

6. A Riemannian manifold with sectional curvature $\leqslant \kappa$ is a CAT(κ) space (see Theorem II 1.A.6, [32]).

7. An R-tree is a CAT(κ) space for any κ. An R-tree is a metric space (X, d) such that:

   - there is a unique geodesic between any pair of points;

   - if $[y, x] \cap [x, z] = \{x\}$ then $[y, z] = [y, x] \cup [x, z]$.

   In fact any geodesic triangle in an R-tree space is degenerate in the sense that the angle at any of the vertices is either 0 or π. And it is π whenever that vertex lies in the geodesic between the other two vertices. For more on R-tree spaces refer to [71] and references therein.

For an extensive treatment of these spaces and the important role they play in mathematics one could refer to Bridson and Haefliger [32] or the book by Burago et al. [41].

2.4. Hadamard Spaces

2.4.1 Definition of Hadamard space

A complete CAT(0) space is called a *Hadamard space* and we denote it by (H, d). In certain circles these spaces are also known by the acronym NPC spaces (complete metric spaces of nonpositive curvature, see [15, Ballman], [109, Sturm]). The most trivial example of a Hadamard space is a Euclidean space $\mathbb{E}^n$ and in general the only Banach space which is Hadamard is the Hilbert space. R-trees are another example of a Hadamard space and so are phylogenetic tree spaces. Some of the most important Hadamard spaces are *Hadamard manifolds*. A Hadamard manifold is a Riemannian manifold of nonnegative sectional curvature (see [48], [102] for fundamentals in differential geometry). In particular the model spaces M_κ^n are Hadamard spaces whenever $\kappa \leqslant 0$. Another interesting example of a Hadamard space is the so called *Hilbert ball*. If $(\mathcal{H}, \|\cdot\|)$ is a complex Hilbert space equipped with its canonical inner product $\langle\cdot,\cdot\rangle$ and $\mathbb{B} := \{x \in \mathcal{H} | \|x\| < 1\}$ is the unit open ball then the metric space $(\mathbb{B}, d)$ where $d(x, y) := \operatorname{arc\,tanh}\sqrt{1 - \sigma(x, y)}$ and $\sigma(x, y) := (1 - \|x\|^2)(1 - \|y\|^2)/(1 - \langle x, y\rangle)$ is a Hadamard space (for details see Example 1.2.13 [43]). Define the mapping

$$W(x, y, t) := tx \oplus (1 - t)y, \quad \forall x, y \in H,\ t \in [0, 1] \tag{2.23}$$

where $W(x, y, t) \in H$ is the point on the geodesic segment $[x, y]$ connecting x with y such that $d(x, W(x, y, t)) = (1 - t)d(x, y)$ and $d(y, W(x, y, t)) = td(x, y)$. It can be shown that the CAT(0) inequality (2.22) is equivalent to

$$d(W(x, y, t), z)^2 \leqslant (1 - t)d(x, z)^2 + td(y, z)^2 - t(1 - t)d(x, y)^2,\ \ \forall z \in H. \tag{2.24}$$

Note that inequality (2.24) can alternatively be obtained by a celebrated result of Bruhat and Tits [40] which states that in a Hadamard space (H, d) for any two points $x, y \in H$ there exists a point $m' \in H$ such that

$$d(m', z)^2 \leqslant \frac{1}{2}d(x, z)^2 + \frac{1}{2}d(y, z)^2 - \frac{1}{4}d(x, y)^2,\ \ \forall z \in H. \tag{2.25}$$

Elementary calculations show that $m' = W(x, y, 1/2)$ is just the metric midpoint of x, y. An iterative application of this property (2.25) and completeness of H yields again (2.24). Therefore in view of strongly convex metric spaces, a Hadamard space is strongly convex. Consequently from Remark 1 it follows that a Hadamard space is a convex metric space where their convex structure W is given by (2.23).

2.4.2 Some fundamental characterization results

We end this chapter with the following results which tie together complete convex metric spaces admitting a strongly convex structure with Hadamard spaces (compare with [43, Theorem 1.3.3]).

Theorem 2.7. *A complete metric space (X, d) is a Hadamard space if and only if (X, d) admits a strongly convex structure W.*

Proof. Let (X, d) be a complete metric space that is also Hadamard then (2.24) holds true. By definition (2.14) it follows that (X, d) is strongly convex. Now suppose that (X, d) is a complete metric space admitting a strongly convex structure W. By Remark 1 it follows that (X, d) is a convex metric space. Then Proposition 2.5 implies that (X, d) is a geodesic space. Let $\Delta \subseteq X$ be a geodesic triangle with vertices $p, q, r \in X$. Denote by $\overline{\Delta}$ its comparison triangle in $\mathbb{E}^2$ with vertices $\overline{p}, \overline{q}, \overline{r}$. Definition (2.14) implies

$$d(p, W(q, r, t))^2 \leqslant (1-t)\|\overline{p} - \overline{q}\|^2 + td\|\overline{p} - \overline{r}\|^2 - t(1-t)\|\overline{q} - \overline{r}\|^2.$$

Calculations in $\mathbb{E}^2$ show $d(p, W(q, r, t)) \leqslant \|\overline{p} - (1-t)\overline{q} - t\overline{r}\|$. On the other hand we have $(1-t)\overline{q} + t\overline{r} = \overline{W}(q, r, t)$ hence $d(p, W(q, r, t)) \leqslant \|\overline{p} - \overline{W}(q, r, t)\|$. This confirms the CAT(0) inequality. Since the geodesic triangle Δ is arbitrary then (X, d) must be a CAT(0) space. Assumption (X, d) is complete implies (X, d) is a Hadamard space. □

Theorem 2.8. *Any Hadamard space is a Busemann space and the inclusion is strict.*

Proof. The proof is based on a fundamental characterisation inequality due to Berg and Nikolaev [24, Theorem 1, Corollary 3] which states that a metric space (X, d) is CAT(0) if and only if for any four points $x_0, y_0, x_1, y_1 \in X$ it holds

$$d(x_0, y_1)^2 + d(y_0, x_1)^2 - d(x_0, x_1)^2 - d(y_0, y_1)^2 \leqslant 2d(x_0, y_0)d(x_1, y_1). \tag{2.26}$$

Now let $x_t := W(x_0, x_1, t)$ and $y_t := W(y_0, y_1, t)$ for some $t \in [0, 1]$. Then by strong convexity

$$d(x_t, y_t)^2 \leqslant t^2 d(x_0, y_0)^2 + (1-t)^2 d(x_1, y_1)^2 + t(1-t)(d(x_0, y_1)^2 + d(y_0, x_1)^2 - d(x_0, x_1)^2 - d(y_0, y_1)^2).$$

From inequality (2.26) it follows $d(x_t, y_t)^2 \leqslant (td(x_0, y_0) + (1-t)d(x_1, y_1))^2$ hence

$$d(W(x_0, x_1, t), W(y_0, y_1, t)) \leqslant td(x_0, y_0) + (1-t)d(x_1, y_1).$$

This proves that a Hadamard space is a Busemann space. Note that an example of a Busemann space that is not Hadamard is any strictly convex Banach space which is not a Hilbert space. Take for instance L^p space for $1 < p < 2$. □

By Theorem 2.7 and Propositions 2.3, 2.4 we obtain the following corollaries.

Corollary 2.9. *In a Hadamard space midpoints are unique. In particular any two points $x, y \in H$ determine exactly one geodesic segment with end points x and y.*

Corollary 2.10 ([43, Theorem 2.1.12])**.** *The metric projection $P_S x$ is nonempty and unique for every $x \in H$ whenever $S \subseteq H$ is a closed convex set. Moreover for any $z \in S \setminus P_S x$ the Alexandrov angle satisfies $\angle_{P_S x}([x, P_S x], [P_S x, z]) \geqslant \pi/2$.*

Proof. First part follows immediately from Proposition 2.4 since by Theorem 2.7 a Hadamard space is strongly convex. Now note that for any $x' \in [x, P_S x], z' \in [P_S x, z]$ inequality (2.15) holds with x replaced by x' and z replaced by z'. By representation (2.17) it follows $\cos(\angle_{P_S x}([x, P_S x], [P_S x, z]) \leqslant 0$ or equivalently $\angle_{P_S x}([x, P_S x], [P_S x, z]) \geqslant \pi/2$. □

At this point it is noteworthy to mention one of the most intriguing problems in convex analysis regarding *Chebyshev sets*. Recall that a set S in a metric space (X, d) is called a Chebyshev set whenever it admits unique projections for every point in the space, i.e. for every $x \in X$ there exists a unique point $x^* \in S$ such that $d(x, x^*) = d(x, S)$. It is known that in a Euclidean space a set is Chebyshev if and only if it is closed and convex. However in general metric spaces a Chebyshev set need not be a convex set, for example in geometric spaces of nonpositive curvature. In particular there are nonconvex Chebyshev sets in the real hyperbolic plane $\mathbb{H}^2$ (see [28]). However the problem whether arbitrary Chebyshev sets in a Hilbert space are convex is unsolved (since 1987).

CHAPTER 3

Weak Topology in Hadamard Spaces

3.1. Identification of Weak Topology

3.1.1 Weak convergence

In 1976 Lim [80] introduced the concept of Δ-convergence in a general metric space (X, d). A sequence $(x_n)_{n\in\mathbb{N}}$ in X is said to Δ-converge to x, written $x_n \overset{\Delta}{\to} x$, if

$$\limsup_k d(x_{n_k}, x) \leqslant \limsup_k d(x_{n_k}, y) \tag{3.1}$$

for every subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ and for every $y \in X$. This concept was adopted later by Kirk and Panyanak [72] in the setting of CAT(0) spaces. In Hadamard spaces Δ-convergence has a natural interpretation in terms of the so called asymptotic centers. Let $(x_n)_{n\in\mathbb{N}} \subset H$ be a bounded sequence and $y \in H$. Consider the following function

$$r((x_n)_{n\in\mathbb{N}}, y) := \limsup_n d(x_n, y)^2 \tag{3.2}$$

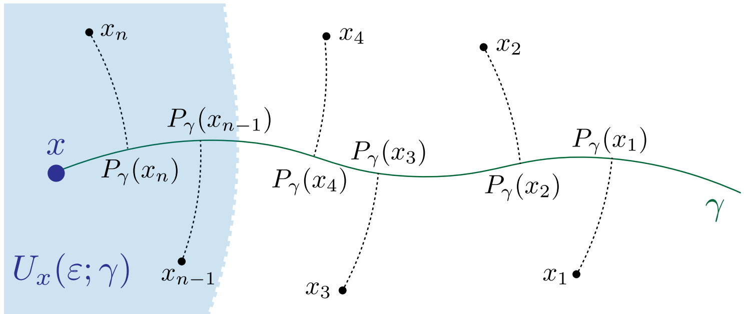
and let

$$r((x_n)_{n\in\mathbb{N}}) := \inf_{y\in H} r((x_n)_{n\in\mathbb{N}}, y) \tag{3.3}$$

denote the asymptotic radius of $(x_n)_{n\in\mathbb{N}}$. Since $d(x_n, \cdot)^2$ is a strongly convex function on H for each n then so is $r((x_n)_{n\in\mathbb{N}}, \cdot)$ as the limit superior of a sequence of strongly convex functions. Therefore there exists a unique minimizer $x \in H$ of $r((x_n)_{n\in\mathbb{N}}, \cdot)$ known as the asymptotic center of the sequence $(x_n)_{n\in\mathbb{N}}$. We say a bounded sequence $x_n \overset{\Delta}{\to} x$ if and only if x is the asymptotic center of every subsequence of $(x_n)_{n\in\mathbb{N}}$ see. e.g., Bačák's work [12, 13, 43]. Δ-convergence in Hadamard spaces is often referred to as the weak convergence, however we reserve this term for later.

It is worth noting that Δ-convergence in CAT(0) spaces shares many properties with the usual notion of weak convergence in Banach spaces. As already noted by Kirk and Panyanak, Δ-convergence in CAT(0) inherits the Opial's and Kadec-Klee properties among

Figure 3.1: Convergence of projections $P_\gamma x_n$ to x along a geodesic γ starting at x. The blue region is part of the elementary set $U_x(\varepsilon;\gamma)$ of our weak topology, which might extend infinitely far to the left of x.



others, see [72]. A form of Banach-Saks property is also satisfied, see [43, Bačak]. In the case of a Hilbert space, the only Banach space which is a CAT(0) space, the notion of Δ-convergence coincides with the usual weak convergence. In Banach spaces, an important consequence of the Banach-Steinhaus theorem is that a weakly convergent sequence in a Banach space must be bounded. However, in a Hadamard space we need to put the boundedness assumption in the definition of Δ-convergence so that the expression in (3.2) is a finite quantity (see Bačak [43]).

In our work we define weak convergence based on a notion introduced by Jost [61][1].

Definition 3.1. *A sequence $(x_n)_{n\in\mathbb{N}} \subseteq H$ converges weakly to an element $x \in H$, and we write $x_n \overset{w}{\to} x$, if and only if $P_\gamma x_n \to x$ as $n \to +\infty$ along any geodesic $\gamma : [0,1] \to H$ starting at x, where P_γ denotes the projection onto γ (see Figure 3.1).*

Note that this notion of weak convergence does not require that the sequence be bounded. Indeed consider a simplicial tree of countably many rays of finite but increasing (unbounded) length all meeting at a common vertex. If x_n is the tip of the geodesic ray γ_n and x is the common vertex then $x_n \overset{w}{\to} x$. By construction $(x_n)_{n\in\mathbb{N}}$ is unbounded.

An open question related to Δ-convergence in a Hadamard space (H,d) has been the identification of a topology τ_Δ on H such that for a given bounded sequence $(x_n)_{n\in\mathbb{N}} \subset H$ and $x \in H$ we have $x_n \overset{\Delta}{\to} x$ if and only if $x_n \overset{\tau_\Delta}{\to} x$. In view of Definition 3.1 it suffices to identify a weak topology τ_w which generates the weak convergence, since τ_Δ would then be the restriction of τ_w on bounded sets. Motivated by a suggestion of Bačak we construct a topology τ_w, which is weaker than the usual metric topology. It holds that the weak convergence yields convergence in τ_w. Moreover if the underlying Hadamard space is *weakly proper* then convergence in τ_w implies weak convergence and thus τ_w would be the correct topology for our weak convergence. Although the construction of τ_w turns out to be rather simple, its construction has important consequences. In fact, we will show how this topology can be used in order to obtain certain compactness results

[1] This notion was also considered by Espínola and Fernández-León [50] and in a slightly different form by Sosov [105].

that are similar to those known from the linear setting of Banach spaces. In particular, we show that in a separable Hadamard space weak compactness and weak sequential compactness are equivalent on bounded sets. If additionally the space is weakly proper and satisfies a certain regularity condition then this equivalence holds also for unbounded sets. Furthermore a bounded closed convex set C in a separable Hadamard space is weakly compact independent of whether the Hadamard space is weakly proper or not. We also suggest a notion of dual spaces corresponding to this topology which in the case of a Hilbert space coincides with the usual dual space. Next we introduce the space of geodesic segments and study a corresponding notion of weak convergence. We show that the space of geodesic segments and the underlying Hadamard space are homeomorphic when equipped with their respective weak topologies. Later we establish existence results for the so-called steepest descent direction in a Hadamard space. Finally, we offer a comparison of our notion to previous notions of weak topologies in Hadamard spaces. In particular, we will compare to the work of Monod [90], which offers a notion of weak topology that is weaker than ours, at least in the case of weakly proper spaces, but which does not yield convergence along geodesics and is thus too weak for our purposes. Likewise, we will compare the work of Kakavandi [64], who offers a topology that is stronger than ours but for which it is unknown whether compactness results hold in general.

3.1.2 Construction of open sets

In order to construct a desired weak topology τ_w, we build on a suggestion by Bačak [13]. A constant speed *geodesic* $\gamma : [0,1] \to H$ is a curve that satisfies $d(\gamma(s), \gamma(t)) = |s-t| d(\gamma(0), \gamma(1))$ for all $s, t \in [0,1]$. For a given $x_0 \in H$ we define $\Gamma_{x_0}(H)$ to be the set of all constant speed geodesics $\gamma : [0,1] \to H$ such that $\gamma(0) = x_0$. A set $U \subset H$ is *weakly open* if it satisfies the following property: for every $x_0 \in U$ there exists some $\varepsilon > 0$ and a finite family of geodesics $\gamma_1, \gamma_2, ..., \gamma_n \in \Gamma_{x_0}(H)$ such that the set

$$U_{x_0}(\varepsilon; \gamma_1, ..., \gamma_n) := \{x \in H : d(x_0, P_{\gamma_i} x) < \varepsilon \ \ \forall i = 1, 2, ..., n\} \tag{3.4}$$

is contained in U.

Proposition 3.2. *The collection of weakly open sets U together with the empty set $\emptyset$ define a topology τ_w on H and we call it the weak topology on H.*

Proof. It is clear that $H \in \tau_w$ since if $x \in H$ then for any finite family of geodesic segments $\gamma_1, ..., \gamma_n \in \Gamma_x(H)$ the set $U_x(\varepsilon; \gamma_1, ..., \gamma_n) \subseteq H$ for every $\varepsilon > 0$. The empty set $\emptyset$ is in τ_w by definition.

Moreover for any collection $\{U_i\}_{i \in I}$ where I is some index set and U_i is weakly open for all $i \in I$ its union is weakly open. To see this let $x \in \bigcup_{i \in I} U_i$ then $x \in U_j$ for some $j \in I$. Since U_j is weakly open then there exist $\varepsilon > 0$ and $\gamma_1, ..., \gamma_n \in \Gamma_x(H)$ such that $U_x(\varepsilon; \gamma_1, ..., \gamma_n) \subseteq U_j \subseteq \bigcup_{i \in I} U_i$. Hence $\bigcup_{i \in I} U_i$ is weakly open.

Now let U_1 and U_2 be two weakly open sets and let $x \in U_1 \cap U_2$. Then there exist $\varepsilon_1, \varepsilon_2 > 0$ and geodesic segments $\gamma_1, ..., \gamma_n, \eta_1, ..., \eta_m \in \Gamma_x(H)$ such that $U_x(\varepsilon_1; \gamma_1, ..., \gamma_n) \subseteq U_1$ and

$U_x(\varepsilon_2; \eta_1, ..., \eta_m) \subseteq U_2$. Let $\varepsilon := \min\{\varepsilon_1, \varepsilon_2\}$ and consider the set $U_x(\varepsilon; \gamma_1, ..., \gamma_n, \eta_1, ..., \eta_m)$. By construction it follows that

$$U_x(\varepsilon; \gamma_1, ..., \gamma_n, \eta_1, ..., \eta_m) \subseteq U_x(\varepsilon_1; \gamma_1, ..., \gamma_n) \cap U_x(\varepsilon_2; \eta_1, ..., \eta_m) \subseteq U_1 \cap U_2$$

Therefore $U_1 \cap U_2$ is weakly open. □

To justify the name weakly open we need to show that any set U is open in the usual metric topology. The next lemma solves this issue:

Lemma 3.3. *Let (H, d) be a Hadmard space. Then the sets $U_{x_0}(\varepsilon; \gamma)$ are open in the metric topology for every geodesic $\gamma : [0, 1] \to H$ whose image contains x_0.*

Proof. A well known inequality due to Reshetnyak [98] (see also [43, Theorem 1.3.3.]) states that for any four points $x, y, u, v \in X$ where X is a CAT(0) space it holds that

$$d(x, u)^2 + d(y, v)^2 \leqslant d(x, y)^2 + d(u, v)^2 + d(x, v)^2 + d(y, u)^2. \tag{3.5}$$

If $C \subseteq H$ is a closed convex set and P_C denotes the projection operator onto C, then the inequality

$$d(x, z)^2 \geqslant d(x, P_C x)^2 + d(P_C x, z)^2, \quad \forall z \in C \tag{3.6}$$

holds true (see [43, Theorem 2.1.12]). Let $x, y \in H$. Applying (3.6) twice with $z = P_C x$ and $z = P_C y$ yields

$$2d(P_C x, P_C y)^2 + d(x, P_C x)^2 + d(y, P_C y)^2 \leqslant d(x, P_C y)^2 + d(y, P_C x)^2. \tag{3.7}$$

Applying (3.5) to the right side of the last inequality gives

$$d(x, P_C y)^2 + d(y, P_C x)^2 \leqslant d(x, y)^2 + d(P_C y, P_C x)^2 + d(x, P_C x)^2 + d(y, P_C y)^2. \tag{3.8}$$

Then (3.7) and (3.8) imply $d(P_C x, P_C y) \leqslant d(x, y)$. Therefore, in a Hadamard space P_C is a *nonexpansive operator*. In particular, P_γ is nonexpansive since every geodesic segment γ is a closed convex set. This implies

$$d(P_\gamma x, P_\gamma y) \leqslant d(x, y) \;\; \forall x, y \in H. \tag{3.9}$$

Now let $x \in U_{x_0}(\varepsilon; \gamma)$. Then there exists some $s < \varepsilon$ such that $d(x_0, P_\gamma x) = s$. We claim that the open geodesic ball $B(x, \varepsilon - s) := \{y \in H : d(x, y) < \varepsilon - s\}$ is contained in $U_{x_0}(\varepsilon; \gamma)$. Let $y \in B(x, \varepsilon - s)$ then

$$d(x_0, P_\gamma y) \leqslant d(x_0, P_\gamma x) + d(P_\gamma x, P_\gamma y) \leqslant d(x_0, P_\gamma x) + d(x, y),$$

where the first inequality follows from the triangle inequality and the second one from (3.9). Therefore

$$d(x_0, P_\gamma y) \leqslant d(x_0, P_\gamma x) + d(x, y) < s + \varepsilon - s = \varepsilon$$

implies $B(x, \varepsilon - s) \subset U_{x_0}(\varepsilon; \gamma)$. □

We say a Hadamard space is *weakly proper* if for any elementary set $U_x(\varepsilon;\gamma)$ there exists a weakly open set $V \in \tau_w$ containing x such that $V \subseteq U_x(\varepsilon;\gamma)$ and this holds for all $x \in H, \gamma \in \Gamma_x(H)$ and $\varepsilon > 0$. Essentially this property requires that elementary sets $U_x(\varepsilon;\gamma)$ to have nonempty interior in τ_w. Note that any Hilbert space is weakly proper. In fact it is even more, the sets $U_x(\varepsilon;\gamma)$ are open in τ_w. To see this take any set $U_x(\varepsilon;\gamma)$ and let $y \in U_x(\varepsilon;\gamma)$. Then there exists a geodesic segment η emanating from y entirely included in $U_x(\varepsilon;\gamma)$ such that it lies in a common plane with γ and it is parallel to it. Then by usual rules of Euclidean geometry it follows that $U_y(\delta;\eta) \subseteq U_x(\varepsilon;\gamma)$ for small enough $\delta > 0$. It is of interest to know whether weakly properness is a universal property for Hadamard spaces.

Theorem 3.4. *Consider a sequence $(x_n)_{n\in\mathbb{N}} \subset H$, and let $x \in H$. If $x_n \xrightarrow{w} x$ then $x_n \xrightarrow{\tau_w} x$. Moreover provided that the Hadamard space is weakly proper then $x_n \xrightarrow{\tau_w} x$ implies $x_n \xrightarrow{w} x$.*

Proof. Let $x_n \xrightarrow{w} x$. Then for every $\gamma \in \Gamma_x(H)$ we have that $\lim_{n\to\infty} d(x, P_\gamma x_n) = 0$, or equivalently $x_n \in U_x(\varepsilon;\gamma)$ for all sufficiently large n. Let $U \in \tau_w$ containing x. Then there exist $\gamma_1, ..., \gamma_n \in \Gamma_x(H)$ and $\varepsilon > 0$ such that $U_x(\varepsilon;\gamma_1,...,\gamma_n) \subseteq U$. Since $x_n \in U_x(\varepsilon;\gamma_1,...,\gamma_n)$ for all sufficiently large n then $x_n \in U$ for all sufficiently large n.

Let H be weakly proper. Suppose that $x_n \xrightarrow{\tau_w} x$ but $x_n \not\xrightarrow{w} x$. Then there exists $\gamma \in \Gamma_x(H)$ such that $\lim_n P_\gamma x_n \neq x$ i.e. there is $\varepsilon > 0$ such that $x_n \notin U_x(\varepsilon;\gamma)$ for infinitely many n. Weakly properness implies that there is an open set $V \in \tau_w$ containing x such that $V \subseteq U_x(\varepsilon;\gamma)$. Therefore $x_n \notin V$ infinitely many n. However this contradicts $x_n \xrightarrow{\tau_w} x$.

□

Lemma 3.5. *A weakly proper Hadamard space is Hausdorff with respect to topology τ_w.*

Proof. Let $x, y \in H$ be two distinct points. Since H is Hadamard, there exist unique geodesics $\gamma, \tilde{\gamma} : [0,1] \to H$ connecting x with y such that $\gamma(0) = x$, $\gamma(1) = y$ and $\tilde{\gamma}(0) = y$, $\tilde{\gamma}(1) = x$. Let $l_\gamma := d(x,y)$, and let $\varepsilon \in (0, l_\gamma)$. Note that $l_\gamma = l_{\tilde{\gamma}}$ and $P_\gamma z = P_{\tilde{\gamma}} z$ for all $z \in H$. Define the sets

$$U_x(\varepsilon;\gamma) := \{z \in H : d(x, P_\gamma z) < \varepsilon\}$$

and

$$U_y(l_\gamma - \varepsilon;\tilde{\gamma}) := \{z \in H : d(y, P_{\tilde{\gamma}} z) < l_\gamma - \varepsilon\}.$$

Suppose there is some $z_0 \in U_x(\varepsilon;\gamma) \cap U_y(l_\gamma - \varepsilon;\tilde{\gamma})$ then $d(x, P_\gamma z_0) < \varepsilon$ and $d(y, P_{\tilde{\gamma}} z_0) < l_\gamma - \varepsilon$ would imply

$$l_\gamma = d(x,y) \leqslant d(x, P_\gamma z_0) + d(y, P_\gamma z_0) = d(x, P_\gamma z_0) + d(y, P_{\tilde{\gamma}} z_0) < l_\gamma$$

which is impossible. Therefore $U_x(\varepsilon;\gamma) \cap U_y(l_\gamma - \varepsilon;\tilde{\gamma}) = \emptyset$. Since (H,d) is weakly proper there are $V, W \in \tau_w$ such that $x \in V \subseteq U_x(\varepsilon;\gamma)$ and $y \in W \subseteq U_y(l_\gamma - \varepsilon;\tilde{\gamma})$. Then $V \cap W \subseteq U_x(\varepsilon;\gamma) \cap U_y(l_\gamma - \varepsilon;\tilde{\gamma}) = \emptyset$. □

3.1.3 Convex sets and compactness

We say a set $C \subseteq H$ is weakly closed if it is closed with respect to τ_w. It is evident that a weakly closed set is closed.

Theorem 3.6. *Let (H, d) be a weakly proper space. If $C \subseteq H$ is a closed convex set then C is weakly closed.*

Proof. Let $y \in H \setminus C$. Then $P_C y$ exists and is unique. Let $\gamma : [0, 1] \to H$ be the geodesic connecting y with $P_C y$ such that $\gamma(0) = y$ and $\gamma(1) = P_C y$. For $\varepsilon \in (0, l(\gamma))$, where $l(\gamma) := d(y, P_C y)$ is the length of the geodesic γ, consider the weakly open set $U_y(\varepsilon; \gamma)$. Since (H, d) is weakly proper space it suffices to prove that $U_y(\varepsilon; \gamma) \cap C = \emptyset$. Let $x \in C$, and let $z = P_\gamma(x)$ be the projection of x to γ. Since $z \in \gamma$, by Corollary 2.10 we have that $P_C z = P_C y$. Since both C and γ are strongly closed and convex, we have the following quadratic inequalities

$$d(z, x)^2 \geqslant d(x, P_C z)^2 + d(P_C z, z)^2$$

and

$$d(x, P_C y)^2 \geqslant d(x, P_\gamma x)^2 + d(P_\gamma x, P_C y)^2.$$

Using that $z = P_\gamma x$ and $P_C z = P_C y$ then implies that $P_\gamma x = P_C y$. In particular from $d(y, P_\gamma x) = d(y, P_C y) > \varepsilon$ for all $x \in C$ it follows that $U_y(\varepsilon; \gamma) \cap C = \emptyset$. □

Theorem 3.7 (Mazur's Lemma for Hadamard spaces)**.** *Let (H, d) be weakly proper and $(x_n)_{n \in \mathbb{N}} \subseteq H$ a sequence such that $x_n \overset{w}{\to} x$ for some $x \in H$. Then there exists some function $N : \mathbb{N} \to \mathbb{N}$ and a sequence $(y_n) \subseteq H$ such that $y_n \in \operatorname{co}(\{x_1, x_2, ..., x_{N(n)}\})$ for all $n \in \mathbb{N}$ and $y_n \to x$.*

Proof. By virtue of Theorem 3.4 convergence $x_n \overset{w}{\to} x$ implies $x_n \overset{\tau_w}{\to} x$, therefore $x \in \operatorname{wcl}\{x_1, x_2, ...\}$, where $\operatorname{wcl}\{x_1, x_2, ...\}$ denotes the weak closure of $\{x_1, x_2, ...\}$. Moreover, $\{x_1, x_2, ...\} \subseteq \operatorname{co}(\{x_1, x_2, ...\})$ implies that $\operatorname{wcl}\{x_1, x_2, ...\} \subseteq \operatorname{wcl}\operatorname{co}(\{x_1, x_2, ...\})$. Hence $x \in \operatorname{wcl}\operatorname{co}(\{x_1, x_2, ...\})$. The strong closure $\operatorname{cl}\operatorname{co}(\{x_1, x_2, ...\})$ of the convex set $\operatorname{co}(\{x_1, x_2, ...\})$ is a closed convex set. Indeed, if C is convex, then so is $\operatorname{cl} C$. In order to see this, let $u, v \in \operatorname{cl} C$. Consider sequences (u_n) and (v_n) in C such that $u_n \to u$ and $v_n \to v$. Let γ be the geodesic connecting u with v, and let γ_n be the geodesics connecting u_n with v_n. Then strong convexity of squared distance function in a Hadamard space implies that for each $t \in [0, 1]$ we have

$$\begin{aligned} d(\gamma(t), \gamma_n(t))^2 \leq & (1-t) d(u, \gamma_n(t))^2 + t d(v, \gamma_n(t))^2 - t(1-t) d(u, v)^2 \\ \leq & (1-t)\left((1-t) d(u, u_n)^2 + t d(u, v_n)^2 - t(1-t) d(u_n, v_n)^2\right) \\ & + t\left((1-t) d(v, u_n)^2 + t d(v, v_n)^2 - t(1-t) d(u_n, v_n)^2\right) \\ & - t(1-t) d(u, v)^2. \end{aligned}$$

Taking the limit $n \to \infty$ yields $d(\gamma(t), \gamma_n(t)) \to 0$. Hence $\operatorname{cl} C$ and in turn $\operatorname{cl}\operatorname{co}(\{x_1, x_2, ...\})$ are indeed convex. Since (H, d) is weakly proper then by Theorem 3.6 we obtain that $\operatorname{cl}\operatorname{co}(\{x_1, x_2, ...\})$ is weakly closed. It follows that

$$x \in \operatorname{wcl}\operatorname{co}(\{x_1, x_2, ...\}) \subseteq \operatorname{cl}\operatorname{co}\{x_1, x_2, ...\}.$$

Then there exists some sequence $(y_n) \subseteq \mathrm{co}(\{x_1, x_2, ...\})$ such that $y_n \to x$. Additionally, we have that $\mathrm{co}(\{x_1, x_2, ...\}) = \bigcup_{k\in\mathbb{N}} \mathrm{co}(\{x_1, x_2, ..., x_k\})$. Hence $y_n \in \mathrm{co}(\{x_1, x_2, ..., x_{k(n)}\})$ for some $k(n)$. For each n set $N(n) := k(n)$. □

A set $K \subseteq H$ is τ_w-sequentially compact if every sequence in K has a τ_w-convergent subsequence. It is clear that when (H, d) is weakly proper then τ_w-sequential convergence coincides with the weak sequential convergence. We say a set K is weakly compact (or compact in τ_w) if for any open cover in τ_w containing K there is a finite subcover which also contains K.

Lemma 3.8 ([43, Proposition 3.1.2]). *Every bounded sequence has a weakly convergent subsequence.*

Lemma 3.9 ([43, Lemma 3.2.1]). *Let $C \subseteq H$ be a closed convex set and $(x_n)_{n\in\mathbb{N}} \subset C$. If $x_n \xrightarrow{w} x$ then $x \in C$.*

Theorem 3.10. *If $K \subseteq H$ a bounded closed convex set then K is weakly sequentially compact.*

Proof. Any sequence $(x_n)_{n\in\mathbb{N}} \subset K$ is bounded. By virtue of Lemma 3.8 $(x_n)_{n\in\mathbb{N}}$ has a weakly convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$. Hence there is some $x \in H$ such that $x_{n_k} \xrightarrow{w} x$ as $k \to +\infty$. Since K is a closed convex set, Lemma 3.9 implies $x \in K$. □

Corollary 3.11. *Let $K \subseteq H$ be bounded then $\mathrm{cl}\,\mathrm{co}(K)$ is weakly sequentially compact.*

Corollary 3.12. *A bounded closed convex set is τ_w-sequentially compact.*

It is known that in a Hausdorff topological space a compact set is always a closed set [87, Theorem 2.12]. By Lemma 3.5 a weakly proper Hadamard space is Hausdorff when equipped with the weak topology τ_w. As result we have the following proposition:

Proposition 3.13. *Let (H, d) be weakly proper. Then a weakly compact set is weakly closed.*

A point $x \in H$ is said to be a *weak limit point* for the set $S \subset H$ if every open set $U \in \tau_w$ of x contains a point of S different from x. A point $x \in H$ is said to be a *weak accumulation point* of S if each open set $U \in \tau_w$ of x contains infinitely many distinct points of S. Clearly every accumulation point of S is a limit point of S.

Proposition 3.14 ([87, Lemma 5.2]). *Let (H, d) be weakly proper and $S \subseteq H$. A point $x \in H$ is a weak accumulation point of S if and only if x is a weak limit point of S.*

Proof. Let x be a weak limit point of S but not a weak accumulation point of S. Then there exists an open set $U \in \tau_w$ containing x such that U contains at most a finite number of points $\{x_1, x_2, ..., x_n\}$ of S. By Lemma 3.5 a weakly proper Hadamard space H is Hausdorff with respect to its weak topology τ_w. For each $i = 1, 2, ..., n$ we can find open sets $U_i \in \tau_w$ containing x and open sets $V_i \in \tau_w$ containing x_i such that $U_i \cap V_i = \emptyset$. Then the open set $U \cap U_1 \cap ... \cap U_n$ contains x and no other points of S distinct from x. This contradicts that x is a weak limit point of S. The reverse implication is evident by definition. □

A slight generalization of sequential compactness is the so called *Bolzano–Weierstrass property*. We say a topological space X has the Bolzano–Weierstrass property if any *infinite* subset of X has at least one accumulation point.

Theorem 3.15 ([87, Theorem 5.3])**.** *A weakly compact set in a weakly proper Hadamard space has the Bolzano–Weierstrass property.*

Proof. Let $K \subset H$ be a weakly compact set. We claim that every infinite set $S \subset K$ has a weak limit point in K. If not, then there is an infinite set $S \subset K$ such that for each $x \in S$ there is $U_x \in \tau_w$ satisfying $U_x \cap S = \{x\}$. In particular this implies that for any open set $U \in \tau_w$ containing x we have $U \cap S \neq \emptyset$. Thus S is weakly closed and consequently S is weakly compact. There are points $x_1, x_2, ..., x_n \in S$ such that $U_{x_1}, U_{x_2}, ..., U_{x_n}$ cover S implying $S = \{x_1, x_2, ..., x_n\}$. This contradicts that S is infinite set. From Proposition 3.14 it follows that S has at least one weak accumulation point. □

In what follows we make a regularity assumption on the Hadamard space (H, d).

Assumption 3.1. *Let $\{y_n\}_{n\in\mathbb{N}} \subset H$ be a dense set and $x \in H$. If $(x_k)_{k\in\mathbb{N}}$ is a sequence in H such that $\lim_k P_{[x,y_n]}x_k = x$ then $\lim_k P_\gamma x_k = x$ for any $\gamma \in \Gamma_x(H)$.*

Theorem 3.16. *Let (H, d) be a separable weakly proper Hadamard space that satisfies Assumption 3.1. Then a set $S \subseteq H$ satisfying the Bolzano–Weierstrass property is weakly sequentially compact.*

Proof. Let (H, d) be weakly proper and let $S \subseteq H$ satisfy the Bolzano–Weierstrass property. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in S. Then $(x_k)_{k\in\mathbb{N}}$ has a weak accumulation point $x \in S$, hence x is also a weak limit point. Since H is separable there exits a dense countable set $\{y_n\}$ in H. Let $\gamma_n : [0, 1] \to H$ denote the geodesic connecting x with y_n for every $n \in \mathbb{N}$. Now consider the family of sets V_n defined as

$$V_n := \bigcap_{i=1}^{n} U_x(1/n; \gamma_i) \quad \text{where} \quad U_x(1/n; \gamma_i) := \{y \in H \mid d(x, P_{\gamma_i} y) < 1/n\}. \tag{3.10}$$

Since (H, d) is weakly proper then there are $U_{n,i} \in \tau_w$ containing x such that $U_{n,i} \subseteq U_x(1/n; \gamma_i)$ for every $i = 1, 2, ..., n$ and for every $n \in \mathbb{N}$. Denote by $U_n := \bigcap_{i=1}^{n} U_{n,i}$ then $U_n \subseteq V_n$ for all $n \in \mathbb{N}$. Since U_n is a weakly open set containing x, it must contain at least one element $x_{k(n)}$. By considering a subsequence (and possibly re-numbering) we may assume that $x_n \in U_n$ for all n. In particular $x_n \in V_n$ for all n. Since the sets V_n are nested, we have that $x_m \in V_n$ for all $m \geqslant n$. This means that $\lim_m P_{\gamma_i} x_m = x$ for all $i = 1, 2, ..., n$ and hence $\lim_m P_{\gamma_n} x_m = x$ for all $n \in \mathbb{N}$. By Assumption 3.1 then $\lim_m P_\gamma x_m = x$ for all $\gamma \in \Gamma_x(H)$.

□

Remark 3.17. *Note that a bounded weakly compact set K in a weakly proper Hadamard space H is weakly sequentially compact. This is evident since any sequence in K is bounded and by Lemma 3.8 it has a weakly convergent subsequence. By Proposition 3.13 the set K is weakly closed, hence it contains the weak limits of its subsequences.*

A topological space X is a *Lindelöf space* if every open cover in X has a countable subcover. For metric spaces that are *separable* we have the following equivalence.

Lemma 3.18 ([87, Theorem 6.7])**.** *A metric space (X, d) is separable if and only if it is Lindelöf.*

Theorem 3.19. *Let (H, d) be separable. If a set in H is τ_w-sequentially compact then it is weakly compact.*

Proof. The proof proceeds by contradiction. Suppose that $K \subseteq H$ is weakly sequentially compact but not weakly compact. Then there exists some open cover $\{U_i\}_{i\in I}$ of K in τ_w that has no finite subcover. By assumption (H, d) is separable. Hence (H, d) is a Lindelöf space, i.e., every open cover (in the strong topology) has a countable subcover; see, e.g., [87, Theorem 6.7]. Since $\{U_i\}_{i\in I}$ is an open cover in the usual metric topology, there exists a countable subcover $\{U_j\}_{j\in J}$. Let $V_n := \bigcup_{j=1}^{n} U_j$. Then $W_n := H \setminus V_n$ is weakly closed for all n. Moreover, the family of sets W_n satisfies $W_{n+1} \subseteq W_n$. Because V_n cannot cover K, we have that $W_n \cap K$ is nonempty for every $n \in \mathbb{N}$. Let $x_n \in W_n \cap K$. Since K is weakly sequentially compact, and consequently τ_w-sequentially compact, the sequence (x_n) has a subsequence (x_{n_k}) that converges in τ_w to some element $x^* \in K$. Let $\mathcal{U}_w(x^*)$ denote the collection of weakly open sets containing x^*. Then for each $U \in \mathcal{U}_w(x^*)$ and for each $n \in \mathbb{N}$ there exists $m \geqslant n$ such that $U \cap W_m \neq \emptyset$, and in particular $U \cap W_n \neq \emptyset$, implying that $x^* \in \operatorname{wcl} W_n = W_n$. Since n is arbitrary, we have that $x^* \in \bigcap_n W_n$. Hence $x^* \in \bigcap_n W_n \cap K$. Therefore, $\bigcap_{n\in\mathbb{N}} W_n \cap K \neq \emptyset$, which together with $\bigcap_{n\in\mathbb{N}} W_n \cap K \subsetneq K$, yields that $K \supsetneq K \setminus (\bigcap_{n\in\mathbb{N}} W_n \cap K) = K \setminus \bigcap_{n\in\mathbb{N}}(K \setminus V_n) = \bigcup_{n\in\mathbb{N}}(K \cap V_n) = K$.

□

Remark 3.20. *Notice that the previous proof also applies to the more general setting of a topology that is weaker than the metric topology in any separable metric space.*

Corollary 3.21. *Let (H, d) be a separable weakly proper space. If $K \subseteq H$ is weakly sequentially compact then it is weakly compact.*

As a result to Theorem 3.15, Theorem 3.16 and Theorem 3.19 we get the following:

Theorem 3.22. *In a separable weakly proper Hadamard space (H, d) satisfying Assumption 3.1 the weak compactness, weak sequential compactness and Bolzano–Weierstrass property are equivalent.*

The last theorem is an important result in functional analysis which is known as the Eberlein-Šmulian Theorem (see [114, Whitley] or for a non-standard proof [19, S. Barinella and Ng. Siu-Ah]). The Eberlein-Šmulyan Theorem holds in any Banach space, separable or not, and therefore in any Hilbert space.

Theorem 3.23. *A bounded closed convex set K in a separable Hadamard space is weakly compact.*

Proof. Let $K \subseteq H$ be a bounded closed convex set. By Theorem 3.10 it follows that K is weakly sequentially compact and Corollary 3.12 implies that K is τ_w-sequentially compact. By virtue of Theorem 3.19 we obtain that K is weakly compact. □

3.1.4 Locally compact space

A topological space (X,τ) is said to be locally compact if for every $x \in X$ there exists an open set $U \in \tau$ and a compact set K such that $x \in U \subseteq K$.

Theorem 3.24 ([32, Proposition 3.7, Corollary 3.8 and Remark 3.9])**.** *The following statements are equivalent:*

1. *Hadamard space (H,d) is locally compact.*
2. *Every closed and bounded subset of (H,d) is compact.*
3. *Every bounded sequence in (H,d) has a convergent subsequence.*

Lemma 3.25. *Let (H,d) be locally compact. If $x_n \overset{w}{\to} x$ then $(x_n)_{n\in\mathbb{N}}$ is bounded.*

Proof. Let $x_n \overset{w}{\to} x$, but suppose that (x_n) is unbounded. Then we can wolog assume that $d(x_n,x) \geq R > 0$ for all $n \in \mathbb{N}$. Consider the closed geodesic ball $C := \mathbb{B}(x,R)$. Denote by $P_C x_n$ the projection of x_n onto C for every $n \in \mathbb{N}$. By assumption (H,d) is locally compact. Theorem 3.24 implies that C is compact, and in particular the boundary $\partial C := \{y \in H \,|\, d(x,y) = R\}$ is compact. Hence, $(P_C x_n)$ is a sequence in the compact set ∂C. There exits a subsequence $(P_C x_{n_k})$ converging to some element $z \in \partial C$. Let $\gamma : [0,1] \to H$ denote the geodesic segment connecting x with z. Evidently, $\gamma \subset C$. Denote by $\gamma_k : [0,1] \to H$ the geodesic segment connecting x with $P_C x_{n_k}$ for each $k \in \mathbb{N}$. Let $P_\gamma x_{n_k}$ denote the projection of x_{n_k} onto the geodesic segment γ. From the triangle inequality we obtain

$$d(P_C x_{n_k}, z) \geqslant |d(x_{n_k}, P_C x_{n_k}) - d(x_{n_k}, z)|,$$

which implies that $\lim_k |d(x_{n_k}, P_C x_{n_k}) - d(x_{n_k}, z)| = 0$. Since both C and γ are strongly closed and convex, we have the following quadratic inequalities (see, e.g., [43, Theorem 2.1.12]):

$$\begin{aligned} d(x_{n_k}, P_C x_{n_k})^2 + d(P_C x_{n_k}, P_\gamma x_{n_k})^2 &\leqslant d(x_{n_k}, P_\gamma x_{n_k})^2, \\ d(x_{n_k}, P_\gamma x_{n_k})^2 + d(P_\gamma x_{n_k}, z)^2 &\leqslant d(x_{n_k}, z)^2, \end{aligned}$$

implying that $d(x_{n_k}, P_C x_{n_k}) \leqslant d(x_{n_k}, P_\gamma x_{n_k}) \leqslant d(x_{n_k}, z)$. Therefore, we have that

$$\lim_k |d(x_{n_k}, P_\gamma x_{n_k}) - d(x_{n_k}, P_C x_{n_k})| = 0. \tag{3.11}$$

By assumption $x_n \overset{w}{\to} x$, and in particular, $x_{n_k} \overset{w}{\to} x$. Therefore, it follows that $P_\gamma x_{n_k} \to x$. Consider the geodesic segment $\eta_k : [0,1] \to H$ connecting x with x_{n_k}. Then there exists $z_k \in \eta_k$ such that $z_k \in \partial C$ for every $k \in \mathbb{N}$. Since $z_k \in \partial C$, we obtain that $d(x_{n_k}, P_C x_{n_k}) \leqslant d(x_{n_k}, z_k)$ and thus

$$d(x_{n_k}, x) = d(x_{n_k}, z_k) + d(z_k, x) \geqslant d(x_{n_k}, P_C x_{n_k}) + R, \quad \forall k \in \mathbb{N},$$

which in turn implies that $|d(x_{n_k}, x) - d(x_{n_k}, P_C x_{n_k})| \geqslant R > 0$. By the triangle inequality we again have

$$d(P_\gamma x_{n_k}, x) \geqslant |d(x_{n_k}, x) - d(x_{n_k}, P_\gamma x_{n_k})|, \quad \forall k \in \mathbb{N}.$$

Therefore, $\lim_k P_\gamma x_{n_k} = x$ implies that $\lim_k |d(x_{n_k}, x) - d(x_{n_k}, P_\gamma x_{n_k})| = 0$. On the other hand we have

$$\begin{aligned} 0 < R &\leqslant |d(x_{n_k}, x) - d(x_{n_k}, P_C x_{n_k})| \\ &\leqslant |d(x_{n_k}, x) - d(x_{n_k}, P_\gamma x_{n_k})| + |d(x_{n_k}, P_\gamma x_{n_k}) - d(x_{n_k}, P_C x_{n_k})|, \end{aligned}$$

which together with (3.11) delivers a contradiction since the right side vanishes as $k \uparrow +\infty$.

□

Theorem 3.26. *Weak and strong convergence coincide in a locally compact space.*

Proof. It is evident that whether or not the space is locally compact, if $x_n \to x$ then from the inequality $d(x, P_\gamma x_n) \leqslant d(x, x_n)$ for every $\gamma \in \Gamma_x(H)$ follows that $x_n \stackrel{w}{\to} x$. Now suppose that (H, d) is a locally compact Hadamard space, $x_n \stackrel{w}{\to} x$ but $x_n \not\to x$. Then there is $\varepsilon > 0$ and a subsequence $(x_{n_k})_{k\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}$ such that $d(x, x_{n_k}) \geqslant \varepsilon$ for all $k \in \mathbb{N}$. By Lemma 3.25 the sequence $(x_n)_{n\in\mathbb{N}}$ is bounded and in particular so is the subsequence $(x_{n_k})_{k\in\mathbb{N}}$. By Theorem 3.24 there is a subsequence $(x_{n_{k_l}})_{l\in\mathbb{N}} \subseteq (x_{n_k})_{k\in\mathbb{N}}$ converging to some $y \in H$ and in particular by the first part of this theorem $x_{n_{k_l}} \stackrel{w}{\to} y$. By uniqueness of the weak limits it follows that $y = x$. But $d(x, x_{n_{k_l}}) \geqslant \varepsilon > 0$, therefore a contradiction. □

Note that the arguments in Lemma 3.25 and Theorem 3.26 can be equally written in terms of nets (see Chapter 4 for the definition). By a fundamental result in topology that two topologies coincide when they have the same convergent nets it follows that the metric topology topologizes weak convergence in a locally compact Hadamard space.

A topological space is said to be *paracompact* if every open cover has an open *refinement* that is *locally finite*. A well known theorem of Stone [107, Theorem 1] states that every metric space is paracompact. By [106, Lemma A.1 pg.460] a connected, locally compact, paracompact space is σ-compact i.e. the entire space can be written as a countable union of compact subspaces. It is evident that a σ-compact space is a Lindelöf space. In view of Lemma 3.18 it follows that a connected, locally compact metric space is always separable. In particular a locally compact Hadamard space is separable.

Theorem 3.27. *Weak topology and strong topology coincide in a locally compact space whenever Assumption 3.1 holds.*

Proof. It is clear that a weakly open set is open. For the converse direction we use an approach similar to the proof of Theorem 3.16. Let U be an open set. Suppose that U is not weakly open. In particular there is some $x \in U$ such that for any set $V := U_x(\varepsilon; \gamma_1, ..., \gamma_n)$ we have $V \setminus U \neq \emptyset$. Since H is a locally compact, connected metric space, hence separable then by definition there is a countable set S that is dense in H. Let $S = \{y_n\}_{n\in\mathbb{N}}$ and denote by $\gamma_n : [0,1] \to H$ the geodesic connecting x with y_n for every

$n \in \mathbb{N}$. Consider the family of sets V_n defined as in (3.10). Since $V_n \setminus U \neq \emptyset$ then there exists $x_n \in V_n \setminus U$, in particular $x_n \in V_n$ for all $n \in \mathbb{N}$. Moreover by construction $x_m \in V_n$ whenever $m \geqslant n$ implying $\lim_m P_{\gamma_i} x_m = x$ for all $i = 1, 2, ..., n$ i.e. $\lim_m P_{\gamma_n} x_m = x$ for all $n \in \mathbb{N}$. Because the space satisfies Assumption 3.1 then $P_\gamma x_m = x$ for any $\gamma \in \Gamma_x(H)$. Because of locally compactness Lemma 3.25 implies that the sequence $(x_n)_{n\in\mathbb{N}}$ is bounded. By Theorem 3.24 there is a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$. By construction $x_{n_k} \notin U$ for all $k \in \mathbb{N}$ hence $x_{n_k} \to y$ for some $y \neq x$. But this means $x_{n_k} \xrightarrow{w} y$. This is impossible. □

Remark 3.28. *By Lemma 3.3 the elementary sets $U_x(\varepsilon;\gamma)$ are open in the usual metric topology. This enables us to construct another topology which is weaker than the metric topology and somewhat different from τ_w. If $\mathcal{U}$ is the collection of all finite intersections of sets of the form $U_x(\varepsilon;\gamma)$, where x and γ vary, let τ_{BW} be the smallest topology generated by $\mathcal{U}$ i.e. any $U \in \tau_{BW}$ is the union of sets from $\mathcal{U}$. Then any set $U \in \tau_{BW}$ is open in the usual metric topology. Moreover convergence τ_{BW} implies weak convergence (convergence along geodesics) for if $x_n \xrightarrow{\tau_{BW}} x$ then for any $\gamma \in \Gamma_x(H)$ and any $\varepsilon > 0$ we have $x_n \in U_x(\varepsilon;\gamma)$ for all large enough n i.e. $\lim_{n\to\infty} d(x, P_\gamma x_n) = 0$ for any $\gamma \in \Gamma_x(H)$; therefore, $x_n \xrightarrow{w} x$. In particular it holds $\tau_w \subseteq \tau_{BW}$ and they would coincide if and only if the elementary sets $U_x(\varepsilon;\gamma)$ are open in τ_w. Topology τ_{BW} enjoys some properties without any additional assumptions on the space. For example it is Hausdorff and a closed convex set is τ_{BW}-closed (note that the proofs for these claims would be identical to the ones for τ_w but without the assumption of weak properness). Moreover in a locally compact space $\tau_{BW} = \tau_S$. However at this point it is not clear whether a bounded sequence would have a convergent subsequence in τ_{BW}. As a result of this it is not clear whether Theorem 3.23 would still hold. Also we do not know for sure if a τ_{BW}-compact set would be τ_{BW}-sequentially compact. A deeper inquiry is needed to study τ_{BW} and its relationships with τ_w and the notion of weak convergence.*

3.2. The Dual Space

3.2.1 Construction of the dual and weak-* topology

For a given $x \in H$ each function $\phi_\gamma(x;\cdot) := d(x, P_\gamma \cdot)$ corresponds to an element $\gamma \in \Gamma_x(H)$. Let

$$\|\phi_\gamma - \phi_\eta\|_\infty := \sup_{y\in H\setminus\{x\}} \frac{|\phi_\gamma(x;y) - \phi_\eta(x;y)|}{d(x,y)}. \tag{3.12}$$

Let $H_x^* := \{\phi_\gamma(x;\cdot) : \gamma \in \Gamma_x(H)\}$, and let $d_*(\phi_\gamma, \phi_\eta) := \|\phi_\gamma - \phi_\eta\|_\infty$. Notice that due to the fact that $\phi_\gamma(x;y) \leq d(x,y)$, the value of $d_*(\phi_\gamma, \phi_\eta)$ is finite.

Lemma 3.29. *(H_x^*, d_*) is a metric space.*

Proof. Symmetry and the triangle inequality are evident. In order to show definiteness, suppose that $d_*(\phi_\gamma, \phi_\eta) = 0$, then this means $\phi_\gamma(x;y) = \phi_\eta(x;y)$ for all $y \in H$. In particular, for $y_1 = \gamma(1)$ and $y_2 = \eta(1)$ we obtain that $\phi_\eta(x;\gamma(1)) = d(x,\gamma(1))$

and $\phi_\gamma(x;\eta(1)) = d(x,\eta(1))$. On the other hand, we have the equation $d(x,\gamma(1)) = d(x,P_\gamma\eta(1)) + d(P_\gamma\eta(1),\gamma(1))$ and analogously the equation $d(x,\eta(1)) = d(x,P_\eta\gamma(1)) + d(P_\eta\gamma(1),\eta(1))$. These two equations imply that $d(P_\gamma\eta(1),\gamma(1)) = d(P_\eta\gamma(1),\eta(1)) = 0$. We claim that $\gamma(1) = \eta(1)$. If not then by Corollary 2.10 the Alexandrov angle $\angle_{\gamma(1)}([x,\gamma(1)],[\gamma(1),\eta(1)]) \geqslant \pi/2$ and analogously $\angle_{\eta(1)}([x,\eta(1)],[\gamma(1),\eta(1)]) \geqslant \pi/2$. On the other hand by the cosine formula for Euclidean triangles, the comparison angles $\overline{\angle}_{\gamma(1)}(x,\eta(1))$ and $\overline{\angle}_{\eta(1)}(x,\gamma(1))$ satisfy $\overline{\angle}_{\gamma(1)}(x,\eta(1)) \geqslant \angle_{\gamma(1)}([x,\gamma(1)],[\gamma(1),\eta(1)])$ and $\overline{\angle}_{\eta(1)}(x,\gamma(1)) \geqslant \angle_{\eta(1)}([x,\eta(1)],[\gamma(1),\eta(1)])$. This in turn would imply $\overline{\angle}_{\gamma(1)}(x,\eta(1)) \geqslant \pi/2$ and $\overline{\angle}_{\eta(1)}(x,\gamma(1)) \geqslant \pi/2$, which is impossible. □

We call (H_x^*, d_*) the *dual* of H at x. We introduce a topology that is weaker than the one induced by $d_*(\cdot,\cdot)$. The attend concept of weak-* convergence is defined as follows: A sequence $(\phi_{\gamma_n})_{n\in\mathbb{N}} \subseteq H_x^*$ is said to weak-* converge to some $\phi_\gamma \in H_x^*$ if and only if $\lim_{n\to\infty}\phi_{\gamma_n}(x;y) = \phi_\gamma(x;y)$ for all $y \in H$. It is obvious that strong convergence, i.e. convergence with respect to d_*, implies weak-* convergence. Weak-* convergence gives rise to a topology which we call the weak-* topology on H_x^*.

Proposition 3.30. *A basis for the weak-* topology on H_x^* is determined by the sets*

$$U_\gamma(\varepsilon; y_1, y_2, ..., y_n) := \{\phi_\eta \in H_x^* : |\phi_\gamma(x;y_i) - \phi_\eta(x;y_i)| < \varepsilon, \forall i = 1,2,..,n\} \tag{3.13}$$

where $y_i \in H$, $\forall i = 1,2,...,n$; $n \in \mathbb{N}$ i.e., any open set in the weak- topology is a union of sets of the form* (3.13). *We denote this topology by τ_{w^*}.*

Proof. Note that $\phi_\gamma \in U_\gamma(\varepsilon; y_1, y_2, ..., y_n)$ for all $\varepsilon > 0$ and $y_1, y_2, ..., y_n \in H$. Hence

$$H_x^* \subseteq \bigcup_{\gamma\in\Gamma_x(H)} U_\gamma(\varepsilon; y_1, y_2, ..., y_n) \subseteq \bigcup_{\gamma\in\Gamma_x(H),\varepsilon>0,y_1,...,y_n\in H} U_\gamma(\varepsilon; y_1, y_2, ..., y_n).$$

Therefore the collection of sets $\{U_\gamma(\varepsilon; y_1, y_2, ..., y_n) : \gamma \in \Gamma_x(H), \varepsilon > 0, y_i \in H, i = 1,2,...,n\}$ covers H_x^*.

Need to check the intersection property. Let $U_{\gamma_1}(\varepsilon_1; y_1, y_2, ..., y_n)$ and $U_{\gamma_2}(\varepsilon_2; z_1, z_2, ..., z_m)$ and $\phi_\eta \in U_{\gamma_1}(\varepsilon_1; y_1, y_2, ..., y_n) \cap U_{\gamma_2}(\varepsilon_2; z_1, z_2, ..., z_m)$. There are $0 \leqslant s_1, s_2, ..., s_n < \varepsilon_1$ and $0 \leqslant t_1, t_2, ..., t_m < \varepsilon_2$ such that

$$|\phi_{\gamma_1}(x;y_i) - \phi_\eta(x;y_i)| = s_i$$
$$|\phi_{\gamma_2}(x;z_j) - \phi_\eta(x;z_j)| = t_j$$

for $i = 1,2,...,n$ and $j = 1,2,...,m$. Let $s := \max\{s_i : i = 1,2,...,n\}$ and $t := \max\{t_j : j = 1,2,...,m\}$. Consider the sets $U_\eta(\varepsilon_1 - s; y_1, ..., y_n)$ and $U_\eta(\varepsilon_2 - t; z_1, z_2, ..., z_m)$. Let $\phi_\sigma \in U_\eta(\varepsilon_1 - s; y_1, ..., y_n)$ then

$$\begin{aligned}|\phi_{\gamma_1}(x;y_i) - \phi_\sigma(x;y_i)| &\leqslant |\phi_{\gamma_1}(x;y_i) - \phi_\eta(x;y_i)| + |\phi_\eta(x;y_i) - \phi_\sigma(x;y_i)| \\ &< s_i + (\varepsilon_1 - s) \leqslant s + (\varepsilon_1 - s) = \varepsilon_1,\ \forall i = 1,2,...,n\end{aligned}$$

implying $\phi_\sigma \in U_{\gamma_1}(\varepsilon_1; y_1, ..., y_n)$. Hence $U_\eta(\varepsilon_1 - s; y_1, ..., y_n) \subseteq U_{\gamma_1}(\varepsilon_1; y_1, ..., y_n)$. Similarly $U_\eta(\varepsilon_2 - t; z_1, ..., z_m) \subseteq U_{\gamma_2}(\varepsilon_2; z_1, ..., z_m)$. Now let $\delta := \min\{\varepsilon_1 - s, \varepsilon_2 - t\}$ then

$$\begin{aligned}\phi_\eta \in U_\eta(\delta; y_1, ..., y_n, z_1, ..., z_m) &= U_\eta(\delta; y_1, ..., y_n) \cap U_\eta(\delta; z_1, ..., z_m)\\ &\subseteq U_\eta(\varepsilon_1 - s; y_1, ..., y_n) \cap U_\eta(\varepsilon_2 - t; z_1, ..., z_m)\\ &\subseteq U_{\gamma_1}(\varepsilon_1; y_1, ..., y_n) \cap U_{\gamma_2}(\varepsilon_2; z_1, ..., z_m)\end{aligned}$$

as desired. This completes the proof. □

Theorem 3.31. *The following properties hold:*

1. *A sequence $(\phi_{\gamma_n})_{n\in\mathbb{N}}$ weak-* converges to ϕ_γ if and only if $\phi_{\gamma_n} \overset{\tau_{w^*}}{\to} \phi_\gamma$.*
2. *(H_x^*, τ_{w^*}) is a Hausdorff space.*
3. *A weak-* closed set in H_x^* is closed.*

Proof. The first property follows from the definition of weak-* topology and Proposition 3.30. For the second property it suffices to show that for any two distinct elements ϕ_γ and ϕ_η there is $\varepsilon > 0$ such that the open sets $U_\gamma(\varepsilon; y)$ and $U_\eta(\varepsilon; y)$ have empty intersection for some $y \in H$. Let $\varepsilon := |\phi_\gamma(x; y) - \phi_\eta(x; y)|/2$ and suppose there is $\phi_\mu \in U_\gamma(\varepsilon; y) \cap U_\eta(\varepsilon; y)$ then

$$\begin{aligned}|\phi_\gamma(x;y) - \phi_\eta(x;y)| &\leqslant |\phi_\gamma(x;y) - \phi_\mu(x;y)| + |\phi_\mu(x;y) - \phi_\eta(x;y)|\\ &< \frac{1}{2}|\phi_\gamma(x;y) - \phi_\eta(x;y)| + \frac{1}{2}|\phi_\gamma(x;y) - \phi_\eta(x;y)|\\ &= |\phi_\gamma(x;y) - \phi_\eta(x;y)|\end{aligned}$$

which is impossible. In order to show the third property, let $A \subseteq H_x^*$ be a weak-* closed set, and let ϕ_{γ_n} be a sequence in A. Suppose that $\phi_{\gamma_n} \to \phi_\gamma$ in the strong topology, then $\phi_{\gamma_n} \overset{w^*}{\to} \phi_\gamma$, which implies that $\phi_\gamma \in A$. Therefore A is (strongly) closed. □

A direct consequence of Theorem 3.31 (2) is the following result.

Corollary 3.32. *A weak-* compact set in H_x^* has the Bolzano-Weierstrass property, the set of weak-* accumulation points coincides with the set of weak-* limit points, and a weak-* compact set is always weak-* closed.*

For $\alpha \in [0, 1]$ define the convex combination of $\phi_{\gamma_1}(x; \cdot)$ and $\phi_{\gamma_2}(x; \cdot)$ as $\phi_{\gamma_\alpha}(x; \cdot)$ where $\gamma_\alpha := (1 - \alpha)\gamma_1 \oplus \alpha\gamma_2$ is the geodesic segment in $\Gamma_x(H)$ that connects x with the point $(1-\alpha)\gamma_1(1) \oplus \alpha\gamma_2(1)$. This is well-defined since H is a uniquely geodesic space. It is worth noting that this convex operation in H_x^* is not defined in terms of the simple pointwise addition '+' of functions but in terms of convex combination along geodesics $\oplus$. However for H_x^* to be a convex metric space we need $W(\gamma_1, \gamma_2, \alpha) := \gamma_\alpha$ to be a convex structure. But it is not clear whether inequality (2.1) holds for the metric $d_*(\cdot, \cdot)$. In this setting a set $C \subseteq H_x^*$ is said to be convex whenever $\phi_{\gamma_1}, \phi_{\gamma_2} \in C$ and $\alpha \in [0, 1]$ imply $\phi_{\gamma_\alpha} \in C$. It is not clear to us whether the following properties hold:

Question 3.33. *Is a closed convex set weak-* closed?*

Question 3.34. *Is a bounded, closed and convex set weak-* compact?*

3.2.2 The case of a Hilbert space

Let H be a Hilbert space $(\mathcal{H}, \|\cdot\|)$ equipped with its canonical norm. Note that for every $x \in \mathcal{H}$ each geodesic segment $\gamma \in \Gamma_x(\mathcal{H})$ corresponds to a unique line l passing through x. Given $\gamma, \eta \in \Gamma_x(\mathcal{H})$ we say γ is equivalent to η, and write $\gamma \sim \eta$, if and only if γ, η belong to the same line l. Let $[l]$ denote the equivalence class of all geodesic segments $\gamma \in \Gamma_x(\mathcal{H})$ sharing the same line l.

Our aim is to show that our dual $\mathcal{H}_x^*$ coincides with the usual notion of the dual of a Hilbert space. We require the following lemma.

Lemma 3.35. *A sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ weakly converges (in our sense) to $x \in \mathcal{H}$ if and only if $P_l x_n \to x$ as $n \uparrow +\infty$ for all lines l containing x.*

Proof. Suppose that $\lim_n P_l x_n = x$ for all lines l containing x. Let $\gamma \in \Gamma_x(\mathcal{H})$ such that $\gamma \subset l$. Then all but finitely many of the terms $P_l x_n$ are in the image of γ, i.e., $P_l x_n = P_\gamma x_n$ for all sufficiently large n. This means $\lim_n P_\gamma x_n = \lim_n P_l x_n = x$. Since this holds for any l containing x and for any $\gamma \in [l]$ then $\lim_n P_\gamma x_n = x$ for any geodesic segment $\gamma \in \Gamma_x(\mathcal{H})$. Now let $x_n \overset{w}{\to} x$. By definition $\lim_n P_\gamma x_n = x$ for all $\gamma \in \Gamma_x(\mathcal{H})$. Each $\gamma \in \Gamma_x(\mathcal{H})$ determines a unique line l containing x. Then a similar argument shows that $\lim_n P_l x_n = x$ for all lines containing x. □

To the collection $\{\phi_\gamma(x;\cdot)\}_{\gamma\in[l]}$ we can associate a function $\phi_{[l]}(x;\cdot)$ defined as

$$\phi_{[l]}(x;y) = \|x - P_l y\|, \quad \forall y \in \mathcal{H}. \tag{3.14}$$

By Lemma 3.35 it follows that $\lim_n \phi_\gamma(x;x_n) = 0$ for all $\gamma \in [l]$ if and only if $\lim_n \phi_{[l]}(x;x_n) = 0$. Therefore it is sufficient to restrict to the family of functions $\{\phi_{[l]}(x;\cdot)\}_L$ where L is the set of all lines containing x. Clearly $\{\phi_{[l]}(x;\cdot)\}_L = H_x^*/\sim$. For a given $y \in H$ and $z_l \in l$ let θ be the the angle between the vectors $y - x$ and $z_l - x$. Then from the cosine formula for inner product we get

$$\langle y - x, z_l - x\rangle = \|y - x\|\|z_l - x\| \cos\theta \tag{3.15}$$

Realizing that $\|y - x\| \cos\theta = \pm\|P_\gamma y - x\|$ then follows

$$\pm\, \phi_{[l]}(x;y) = \frac{1}{\|z_l - x\|}\langle y - x, z_l - x\rangle, \quad \forall y \in \mathcal{H} \tag{3.16}$$

Using the linearity of the inner product one can rewrite (3.16) as

$$\phi_{[l]}(x;y) = \langle y - x, u_l\rangle, \quad \forall y \in \mathcal{H} \text{ where } u_l := \pm\frac{z_l - x}{\|z_l - x\|} \tag{3.17}$$

From (3.17) it is evident that $H_x^*/\sim$ together with its quotient metric coincides with the dual $\mathcal{H}^*$ of the Hilbert space $\mathcal{H}$.

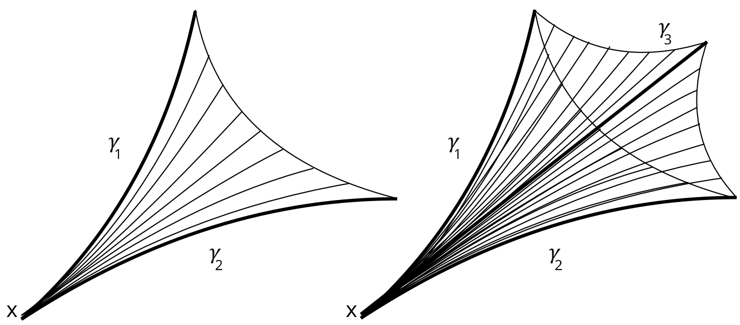


Figure 3.2: A geodesic segment in $\Gamma_x(H)$ (left) and a geodesic triangle in $\Gamma_x(H)$ (right).

3.3. The Space of Geodesic Segments

3.3.1 Metric space of geodesic segments

For each $x \in H$ the collection of geodesic segments $\Gamma_x(H)$ can be turned into a metric space by equipping it with the metric $d_1(\gamma, \eta) := \sup_{t\in[0,1]} d(\gamma(t), \eta(t))$ where $\gamma, \eta \in \Gamma_x(H)$. We call $(\Gamma_x(H), d_1)$ the metric space of geodesic segments and $\Gamma(H) := \bigcup_{x\in H} \Gamma_x(H)$ the bundle of geodesic segments in H. Furthermore we let e_x denote the trivial geodesic (of zero length) starting at x. Let $\psi : \Gamma_x(H) \to H$ be the mapping defined as $\psi(\gamma) := \gamma(1)$ for every $\gamma \in \Gamma_x(H)$.

Theorem 3.36. *The mapping ψ is a global isometry from $\Gamma_x(H)$ onto H for every $x \in H$. Therefore, the metric space $(\Gamma_x(H), d_1)$ is a Hadamard space for every $x \in H$.*

Proof. Clearly, ψ is a bijection due to the fact that any two points can be connected by a unique geodesic. It thus suffices to show that $d_1(\gamma, \eta) = d(\gamma(1), \eta(1))$. Consider the comparison triangle $\Delta(\bar{x}, \bar{\gamma}(1), \bar{\eta}(1))$. Then the distance $d(\gamma(t), \eta(t))$ is bounded above by the distance between the corresponding points in the comparison triangle, which is in turn is bounded above by the distance between $\bar{\gamma}(1)$ and $\bar{\eta}(1)$. This implies $d(\gamma(t), \eta(t)) \leq d(\gamma(1), \eta(1))$, which proves the claim. □

Given two geodesics $\gamma_1, \gamma_2 \in \Gamma_x(H)$ and $t \in [0, 1]$, consider the point $y_t := (1-t)\gamma_1(1) \oplus t\gamma_2(1)$ and the unique geodesic $\gamma_t \in \Gamma_x(H)$ that connects x with y_t. We call this geodesic $\gamma_t := (1-t)\gamma_1 \oplus t\gamma_2$. Figure 3.2 depicts a geodesic segment and a geodesic triangle in $\Gamma_x(H)$.

Corollary 3.37. *$(\Gamma_x(H), d_1)$ is a complete convex metric space where its convex structure W satisfies $W(\gamma_1, \gamma_2, t) = \gamma_t$ for all $\gamma_1, \gamma_2 \in \Gamma_x(H)$ and $t \in [0, 1]$.*

3.3.2 Weak convergence in $\Gamma_x(H)$

The metric $d_1(\cdot,\cdot)$ induces a strong topology where an open ball of radius $\varepsilon > 0$ centered at $\gamma \in \Gamma_x(H)$ is of the form $\mathbb{B}(\gamma, \varepsilon) := \{\eta \in \Gamma_x(H) : d_1(\gamma, \eta) < \varepsilon\}$. We can equip the metric space $\Gamma_x(H)$ also with a weak topology in the following way. We say a sequence $(\gamma_n)_{n\in\mathbb{N}} \subseteq \Gamma_x(H)$ weak-Γ converges to some element $\gamma \in \Gamma_x(H)$ and we denote it by $\gamma_n \overset{w_\Gamma}{\to} \gamma$ iff $\gamma_n(1) \overset{w}{\to} \gamma(1)$[2]. We say a set $U \subseteq \Gamma_x(H)$ is weak-Γ open if for any $\gamma \in U$ there exists $\varepsilon > 0$ and a finite family of geodesics $\gamma_1, ..., \gamma_n \in \Gamma_{\gamma(1)}(H)$ such that the set

$$U_\gamma(\varepsilon; \gamma_1, \gamma_2, ..., \gamma_n) := \{\eta \in \Gamma_x(H) : d(\gamma(1), P_{\gamma_i}\eta(1)) < \varepsilon, \forall i\}, \ \ \gamma_i \in \Gamma_{\gamma(1)}(H), n \in \mathbb{N} \tag{3.18}$$

is in U. It can be shown by exact same argument as in Proposition 3.2 that the collection of such sets U indeed defines a topology in $\Gamma_x(H)$. We wish to call it weak-Γ topology and denote it by τ_{w_Γ}. Weak-Γ convergence implies convergence in weak-Γ topology in the same way as weak convergence implies convergence in the weak topology for the underlying Hadamard space (H, d).

Using the isometry ψ defined in the previous section we can rewrite the definition of weak-Γ convergence as $\gamma_n \overset{w_\Gamma}{\to} \gamma$ iff $\psi(\gamma_n) \overset{w}{\to} \psi(\gamma)$. Similarly $\gamma_n \overset{\tau_{w_\Gamma}}{\to} \gamma$ iff $\psi(\gamma_n) \overset{\tau_w}{\to} \psi(\gamma)$. The mapping ψ enjoys also the property of being a homeomorphism between the topological spaces $(\Gamma_x(H), \tau_{w_\Gamma})$ and (H, τ_w).

Proposition 3.38. *The topological spaces* $(\Gamma_x(H), \tau_{w_\Gamma})$ *and* (H, τ_w) *are homeomorphic.*

Proof. It suffices to prove that the mapping ψ is a bicontinuous function i.e. it is bijective and it has a continuous inverse ψ^{-1}. Bijection follows from Proposition 3.36 since ψ is a global isometry. Now let $\gamma_n \overset{\tau_{w_\Gamma}}{\to} \gamma$ then by definition $\psi(\gamma_n) \overset{\tau_w}{\to} \psi(\gamma)$ hence ψ is continuous. For the other direction take any sequence $(y_n)_{n\in\mathbb{N}} \subseteq H$ such that $y_n \overset{\tau_w}{\to} y$ for some $y \in H$. Since H is a uniquely geodesic space for each y_n there is a unique $\gamma_n \in \Gamma_x(H)$ connecting it to x. Similarly let $\gamma \in \Gamma_x(H)$ be the geodesic segment connecting y with x. Then $y_n \overset{\tau_w}{\to} y$ means $\gamma_n(1) \overset{\tau_w}{\to} \gamma(1)$ or equivalently $\psi(\gamma_n) \overset{\tau_w}{\to} \psi(\gamma)$. By definition of weak-$\Gamma$ topology then $\gamma_n \overset{\tau_{w_\Gamma}}{\to} \gamma$ or equivalently $\psi^{-1}(y_n) \overset{\tau_{w_\Gamma}}{\to} \psi^{-1}(y)$. Hence ψ^{-1} is a continous mapping. $\square$

The last proposition implies that topologically $(\Gamma_x(H), \tau_{w_\Gamma})$ and (H, τ_w) are indistinguishable. As a direct result of this we have the following theorem:

Theorem 3.39. *The following statements are true:*

1. $(\Gamma_x(H), \tau_{w_\Gamma})$ *is Hausdorff whenever* (H, d) *is weakly proper (Lemma 3.5).*
2. *Any bounded sequence has a weak-*Γ *convergent subsequence (Lemma 3.8).*
3. *A closed convex set in* $\Gamma_x(H)$ *is weak-*Γ *sequentially closed. If additionally the set is also bounded then it is weak-*Γ *sequentially compact (Lemma 3.9, Theorem 3.19).*

[2]Not to be confused with Γ-convergence related to Mosco convergence (see Chapter 6)

4. *A bounded closed convex set in* $\Gamma_x(H)$ *is weak-*Γ *compact whenever* H *is separable (Theorem 3.23).*

Note that in case H is locally compact then so is the space $\Gamma_x(H)$ and weak-Γ topology coincides with the metric topology. In particular any closed and bounded set in $\Gamma_x(H)$ is compact and any bounded sequence $(\gamma_n)_{n\in\mathbb{N}} \subseteq \Gamma_x(H)$ has a convergent subsequence.

3.3.3 Existence of the steepest descent

A function $f : H \to \mathbb{R}$ is said to be geodesically differentiable at $x \in H$ along geodesic $\gamma \in \Gamma_x(H) \setminus \{e_x\}$ if the following limit exists

$$f'(x;\gamma) := \lim_{y \xrightarrow{\gamma} x} \frac{f(y) - f(x)}{d(y,x)}. \tag{3.19}$$

In case the limit in (3.19) exists for all $\gamma \in \Gamma_x(H) \setminus \{e_x\}$ then we simply say that f is geodesically differentiable at x. Note that this limit could vary from one geodesic to another and that $f'(x;\gamma) = f'(x;\eta)$ whenever $\eta \subseteq \gamma$. If (3.19) holds for all $x \in H$ then f is said to be geodesically differentiable on H. Furthermore we define

$$f\,'(x;e_x) := \inf_{(\gamma_n)\subset\Gamma_x(H)\setminus\{e_x\}} \liminf_{\gamma_n \to e_x} f\,'(x;\gamma_n). \tag{3.20}$$

For an illustrative example consider the Poincaré half-plane $\mathbb{H}^2 := \{x = (x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\}$ equipped with the metric $d(x,y) := \operatorname{arc\,cosh}(1 + \|x-y\|^2/2x_2y_2)$. Let $f : \mathbb{H}^2 \to \mathbb{R}$ be given by $f(x) = x_1 + x_2$. Take $x = (0, x_2)$ for some $x_2 > 0$. Recall that geodesics in $\mathbb{H}^2$ are either straight lines perpendicular to the horizontal axis or half circles centered on the horizontal axis. Let $\gamma, \eta : [0,1] \to \mathbb{H}^2$ be two geodesic segments starting from x; γ up in the ordinate axis and η lying in the second quadrant of $\mathbb{R}^2$ and part of the circle centered at the origin in $\mathbb{R}^2$ of radius x_2. By definition (3.19) we obtain

$$\begin{aligned} f'(x;\gamma) = \lim_{y \xrightarrow{\gamma} x} \frac{f(y)-f(x)}{d(y,x)} &= \lim_{y_2 \downarrow x_2} \frac{y_2 - x_2}{\operatorname{arc\,cosh}(1 + |y_2 - x_2|^2/2x_2y_2)} \\ &= \lim_{y_2 \downarrow x_2} \frac{y_2 - x_2}{|\ln y_2 - \ln x_2|} = \left(\lim_{y_2 \downarrow x_2} \frac{\ln y_2 - \ln x_2}{y_2 - x_2} \right)^{-1} = x_2 \end{aligned}$$

Now lets compute the derivative of f at x along η. Since the geodesic segment η is an arc of a circle it is helpful to write $y \in \eta$ in polar coordinates, i.e. $y = (x_2 \cos\theta, x_2 \sin\theta)$. Evidently $y \xrightarrow{\eta} x$ is equivalent to $\theta \downarrow \pi/2$. Again by definition (3.19) we obtain

$$f'(x;\eta) = \lim_{y \xrightarrow{\eta} x} \frac{f(y)-f(x)}{d(y,x)} = \lim_{\theta \downarrow \pi/2} \frac{x_2(\cos\theta + \sin\theta - 1)}{\operatorname{arc\,cosh}(1 + 2\sin^2(\theta/2 - \pi/4)/\sin\theta)} = -x_2$$

where the last limit follows by means of L'Hôpital's rule. Therefore the derivatives of a function at a point along different geodesic segments emanating from that point need not be equal.

Denote by $\overline{\mathbb{B}}(e_x)$ and $\mathbb{S}(e_x)$ the closed geodesic unit ball and the geodesic unit sphere respectively in $\Gamma_x(H)$ centered at e_x. A geodesic segment $\gamma_{\min} \in \mathbb{S}(e_x)$ of positive length is said to be the *direction of steepest descent* of the function f at $x \in H$ if $f'(x;\gamma_{\min}) = \inf_{\gamma\in\mathbb{S}(e_x)} f'(x;\gamma)$. Likewise a geodesic segment $\gamma_{\max} \in \mathbb{S}(e_x)$ is said to be the *direction of steepest ascent* of the function f at $x \in H$ if $f'(x;\gamma_{\max}) = \sup_{\gamma\in\mathbb{S}(e_x)} f'(x;\gamma)$.

Theorem 3.40 (Consequence of Extreme Value Theorem). *Let (H,d) be locally compact and $f : H \to \mathbb{R}$ be a geodesically differentiable function such that $f'(x;\gamma)$ is continuous in γ for each $x \in H$. Then $f'(x;\gamma)$ is bounded on closed bounded sets in $\Gamma_x(H)$ and there are $\gamma_{\max}, \gamma_{\min} \in \mathbb{S}(e_x)$ such that $f'(x;\gamma_{\max}) = \sup_{\gamma\in\mathbb{S}(e_x)} f'(x;\gamma)$ and $f'(x;\gamma_{\min}) = \inf_{\gamma\in\mathbb{S}(e_x)} f'(x;\gamma)$.*

Proof. Let $U \subseteq \Gamma_x(H)$ be some arbitrary closed bounded set and assume without loss of generality that $f'(x;\gamma)$ is not bounded from above on U. Then there is some sequence $(\gamma_n)_{n\in\mathbb{N}} \subseteq U$ such that $\lim_n f'(x;\gamma_n) = +\infty$. On the other hand $\Gamma_x(H)$ is locally compact since H is. This implies that U is compact because it is closed and bounded thus $(\gamma_n)_{n\in\mathbb{N}}$ has a convergent subsequence $(\gamma_{n_k})_{k\in\mathbb{N}}$. Let $\lim_k \gamma_{n_k} = \gamma$ then $\gamma \in U$. Assumption $f'(x;\gamma)$ is continuous in γ implies $\lim_k f'(x;\gamma_{n_k}) = f'(x;\gamma)$ but this contradicts $\lim_n f'(x;\gamma_n) = +\infty$. Therefore $f'(x;\gamma)$ must be bounded from above on U. Analogue arguments for boundedness from below. Now $\mathbb{S}(e_x) = \psi^{-1}(\mathbb{S}(x))$ and by Proposition 3.38 ψ is a homeomorphism so $\mathbb{S}(e_x)$ is compact. By extreme value theorem, since $f'(x;\gamma)$ is continuous in γ, there are $\gamma_{\max}, \gamma_{\min} \in \mathbb{S}(e_x)$ such that $f'(x;\gamma_{\max}) = \sup_{\gamma\in\mathbb{S}(e_x)} f'(x;\gamma)$ and $f'(x;\gamma_{\min}) = \inf_{\gamma\in\mathbb{S}(e_x)} f'(x;\gamma)$. This completes the proof. □

As an example consider the function $f : \mathbb{H}^2 \to \mathbb{R}$ defined as $f(x) := x_2$. We claim that $f'(x;\gamma)$ is continuous in γ for every $x \in \mathbb{H}^2$. For this it suffices that it is continuous at every x on the positive ordinate axis (y-axis) since by the theory of Möbius tranformation the other cases are just horizontal translations of the positive ordinate axis and all half-circle geodesics belonging to the points on this axis. Let $x = (0, x_2)$ for some $x_2 > 0$ and let $\gamma \in \Gamma_x(H)$. Without loss of generality suppose that the geodesic segment γ is part of an arclength of a half-circle geodesic passing through the point x. Note that the vertical line (ordinate axis) can be thought as the perimeter (or part of it) of an 'infinite radius' circle centered at either $+\infty$ or $-\infty$. Assume that the half-circle to which γ belongs is centered on the negative abscissa axis (the other situation follows similarly). Let $\theta_0 \in (0, \pi/2]$ be the smaller angle of the two subtended by the arc of this circle which joins its abscissa intersection points with the point x. Elementary geometry yields $x_2 = r\sin\theta_0$ where $r > 0$ is the radius of this circle. By Definition (3.19) we obtain

$$\begin{aligned} f'(x;\gamma) &= \lim_{y \xrightarrow{\gamma} x} \frac{f(y) - f(x)}{d(y,x)} = \lim_{y \xrightarrow{\gamma} x} \frac{y_2 - x_2}{d(y,x)} \\ &= \lim_{\theta \downarrow \theta_0} \frac{r(\sin\theta - \sin\theta_0)}{\operatorname{arc}\cosh(1 + 2\sin^2((\theta-\theta_0)/2)/\sin\theta_0\sin\theta)} = r\cos\theta_0\sin\theta_0 \end{aligned}$$

where the last limit is computed by applying twice L'Hôpital's rule. Using $x_2 = r\sin\theta_0$ we can rewrite $f'(x;\gamma) = x_2\cos\theta_0$. Now let $(\gamma_k)_{k\in\mathbb{N}}$ be a sequence of geodesic segments emanating from x such that $\lim_k \gamma_k = \gamma$. In particular this means that the Alexandrov's angle $\angle_x(\gamma, \gamma_k)$ vanishes as $k \uparrow +\infty$ which in turn is equivalent to saying that the sequence of circles $(C_k)_{k\in\mathbb{N}}$ associated with the sequence of geodesic segments $(\gamma_k)_{k\in\mathbb{N}}$ approaches the circle C where γ lies. If $\theta_k \in (0, \pi/2]$ is the smaller angle subtended by the arc of each circle C_k which joins its abscissa intersection points with the point x then it follows that $\lim_k \theta_k = \theta_0$. This in turn implies that $|f'(x;\gamma) - f'(x;\gamma_k)| = x_2|\cos\theta_0 - \cos\theta_k| \to 0$ by continuity of $\cos\theta$, i.e. $f'(x;\gamma)$ is continuous in γ. Since x was arbitrary on the ordinate axis then this holds true for any x on this axis and subsequently on all $\mathbb{H}^2$.

We say a function $g : \Gamma_x(H) \to \mathbb{R}$ is convex whenever $g(\gamma_t) \leqslant (1-t)g(\gamma_1) + tg(\gamma_2)$ for any $\gamma_1, \gamma_2 \in \Gamma_x(H)$ and $t \in [0,1]$ where γ_t is the convex combination of γ_1 and γ_2. In case (H,d) is not locally compact then a similar result holds if we additionally assume convexity of $f'(x;\gamma)$ in γ. We then have the following theorem.

Theorem 3.41. *Let (H,d) be a Hadamard space and $f : H \to \mathbb{R}$ be a geodesically differentiable function such that $f'(x;\gamma)$ is convex and lower semicontinuous in γ for each $x \in H$. Then $f'(x;\gamma)$ is bounded from below on bounded sets and there exists $\gamma_{\min} \in \overline{\mathbb{B}}(e_x)$ such that $f'(x;\gamma_{\min}) = \inf_{\gamma\in\overline{\mathbb{B}}(e_x)} f'(x;\gamma)$.*

Proof. Let f be geodesically differentiable. Then $f'(x;\gamma)$ exists and it is well defined for all $x \in H$ and $\gamma \in \Gamma_x(H)$. Consider the sub-level set $\mathrm{lev}_\alpha := \{\gamma \in \Gamma_x(H) : f'(x;\gamma) \leqslant \alpha\}$ where $\alpha \in \mathbb{R}$. The assumption that $f'(x;\gamma)$ is convex in γ implies that lev_α is a convex set. Moreover, since $f'(x;\gamma)$ is lower semicontinuous in γ, the sub-level sets lev_α are closed. By Theorem 3.39 (3) we obtain that lev_α is weak-Γ sequentially closed and consequently $f'(x;\gamma)$ is weak-Γ sequentially lower semicontinuous in γ.

Suppose that $f'(x;\gamma)$ is not bounded from below on bounded sets. Then there exists some bounded sequence $(\gamma_n)_{n\in\mathbb{N}} \subseteq \mathrm{lev}_\alpha$ such that $f'(x;\gamma_n) < -n$ for all $n \in \mathbb{N}$. By Theorem 3.39 (2) the sequence $(\gamma_n)_{n\in\mathbb{N}}$ has a weak-Γ convergent subsequence $(\gamma_{n_k})_{k\in\mathbb{N}}$. Let $\gamma_{n_k} \overset{w\Gamma}{\to} \gamma$ then $\gamma \in \mathrm{lev}_\alpha$. Weak-$\Gamma$ sequential lower semicontinuity of $f'(x;\gamma)$ then implies that $-\infty < f'(x;\gamma) \leqslant \liminf_k f'(x;\gamma_{n_k}) \leqslant \liminf_k(-n_k) = -\infty$, which is a contradiction.

The fact that there is $\gamma_{\min} \in \overline{\mathbb{B}}(e_x)$ such that $f'(x;\gamma_{\min}) = \inf_{\gamma\in\overline{\mathbb{B}}(e_x)} f'(x;\gamma)$ now follows from Theorem 3.39 (3), which implies that the set $\overline{\mathbb{B}}(e_x)$ is weak-Γ sequentially compact. Indeed, $f'(x;\gamma)$ is bounded below on $\overline{\mathbb{B}}(e_x)$. Let $(\gamma_n)_{n\in\mathbb{N}} \subseteq \overline{\mathbb{B}}(e_x)$ be a minimizing sequence, from which we can extract a weak-Γ convergent subsequence. The claim then follows weak-Γ sequential lower semicontinuity of $f'(x;\gamma)$.

□

From the relation a function f is concave iff $-f$ is convex follows the next corollary which we present without proof.

Corollary 3.42. *If $f'(x;\gamma)$ is concave and upper semicontinuous in γ then $f'(x;\gamma)$ is bounded from above on bounded sets and there exists $\gamma_{\max} \in \overline{\mathbb{B}}(e_x)$ such that $f'(x;\gamma_{\max}) = \sup_{\gamma\in\overline{\mathbb{B}}(e_x)} f'(x;\gamma)$.*

For an example of a function f defined on a Hadamard space (H,d) such that $f'(x;\gamma)$ is lower semicontinuous at some $\gamma \in \Gamma_x(H)$ for a given x consider the following. Let $H := \{x \in \mathbb{R}^2 : 0 \leqslant x_1 \leqslant 1, 0 \leqslant x_2 \leqslant 1\} \cup \{x \in \mathbb{R}^2 : -1 \leqslant x_1 \leqslant 0, -1 \leqslant x_2 \leqslant 0\}$ equipped with the length metric d. Then (H,d) is a Hadamard space. Let $f : H \to \mathbb{R}$ be defined as follows $f(y) = 0$ for $y \in \{x \in \mathbb{R}^2 : -1 \leqslant x_1 \leqslant 0, -1 \leqslant x_2 \leqslant 0\} \setminus \{0\}$, $f(y) = c$ for $y \in \{x \in H : x_1 = 0\}$ where c is some constant, and $f(y) := \sqrt{y_1^2 + y_2^2}$ otherwise. Let $\gamma : [0,1] \to H$ be the geodesic emanating from 0 in the ordinate axis of unit length and let $(\gamma_k)_{k\in\mathbb{N}}$ be a sequence of unit length geodesic segments emanating from 0 different from γ and lying in the positive quadrant (square). Let $\theta_k := \pi/2^k$ for $k \in \mathbb{N}$ be the Alexandrov's angle between γ_k and γ then $\lim_k \gamma_k = \gamma$. Direct calculations show that $f'(0;\gamma) = 0$ and $f'(0;\gamma_k) = 1$ for all $k \in \mathbb{N}$. Therefore $0 = f'(0;\gamma) \leqslant \liminf_k f'(0;\gamma_k) = 1$.

3.4. Other Forms of Weak Topology

There have been previous attempts to identify the correct topology corresponding to the weak convergence in Hadamard spaces. However, these attempts have identified topologies which are either too strong or too weak for capturing weak convergence in a Hadamard space. Nevertheless, these attempts have offered other perspectives on the notion of the weak convergence, which we compare to our notion.

3.4.1 Kakavandi's weak topology

Kakavandi [64] proposed a notion of weak topology, which is essentially based on the following observation. In a Hilbert space $(\mathcal{H}, \|\cdot\|)$ equipped with its canonical norm $\|\cdot\|$ a sequence $(x_n)_{n\in\mathbb{N}}$ converges weakly to an element $x \in \mathcal{H}$ iff $\lim_{n\to\infty}\langle x_n, y\rangle = \langle x, y\rangle$ for all $y \in \mathcal{H}$. This is equivalent to $\lim_{n\to\infty}\langle x_n - z, y - z\rangle = \langle x - z, y - z\rangle$ for all $y, z \in \mathcal{H}$. But the identity

$$\langle x - z, y - w\rangle = \frac{1}{2}(|x-y|^2 + |z-w|^2 - |x-w|^2 - |z-y|^2) \tag{3.21}$$

gives rise to the possibility of extending the definition of the weak convergence to a general metric space (X,d) by expressing the right side of (4.6) in terms of the metric $d(\cdot,\cdot)$. Following Berg and Nikolaev [24] consider the Cartesian product $X \times X$ where X is a general metric space. Each pair $(x,y) \in X \times X$ determines a so-called bound vector which is denoted by $\overrightarrow{xy}$. The point x is called the tail of $\overrightarrow{xy}$ and y is called the head. The zero bound vector is $\overrightarrow{0}_x = \overrightarrow{xx}$. The length of a bound vector $\overrightarrow{xy}$ is defined as the metric distance $d(x,y)$. Furthermore, if $\overrightarrow{u} := \overrightarrow{xy}$, then $-\overrightarrow{u} := \overrightarrow{yx}$. Let

$$\langle \overrightarrow{xz}, \overrightarrow{yw}\rangle := \frac{1}{2}(d(x,w)^2 + d(z,y)^2 - d(x,y)^2 - d(z,w)^2)\ . \tag{3.22}$$

Kakavandi's notion of weak convergence is defined in the following way. A sequence $(x_n)_{n\in\mathbb{N}}$ in a Hadamard space (H,d) converges weakly to an element $x \in H$ if and only if $\lim_{n\to\infty}\langle\overrightarrow{xx_n},\overrightarrow{xy}\rangle = 0$ for all $y \in H$. It is clear that this form of convergence coincides with the usual weak convergence in a Hilbert space. The identity (3.22) is what is known as the quasilinearization of a metric $d(\cdot,\cdot)$. It turns out that Kakavandi's form of convergence does not coincide with weak convergence in terms of asymptotic centers or equivalently along geodesics, see Example 4.7 in [64]. There is a natural topology associated to Kakavandi's convergence generated by the sets of the form

$$W(x, y; \varepsilon) := \{z \in H \,|\, |\langle\overrightarrow{xz},\overrightarrow{xy}\rangle| < \varepsilon\}, \quad \text{for any } x, y \in H, \varepsilon > 0. \tag{3.23}$$

More precisely the family of open sets $\{W(x, y; \varepsilon) \,|\, x, y \in H, \varepsilon > 0\}$ forms a subbasis for Kakavandi's topology τ_K. It is evident that $x_n \overset{K}{\to} x$ if and only if $x_n \overset{\tau_K}{\to} x$, for details see [64, Theorem 3.2].

Theorem 3.43. *Let (H, d) be a Hadamard space. Then the followings hold:*

1. *τ_w is coarser than τ_K.*
2. *(H, d) is Hausdorff space with respect to τ_K.*
3. *Kakavandi convergence and weak convergence coincide in a locally compact space.*

Proof. To show (1) let $U \in \tau_w$ be a weakly open set and $x \in U$. By construction of the topology τ_w there exist a finite number of geodesic segments $\{\gamma_i\}_{i=1}^n$ starting from x such that $\bigcap_{i=1}^n U_x(\delta; \gamma_i) \subset U$. For simplicity suppose $n = 1$, i.e., $U_x(\delta; \gamma) \subset U$ for some γ starting at x. Let $y \in \gamma$ such that $d(x, y) = \delta$ and $\varepsilon := \delta^2$. Consider the open set $W(x, y; \varepsilon)$ in τ_K. Let $z \in W(x, y; \varepsilon)$, then

$$|\langle\overrightarrow{xz},\overrightarrow{xy}\rangle| = d(x, z)d(x, y)|\cos\theta| < \varepsilon$$

where θ is the comparison angle of the Alexandrov angle at vertex x between the geodesic segments $[x, y]$ and $[x, z]$ in the comparison triangle $\overline{\Delta}$ determined by the points $\overline{x}, \overline{y}, \overline{z}$. Suppose that $d(x, P_\gamma z) \geqslant \delta$. Without loss of generality say $P_\gamma z = y$. Since γ is a closed convex set then by a property of projections the Alexandrov angle $\angle_y([y, x], [y, z])$ at y between geodesic segments $[y, x]$ and $[y, z]$ is at least $\pi/2$. By nonpositive curvature the comparison angle $\overline{\angle}_y([y, x], [y, z])$ is greater or equal to $\pi/2$. Then the projection of $\overline{z}$ onto the line $\overline{\gamma}$ extending beyond segment $[\overline{x}, \overline{y}]$ lies outside this segment, i.e., $|\overline{x}P_{\overline{\gamma}}\overline{z}| \geqslant |\overline{xy}| = \delta$. On the other hand we have $|\overline{x}P_{\overline{\gamma}}\overline{z}| = |\overline{xz}|\cos\theta = d(x, z)\cos\theta$ implies

$$|\langle\overrightarrow{xz},\overrightarrow{xy}\rangle| = d(x, z)d(x, y)|\cos\theta| = |\overline{xz}||\overline{xy}|\cos\theta = |\overline{x}P_{\overline{\gamma}}\overline{z}||\overline{xy}| \geqslant \delta^2 = \varepsilon.$$

This raises a contradiction.

To prove (2) let $x, y \in H$ be two distinct elements. Let $2\varepsilon := d(x, y)^2$ and consider the sets $W(x, y; \varepsilon)$ and $W(y, x; \varepsilon)$. By construction $x \in W(x, y; \varepsilon)$ and $y \in W(y, x; \varepsilon)$. Let $z \in W(x, y; \varepsilon) \cap W(y, x; \varepsilon)$ then we get the system of inequalities

$$|d(x, y)^2 + d(z, x)^2 - d(z, y)^2| < 2\varepsilon = d(x, y)^2$$
$$|d(y, x)^2 + d(z, y)^2 - d(z, x)^2| < 2\varepsilon = d(x, y)^2$$

which in turn implies $d(z, x) < d(z, y)$ and $d(z, y) < d(z, x)$. This is impossible.

Theorem 3.26 and [64][Proposition 4.4] imply (3). □

The previous theorem shows that the weak topology τ_w is coarser than Kakavandi's topology. Moreover in view of Theorem 3.43 (ii) it follows that a set S that is compact in τ_K is closed in τ_K and that τ_K-accumulation points coincide with τ_K-limit points. In particular in view of Theorem 3.15 one obtains that a compact set in τ_K always satisfies the Bolzano–Weierstrass property.

A theorem of Berg and Nikolaev ([24, Theorem 1, Corollary 3]) states that a metric space (X, d) that satisfies the inequality

$$|\langle \overrightarrow{xy}, \overrightarrow{zw} \rangle| \leqslant d(x,y)d(z,w) \tag{3.24}$$

for any pair of bound vectors $\overrightarrow{xy}, \overrightarrow{zw} \in X \times X$ must necessarily be a CAT(0) space. Since a Hadamard space is a complete CAT(0) space, this inequality automatically holds. In particular we obtain the estimate

$$\begin{aligned} |\langle \overrightarrow{xz}, \overrightarrow{xy_1} \rangle - \langle \overrightarrow{xz}, \overrightarrow{xy_2} \rangle| &= \frac{1}{2}|d(x,y_2)^2 + d(z,y_1)^2 - d(x,y_1)^2 - d(z,y_2)^2| \\ &\leqslant d(x,z)d(y_1,y_2). \end{aligned} \tag{3.25}$$

Suppose that (H, d) is separable. Then there exists some countable dense set $S := \{y_m\}$. Let $(x_n)_{n\in\mathbb{N}} \subseteq H$ be a bounded sequence and $x \in H$ such that $\lim_n \langle \overrightarrow{xx_n}, \overrightarrow{xy_m} \rangle = 0$ for all $m \in \mathbb{N}$. On the other hand for any $y \in H$ there is a sequence $(y_k)_{k\in\mathbb{N}} \subseteq S$ satisfying $y_k \to y$, which in view of 3.25 implies that $\lim_k \langle \overrightarrow{xx_n}, \overrightarrow{xy_k} \rangle = \langle \overrightarrow{xx_n}, \overrightarrow{xy} \rangle$ for every $n \in \mathbb{N}$. Notice that this convergence is uniform in $n \in \mathbb{N}$ since $(x_n)_{n\in\mathbb{N}}$ is bounded. By Moore-Osgood Theorem [101, Theorem 7.11] it follows that

$$\lim_n \langle \overrightarrow{xx_n}, \overrightarrow{xy} \rangle = \lim_n \lim_k \langle \overrightarrow{xx_n}, \overrightarrow{xy_k} \rangle = \lim_k \lim_n \langle \overrightarrow{xx_n}, \overrightarrow{xy_k} \rangle = 0. \tag{3.26}$$

Therefore $x_n \xrightarrow{K} x$. Following similar arguments as in the proof of Theorem 3.16, but instead using for open sets

$$V_n := \bigcap_{i=1}^{n} W(x, y_i; 1/n), \tag{3.27}$$

we obtain that in a separable Hadamard space a bounded set S which satisfies Bolzano–Weierstrass property in τ_K is τ_K-sequentially compact. On the other hand by virtue of Remark 3.20 we have that Theorem 3.19 holds true in τ_K and as a consequence also Theorem 3.22 on bounded sets. Moreover, in contrast to our weak topology τ_w, [64, Example 4.7] shows that a bounded sequence need not have a convergent subsequence in τ_K. Therefore in view of Theorem 3.23 we cannot say for certain whether a compactness result holds for bounded closed convex sets in Kakavandi's topology.

3.4.2 Monod's weak topology

Monod [90] proposed a topology on a Hadamard space (H, d), which we denote by τ_M, in the following way. The topology τ_M is the weakest topology on (H, d) such that any τ_s-closed convex set is τ_M-closed. Here τ_s is the usual metric topology on (H, d). Monod's topology was studied in detail by Kell [69] who refers to it as the co-convex topology. The next theorem relates the topologies discussed so far.

Proposition 3.44. *The following relations hold $\tau_M \subseteq \tau_w \subseteq \tau_K$ whenever the underlying Hadamard space (H,d) is weakly proper. All three topologies coincide with the usual weak topology whenever (H,d) is a Hilbert space.*

Proof. By Theorem 3.43 it follows that $\tau_w \subseteq \tau_K$. Suppose that H is weakly proper then from Theorem 3.6 we have that a convex set is τ_w-closed if and only if it is τ_s-closed. Hence $\tau_M \subseteq \tau_w$. From [90, Example 18] we know that whenever (H,d) is a Hilbert space, then the topologies τ_M and τ_K coincide with the usual weak topology. Since Hilbert spaces are weakly proper and $\tau_M \subseteq \tau_w \subseteq \tau_K$ then all three topologies coincide with the usual weak topology in Hilbert spaces. □

Question 3.45. *Does there exist a non-locally compact Hadamard space distinct from a Hilbert space or a product of a locally compact Hadamard space with a Hilbert space, on which $\tau_M = \tau_K = \tau_w$?*

It is known that if $K \subseteq H$ is compact then the restrictions of τ_M and τ_s to K coincide (see [90, Lemma 17]). Hence, in view of Proposition 3.44, the restrictions of all three weak topologies to a compact set K of a Hadamard space H coincide with the strong topology. An important property of Monod's topology τ_M is that any τ_s-closed convex and bounded set is τ_M-compact [90, Theorem 14]. However, it turns out that in general a Hadamard space is not Hausdorff in τ_M. For example, the Euclidean cone of an infinite dimensional Hilbert space is not a Hausdorff space when equipped with τ_M [69, Example 3.6]).

3.4.3 Geodesically monotone operators

A continuous operator $T : H \to H$ is said to be geodesically monotone if for all $x_0, x_1 \in H$ the real-valued function $\varphi : [0,1] \to \mathbb{R}_+$ defined by $\varphi(\alpha; x_0, x_1) := d(Tx_0, Tx_\alpha)$ is monotone in α where $x_\alpha := (1-\alpha)x_0 \oplus \alpha x_1$ is the convex combination along the geodesic from x_0 to x_1. The next theorem provides a sufficient condition for an arbitrary Hadamard space to be Hausdorff in Monod's topology.

Theorem 3.46. *If the projection P_γ is geodesically monotone for all geodesic segments γ, then (H, τ_M) is Hausdorff.*

Proof. Let $x, y \in H$ be two distinct points and $\gamma : [0,1] \to H$ a geodesic such that $\gamma(0) = x, \gamma(1) = y$. Let $l > 0$ denote the length of γ. For some fixed number $0 < \varepsilon < l$ let $C(x,\varepsilon) := \{z \in H \mid d(x, P_\gamma z) \leqslant \varepsilon\}$. We claim that $C(x,\varepsilon)$ is a closed convex set. Closedness follows immediately since P_γ is nonexpansive and therefore continuous. Let $z_0, z_1 \in C(x,\varepsilon)$ be two distinct elements. Let $z_\alpha := (1-\alpha)z_0 \oplus \alpha z_1$ for some $\alpha \in (0,1)$. By assumption, P_γ is a geodesically monotone operator; thus, $d(P_\gamma z_0, P_\gamma z_\alpha)$ is monotone in α, implying that $P_\gamma z_\alpha \in [P_\gamma z_0, P_\gamma z_1]$. The estimate $d(x, P_\gamma z_\alpha) \leqslant \max\{d(x, P_\gamma z_0), d(x, P_\gamma z_1)\} \leqslant \varepsilon$ implies that $z_\alpha \in C(x,\varepsilon)$. By definition of τ_M it follows that $C(x,\varepsilon)$ is τ_M-closed . Hence $H \setminus C(x,\varepsilon)$ is τ_M-open. By construction $y \in H \setminus C(x,\varepsilon)$. Using the same argument, it

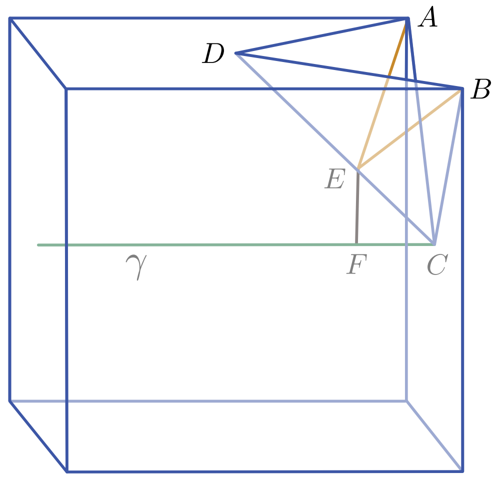


Figure 3.3: Example of a Hadamard space that arises by removing the wedge $(ABCD)$ from a cube. Both endpoints A and B of the geodesic segment $[A, B]$ (orange) connecting A with B project to the point C on the geodesic γ; however, the midpoint E of $[A, B]$ projects to the point F, which is not in the convex hull along γ of the projection of A and B. This illustrates that projection to geodesics might not preserve monotonicity.

follows that $C(y, \varepsilon) := \{z \in H \mid d(y, P_\gamma z) \leqslant l - \varepsilon\}$ is τ_M-closed. Hence, $H \setminus C(y, \varepsilon)$ is a τ_M-open set containing x. It is evident by construction that $(H \setminus C(x, \varepsilon)) \cap (H \setminus C(y, \varepsilon)) = \emptyset$. Therefore (H, τ_M) is a Hausdorff space. □

The contrapositive of this statement together with [69, Example 3.6] shows that the projection P_γ is not a geodesically monotone operator in a general Hadamard space. This is in contrast with projections to geodesic segments in Hilbert spaces, which are always geodesically monotone. Notice furthermore that the converse of Theorem 3.46 is not true: Figure 3.3 provides an example. Another interesting implication from the Example illustrated in Figure 3.3 relates to the so called *normal cone*. Given a closed convex set $C \subseteq H$ and $p \in C$ we define the normal cone at $p \in C$ and denote it by $N(p, C)$ as the set of all elements in H such that the geodesic segment connecting the given element with the point p makes an Aleksandrov angle no less then $\pi/2$ with any geodesic segment connecting p with another point in the set C, i.e.

$$N(p, C) := \{x \in H \,:\, \angle_p([x, p], [p, y]) \geqslant \pi/2, \; \forall y \in C\}. \tag{3.28}$$

Figure 3.3 tells us that $N(p, C)$ is not convex. This is in contrast with a basic result in Hilbert spaces that $N(p, C)$ is always a closed convex set.

CHAPTER 4

Convex Hulls of Compact Sets

4.1. The Closure of Convex Hulls

4.1.1 A characterization result in locally compact spaces

It is known that in a Banach space X given a compact set $K \subset X$ the closure of its convex hull $\operatorname{cl}\operatorname{co} K$ is compact [5, Theorem 5.35, p.185]. It is not known whether such a result carries over to Hadamard spaces. The problem remains widely open even for the simplest case of a set of only three points. This was pointed out first by Gromov in [54]. Motivated by the problem of the mean tree in phylogenetics, where it is can be shown that it lies in the closure of the convex hull of the given set of trees, in this chapter we investigate the closure of convex hulls of compact sets. However first in the setting of a locally compact Hadamard space.

Let $(A, \preceq)$ be a directed set and X a topological space. A *net* $(x_\alpha)_{\alpha\in A}$ in X is a mapping $\psi : A \to X$ defined as $\psi(\alpha) := x_\alpha$. We say a net $(x_\alpha)_{\alpha\in A}$ in X converges to an element x in X if for every neighborhood $U \subseteq X$ of x there exists $\alpha_0 \in A$ such that $x_\alpha \in U$ for every $\alpha \succeq \alpha_0$. The definition of nets can be naturally extended to that of nets of sets in the following way. A *net of sets* $(S_\alpha)_{\alpha\in A}$ in $\mathcal{P}(X)$ (the set of all subsets of X) is a mapping $\psi : A \to \mathcal{P}(X)$ defined as $\psi(\alpha) := S_\alpha$. We say a net of sets $(S_\alpha)_{\alpha\in A}$ is nondecreasing (nonincreasing) if $S_\alpha \subseteq S_\beta$ ($S_\alpha \supseteq S_\beta$) whenever $\alpha \preceq \beta$.

Definition 4.1. *Given a net of sets $(S_\alpha)_{\alpha\in A}$ its Painlevé -Kuratowski outer and inner limit are defined respectively as*

$$Ls_\alpha\, S_\alpha := \{x \in H : \forall V \in \mathcal{N}(x), \forall \alpha_0 \in A, \exists \alpha \succeq \alpha_0, \textit{such that}\; V \cap S_\alpha \neq \emptyset\} \tag{4.1}$$

$$Li_\alpha\, S_\alpha := \{x \in H : \forall V \in \mathcal{N}(x), \exists \alpha_0 \in A, \textit{such that}\; V \cap S_\alpha \neq \emptyset,\ \forall \alpha \succeq \alpha_0\} \tag{4.2}$$

where $\mathcal{N}(x)$ is the collection of neighborhoods at x. If (4.1) *and* (4.2) *coincide then the limit of the sequence $(S_\alpha)_{\alpha\in A}$ exists and we denote it by $Lim_\alpha\, S_\alpha$.*

Note that when $(A, \preceq) = (\mathbb{N}, \leqslant)$ then the above definition coincides with the usual definition of Painlevé–Kuratowski limit of a sequence of sets. The notion of Painlevé-Kuratowski limit becomes useful if one wants to extend the concept of semicontinuity of functions to semicontinuity of set valued mappings. For more details in this direction refer to [100, Rockafellar and Wetts]. It is known that in general any sequence of sets has a subsequence converging either to a nonempty set or the so called *horizon* (see [82, Theorem 3.11]). However for our purposes we only need the following lemma concerning nondecreasing nets of sets.

Lemma 4.2. *Let X be a topological space and $(S_\alpha)_{\alpha\in A}$ be a nondecreasing net of sets in X. Then $\operatorname{Lim}_\alpha S_\alpha = \operatorname{cl}\bigcup_{\alpha\in A} S_\alpha$.*

Proof. Let $x \in \operatorname{cl}\bigcup_{\alpha\in A} S_\alpha$ then for every $V \in \mathcal{N}(x)$ we have $V \cap \left(\bigcup_{\alpha\in A} S_\alpha\right) \neq \emptyset$. So there exists some $\alpha_0 \in A$ such that $V \cap S_{\alpha_0} \neq \emptyset$. Define $N := \{\alpha \in A : \alpha \succeq \alpha_0\}$. Assumption that $(S_\alpha)_{\alpha\in A}$ is a nondecreasing net of sets implies that $V \cap S_\alpha \neq \emptyset$ for all $\alpha \in N$. By definition (4.2) we get $x \in \operatorname{Li}_\alpha S_\alpha$. On the other hand nondecreasing property of the net $(S_\alpha)_{\alpha\in A}$ implies $\operatorname{Ls}_\alpha S_\alpha \subseteq \operatorname{Li}_\alpha S_\alpha$ hence $\operatorname{Li}_\alpha S_\alpha = \operatorname{Ls}_\alpha S_\alpha$. By virtue of Painlevé –Kuratowski definition it follows that $x \in \operatorname{Lim}_\alpha S_\alpha$. Thus $\operatorname{cl}\bigcup_{\alpha\in A} S_\alpha \subseteq \operatorname{Lim}_\alpha S_\alpha$. Now let $x \in \operatorname{Lim}_\alpha S_\alpha$ then again by definition of Painlevé –Kuratowski limit we have $x \in \operatorname{Li}_\alpha S_\alpha$. This means that for all $V \in \mathcal{N}(x)$ we have $V \cap \bigcup_{\alpha\in A} S_\alpha \neq \emptyset$ implying $x \in \operatorname{cl}\bigcup_{\alpha\in A} S_\alpha$. Consequently $\operatorname{Lim}_\alpha S_\alpha \subseteq \operatorname{cl}\bigcup_{\alpha\in A} S_\alpha$. Therefore $\operatorname{Lim}_\alpha S_\alpha = \operatorname{cl}\bigcup_{\alpha\in A} S_\alpha$. □

Definition 4.3. *We say a Hadamard space (H, d) is regular if for any nondecreasing net of compact sets $(K_\alpha)_{\alpha\in A}$ such that $\operatorname{Lim}_\alpha K_\alpha$ is bounded implies $\operatorname{Lim}_\alpha K_\alpha$ is compact.*

A collection of sets $\{S_\alpha\}_{\alpha\in A}$ in a topological space X, where A is now some arbitrary index set, is a *chain* in X if $S_\alpha \subseteq S_\beta$ or $S_\beta \subseteq S_\alpha$ whenever $\alpha \neq \beta$ i.e. the collection of sets $\{S_\alpha\}_{\alpha\in A}$ is a *totally ordered* subset of $\mathcal{P}(X)$. Note that by a well known theorem of Zermelo (see [92, Well-ordering theorem]), the index set A can be equipped with a partial ordering '$\preceq$' such that $(A, \preceq)$ is a well-ordered set. In particular $(A, \preceq)$ is a totally ordered set. Then formally we can say that a chain is a mapping $\psi : A \to \mathcal{P}(X)$ on a totally ordered set $(A, \preceq)$ defined as $\psi(\alpha) := S_\alpha$ such that the collection $\{S_\alpha\}_{\alpha\in A}$ is itself a totally ordered subset of $\mathcal{P}(X)$. It is clear that a chain is a net of sets and moreover they are either nondecreasing or nonincreasing. We refer to them as nondecreasing or nonincreasing chains respectively.

Theorem 4.4. *A Hadamard space (H, d) is locally compact if and only if it is regular.*

Proof. Let (H, d) be a locally compact Hadamard space $(A, \preceq)$ some directed set. Let $(K_\alpha)_{\alpha\in A}$ be a nondecreasing net of compact sets such that $\operatorname{Lim}_\alpha K_\alpha$ is bounded. By Lemma 4.2 $\operatorname{Lim}_\alpha K_\alpha = \operatorname{cl}\bigcup_{\alpha\in A} K_\alpha$. Hence $\operatorname{Lim}_\alpha K_\alpha$ is a closed and bounded subset of (H, d). By Theorem 3.24 it follows that $\operatorname{Lim}_\alpha K_\alpha$ is compact. The nondecreasing net of compact sets $(K_\alpha)_{\alpha\in A}$ was arbitrary, thus (H, d) is regular. Now we show the other direction. Suppose (H,d) is regular. Let $x \in H$ be arbitrary and $\overline{\mathbb{B}}(x,R)$ a closed geodesic ball centered at x with radius $R > 0$. Denote by

$$\mathcal{K} := \{K \subseteq \overline{\mathbb{B}}(x,R) : K \text{ is compact}\}.$$

Clearly $\mathcal{K}$ is nonempty since any finite set of points in $\overline{\mathbb{B}}(x,R)$ is a compact set. Moreover if $(K_\alpha)_{\alpha\in A}$ is a nondecreasing net of sets in $\mathcal{K}$ then by Lemma 4.2 it follows that the limit $K^* := \operatorname{Lim}_\alpha K_\alpha$ exists. By construction $K_\alpha \subseteq \overline{\mathbb{B}}(x,R)$ for every $\alpha \in A$ therefore K^* is bounded. Assumption that (H,d) is regular implies that K^* is a compact set and hence $K^* \in \mathcal{K}$. Define a partial order on $\mathcal{K}$ by $K_1 \preceq K_2$ iff $K_1 \subseteq K_2$. For every chain $(K_\beta)_{\beta\in B}$ in $\mathcal{K}$, where now B is some totally ordered set of indices, we have $K_\beta \preceq K^*$, hence every chain is bounded in $\mathcal{K}$. By Zorn's Lemma $\mathcal{K}$ has at least one maximal element. Denote this element by K^{**}. We claim that K^{**} concides with $\overline{\mathbb{B}}(x,R)$. Suppose that this is not the case. Then there is an element $y \in \overline{\mathbb{B}}(x,R) \setminus K^{**}$. Construct $K'_\beta := K_\beta \cup \{y\}$ where $(K_\beta)_{\beta\in B}$ is an arbitrary chain of compact sets in $\overline{\mathbb{B}}(x,R)$. Clearly $y \in K'_\beta$ and $K'_\beta \in \mathcal{K}$ for all $\beta \in B$. Let $(K')^* := \operatorname{Lim}_\beta K'_\beta$ then maximality of K^{**} requires that $(K')^* \preceq K^{**}$. But this is impossible since by construction $y \in (K')^*$ and $y \notin K^{**}$. Therefore $K^{**} = \overline{\mathbb{B}}(x,R)$ which in turn yields that $\overline{\mathbb{B}}(x,R) \in \mathcal{K}$ and hence $\overline{\mathbb{B}}(x,R)$ is compact. Then for any $R' < R$ the open geodesic ball $\mathbb{B}(x,R')$ is entirely contained in $\overline{\mathbb{B}}(x,R)$. Since $x \in H$ was arbitrary then (H,d) must be locally compact.

□

Lemma 4.5. *Let $K \subseteq H$ be a convex set and $\mathcal{F}_K$ the collection of all finite subsets S contained in K then*

$$K = \bigcup_{S\in\mathcal{F}_K} \operatorname{co} S \tag{4.3}$$

If in addition K is compact then there is an increasing sequence of sets $S_1 \subseteq S_2 \subseteq \ldots$ in $\mathcal{F}_K$ such that

$$K = \operatorname{cl} \bigcup_{n\in\mathbb{N}} \operatorname{co} S_n \tag{4.4}$$

Proof. Note that $S \subseteq K$ implies $\operatorname{co} S \subseteq K$ since K is convex and so

$$\bigcup_{S\in\mathcal{F}_K} \operatorname{co} S \subseteq K$$

Now let $y \in K$ then $\{y\} \in \mathcal{F}_K$ implies $\operatorname{co}\{y\} \in \bigcup_{S\in\mathcal{F}} \operatorname{co} S$. Therefore

$$K \subseteq \bigcup_{S\in\mathcal{F}} \operatorname{co} S$$

This proves identity (4.3). Now suppose in addition that K is compact. Then K is separable, i.e. K contains a countable dense subset $S := \{x_1, x_2, ..., x_n, ...\}$. Let $S_n := \{x_1, x_2, ..., x_n\}$ then $(S_n)_{n\in\mathbb{N}}$ is an increasing sequence in $\mathcal{F}_K$. It follows that $\operatorname{co} S_n \subseteq K$ since K is convex which in turn yields

$$\operatorname{cl} \bigcup_{n\in\mathbb{N}} \operatorname{co} S_n \subseteq K.$$

On the other hand we have the immediate inclusions

$$S_i \subseteq \bigcup_{n\in\mathbb{N}} \operatorname{co} S_n, \quad \forall i \in \mathbb{N}$$

implying

$$\operatorname{cl} S \subseteq \operatorname{cl} \bigcup_{n\in\mathbb{N}} \operatorname{co} S_n.$$

Then the equation $K = \operatorname{cl} S$ implies identity (4.4). □

For a given metric space (X, d) and two bounded sets A, B in X let $d_H(A, B)$ denote the *Hausdorff distance* between A and B defined as

$$d_H(A, B) := \max\{\sup_{y\in A} \inf_{x\in B} d(x, y), \sup_{x\in B} \inf_{y\in A} d(x, y)\}. \tag{4.5}$$

It is known that $d_H(\cdot, \cdot)$ is a *pseudometric* on the set of bounded sets of a metric space (X, d) (see [17]). When X is compact it is known that Painlevé-Kuratowski convergence coincides with Hausdorff distance convergence (see [78], [22, Corollary 5.1.11]).

A geodesic metric space (X, d) is said to have the *geodesic extension* property if for every geodesic $\gamma : [0, 1] \to X$ there exists $\varepsilon > 0$ and a geodesic $\tilde{\gamma} : [0, 1] \to X$ such that $l(\tilde{\gamma}) = l(\gamma) + \varepsilon$ and $\tilde{\gamma}|_{[0,l(\gamma)/l(\tilde{\gamma})]} = \gamma$. In the particular case when (X, d) is a CAT(0) space then geodesic extension property is equivalent to saying that any non-constant geodesic segment can be extended indefinitely to a geodesic line $\gamma : \mathbb{R} \to X$ [32, Lemma 5.8 (2)]. Examples of CAT(0) spaces satisfying geodesic extension property include but are not limited to Hilbert spaces, Hadamard manifolds, polyhedral complexes without free faces [32, Proposition 5.10], and in general any CAT(0) space that is homeomorphic to a finite dimensional manifold [32, Proposition 5.12].

We say a Hadamard space (H, d) satisfies the *finite set property* if for every $x \in H$ there exists an open set U containing x and a finite set S such that $U \subseteq \operatorname{cl} \operatorname{co} S$.

Theorem 4.6. *Let (H, d) be a Hadamard space satisfying the geodesic extension property. Then (H, d) is locally compact if and only if (H, d) satisfies the finite set property and for every compact set K the closure of its convex hull $\operatorname{cl} \operatorname{co} K$ is compact.*

Proof. Let (H, d) be a locally compact space. If $K \subseteq H$ is a compact set then it is in particular bounded. Hence there exists $R > 0$ such that $K \subseteq \mathbb{B}(x, R)$ for some $x \in K$. Clearly $\operatorname{co} K \subseteq \mathbb{B}(x, R)$ and subsequently $\operatorname{cl} \operatorname{co} K \subseteq \overline{\mathbb{B}}(x, R)$. Therefore $\operatorname{cl} \operatorname{co} K$ is a closed bounded set. By Theorem 3.24 we have that $\operatorname{cl} \operatorname{co} K$ is compact. Now suppose that there is some $x \in H$ such that for any open set U containing x there is no finite set S such that $U \subseteq \operatorname{cl} \operatorname{co} S$. Since (H, d) is locally compact by definition there is an open set V containing x and a compact set K such that $V \subseteq K$. Suppose K is convex, else take $\operatorname{cl} \operatorname{co} K$ which is also compact by the first implication. Consider any finite S in K containing x. For each such finite set S one can find a sequence (x_n) such that $x_n \in \mathbb{B}(x, \varepsilon_n) \setminus \operatorname{cl} \operatorname{co} S$ where $\mathbb{B}(x, \varepsilon_n) \subset V$ and $\lim_n \varepsilon_n = 0$. Clearly $(x_n)_{n\in\mathbb{N}} \subset H \setminus \operatorname{cl} \operatorname{co} S$. Then $\lim_n x_n = x$ implies $x \in \operatorname{cl}(H \setminus \operatorname{cl} \operatorname{co} S)$. The obvious inclusion $x \in \operatorname{cl} \operatorname{co} S$ yields that x is in the boundary of

$\operatorname{cl}\operatorname{co} S$ for any finite set S in K containing x. Since K is compact by Lemma 4.5 it follows that there is a collection of increasing sets $(S_n)_{n\in\mathbb{N}}$ such that

$$K = \operatorname{cl} \bigcup_{n\in\mathbb{N}} \operatorname{co} S_n \tag{4.6}$$

Since K is closed and convex the inclusions $\operatorname{cl}\operatorname{co} S_n \subseteq K$ for all $n \in \mathbb{N}$ imply

$$K = \operatorname{cl} \bigcup_{n\in\mathbb{N}} \operatorname{cl}\operatorname{co} S_n. \tag{4.7}$$

Assume without loss of generality that $x \in S_n$ for all $n \in \mathbb{N}$, else we can always add x to S_n and obtain a new sequence of increasing sets satisfying identity (4.6) or equivalently (4.7). By above arguments it follows that x is in the boundary of each $\operatorname{cl}\operatorname{co} S_n$. Moreover in view of Lemma 4.2 we have $\operatorname{Lim}_n \operatorname{cl}\operatorname{co} S_n = K$ and subsequently $\lim_n d_H(K, \operatorname{cl}\operatorname{co} S_n) = 0$ since K is compact. The inclusion $\operatorname{cl}\operatorname{co} S_n \subseteq K$ and definition (4.5) yield

$$d_H(K, \operatorname{cl}\operatorname{co} S_n) = \sup_{y\in K} \inf_{z\in \operatorname{cl}\operatorname{co} S_n} d(z, y).$$

Because $\operatorname{cl}\operatorname{co} S_n$ is a closed convex set then for each $y \in K$ its projection $P_{\operatorname{cl}\operatorname{co} S_n} y$ onto $\operatorname{cl}\operatorname{co} S_n$ exists and it is unique, therefore we obtain

$$d_H(K, \operatorname{cl}\operatorname{co} S_n) = \sup_{y\in K} d(y, P_{\operatorname{cl}\operatorname{co} S_n} y).$$

Let ∂K denote the boundary of K. Then $\partial K \subseteq K$ implies

$$\sup_{y\in K} d(y, P_{\operatorname{cl}\operatorname{co} S_n} y) \geqslant \sup_{y\in \partial K} d(y, P_{\operatorname{cl}\operatorname{co} S_n} y)$$

and thus

$$d_H(K, \operatorname{cl}\operatorname{co} S_n) \geqslant \sup_{y\in \partial K} d(y, P_{\operatorname{cl}\operatorname{co} S_n} y). \tag{4.8}$$

There exists a sequence $(y_k)_{k\in\mathbb{N}} \subseteq K \setminus \{x\}$ such that $\lim_k y_k = x$. Let $P_{\operatorname{cl}\operatorname{co} S_n} y_k$ denote the metric projection of y_k onto $\operatorname{cl}\operatorname{co} S_n$ for every $k \in \mathbb{N}$. Denote by $\gamma_k : [0,1] \to H$ the geodesic segment connecting $P_{\operatorname{cl}\operatorname{co} S_n} y_k$ with y_k. By assumption (H, d) satisfies the geodesic extension property. Therefore there exists geodesic lines $\widetilde{\gamma}_k : \mathbb{R} \to H$ such that $\widetilde{\gamma}_k|_{[0,1]} = \gamma_k$ for every $k \in \mathbb{N}$. Since K is bounded and the image of $\widetilde{\gamma}_k$ is connected then there exists $z_k \in \widetilde{\gamma}_k \cap \partial K$ for every $k \in \mathbb{N}$. From the equation $P_{\operatorname{cl}\operatorname{co} S_n} z_k = P_{\operatorname{cl}\operatorname{co} S_n} y_k$ for all $k \in \mathbb{N}$ we obtain the inequalities

$$d(z_k, y)^2 \geqslant d(z_k, P_{\operatorname{cl}\operatorname{co} S_n} z_k)^2 + d(P_{\operatorname{cl}\operatorname{co} S_n} z_k, y)^2,\ \forall k \in \mathbb{N}, \forall y \in \operatorname{cl}\operatorname{co} S_n.$$

Note that $\lim_k d(P_{\operatorname{cl}\operatorname{co} S_n} z_k, x) = \lim_k d(P_{\operatorname{cl}\operatorname{co} S_n} y_k, x) \leqslant \lim_k d(y_k, x) = 0$ implies that $\lim_k P_{\operatorname{cl}\operatorname{co} S_n} z_k = x$. Since $(z_k)_{k\in\mathbb{N}} \subseteq \partial K$ and ∂K is compact then there is a subsequence $(z_{k_m})_{m\in\mathbb{N}} \subseteq (z_k)_{k\in\mathbb{N}}$ converging to some element $z \in \partial K$. Passing in the limit we obtain

$$\begin{aligned} \lim_m d(z_{k_m}, y)^2 &\geqslant \lim_m d(z_{k_m}, P_{\operatorname{cl}\operatorname{co} S_n} z_{k_m})^2 + \lim_m d(P_{\operatorname{cl}\operatorname{co} S_n} z_{k_m}, y)^2 \\ &\Rightarrow d(z, y)^2 \geqslant d(z, x)^2 + d(x, y)^2,\ \forall y \in \operatorname{cl}\operatorname{co} S_n \end{aligned}$$

In particular when $y = P_{\operatorname{cl}\operatorname{co} S_n} z$ we get

$$d(z, P_{\operatorname{cl}\operatorname{co} S_n} z)^2 \geqslant d(z,x)^2 + d(x, P_{\operatorname{cl}\operatorname{co} S_n} z)^2.$$

On the other hand by the property of the projection onto a closed convex set we have the inequality

$$d(z,x)^2 \geqslant d(z, P_{\operatorname{cl}\operatorname{co} S_n} z)^2 + d(x, P_{\operatorname{cl}\operatorname{co} S_n} z)^2.$$

Together these last two inequalities imply $d(x, P_{\operatorname{cl}\operatorname{co} S_n} z) \leqslant 0$ and hence $x = P_{\operatorname{cl}\operatorname{co} S_n} z$. From (4.8) it follows then

$$d_H(K, \operatorname{cl}\operatorname{co} S_n) \geqslant \sup_{y \in \partial K} d(y, P_{\operatorname{cl}\operatorname{co} S_n} y) \geqslant d(z, P_{\operatorname{cl}\operatorname{co} S_n} z) = d(z,x) \geqslant \inf_{u \in \partial K} d(u,x) = d(x, \partial K).$$

In the limit we obtain $\lim_n d(x, \partial K) \leqslant \lim_n d_H(K, \operatorname{cl}\operatorname{co} S_n) = 0$ or equivalently $x \in \partial K$. However this is impossible since $x \in \operatorname{int} K$. Now we show the opposite direction. Suppose that for every compact set K the closure of its convex hull $\operatorname{cl}\operatorname{co} K$ is compact and that (H,d) satisfies the finite set property. For every $x \in H$ there is an open set U containing x and a finite set S such that $U \subseteq \operatorname{cl}\operatorname{co} S$. But S is finite and therefore compact which in turn yields that $\operatorname{cl}\operatorname{co} S$ is compact. By definition of local compactness it follows that (H,d) is a locally compact space. □

An immediate application of Theorem 4.4 and Theorem 4.6 is the next corollary for Hadamard spaces satisfying the geodesic extension property.

Corollary 4.7. *A Hadamard space (H,d) satisfying the geodesic extension property is regular if and only if (H,d) satisfies the finite set property and for every compact set K the closure of its convex hull $\operatorname{cl}\operatorname{co} K$ is compact.*

Remark 4.8. *Note that Theorem 4.6 and Corollary 4.7 hold in more generality for any complete metric space (X,d) which admits a convex structure W and such that line segments have the extension property. In view of Proposition 2.5 these results are true in any geodesic metric space with the geodesic extension property. On the other hand Theorem 4.4 is even more general. Since no notion of convexity was used in its proof this means that it holds true in a metric space not necessarily admiting a convex structure W.*

4.1.2 The case of weakly compact sets

For a Hadamard space not necessarily locally compact we make the following conclusion for bounded weakly compact sets.

Theorem 4.9. *Let (H,d) be a separable Hadamard space and $K \subseteq H$ a bounded weakly compact set. Then $\operatorname{cl}\operatorname{co} K$ is weakly compact.*

Proof. Assume $K \subseteq H$ is a bounded weakly compact set. Since it is bounded there exists some $R > 0$ such that $K \subseteq \overline{\mathbb{B}}(x,R)$ for some $x \in H$. But $\overline{\mathbb{B}}(x,R)$ is itself closed and convex and therefore $\operatorname{cl}\operatorname{co} K \subseteq \overline{\mathbb{B}}(x,R)$. This implies that $\operatorname{cl}\operatorname{co} K$ is also bounded. By Theorem 3.23 $\operatorname{cl}\operatorname{co} K$ is weakly compact. □

Corollary 4.10. *If $K \subseteq H$ is a compact set then* $\operatorname{cl}\operatorname{co} K$ *is weakly compact whenever* (H, d) *is separable.*

Note that Theorem 4.9 can be regarded as a variant of Krain-Šmulian Theorem for separable Hadamard spaces. Since Krain-Šmulian Theorem holds for any Banach space (see [115, Whitley]) we are lead to the following question.

Question 4.11. *Is Krein-Šmulian Theorem true in a non-separable Hadamard space distinct from a Hilbert space?*

4.2. Threading

4.2.1 Main properties of threading

For $S \subseteq H$ the *threading* of S denoted by $\operatorname{thr} S$ is defined to be the union of all geodesic segments with both endpoints in S. It follows from this definition that if $z \in \operatorname{thr} S$ then there are $x, y \in S$ such that $z \in [x, y]$. In general $[x', y'] \subseteq \operatorname{thr} S$ if and only if there are $x, y \in S$ such that $[x', y'] \subseteq [x, y]$. One can then iteratively define threading of threading of a set and so on. Given a set S we have the following chain of inclusions $S \subseteq \operatorname{thr} S \subseteq \operatorname{thr}^2 S \subseteq ... \subseteq \operatorname{thr}^n S \subseteq ...$ and the equation $\operatorname{thr}^n S = \operatorname{thr}(\operatorname{thr}^{n-1} S)$ for all $n \in \mathbb{N}$ is immediate.

Proposition 4.12. *In general the following rules hold*

$$\operatorname{thr}(S_1 \cap S_2) \subseteq \operatorname{thr} S_1 \cap \operatorname{thr} S_2 \tag{4.9}$$

$$\operatorname{thr} S_1 \cup \operatorname{thr} S_2 \subseteq \operatorname{thr}(S_1 \cup S_2) \tag{4.10}$$

with equality in both if $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$.

Proof. Let $z \in \operatorname{thr}(S_1 \cap S_2)$ then by definition of threading we have $x, y \in S_1 \cap S_2$ such that $z \in [x, y]$. But $x, y \in S_1 \cap S_2$ implies $x, y \in S_1$ and $x, y \in S_2$. Therefore $[x, y] \subseteq \operatorname{thr} S_1$ and $[x, y] \subseteq \operatorname{thr} S_2$ and so $z \in \operatorname{thr} S_1 \cap \operatorname{thr} S_2$. Then identity (4.9) follows. Now let $z \in \operatorname{thr} S_1 \cup \operatorname{thr} S_2$ then $z \in \operatorname{thr} S_1$ or $z \in \operatorname{thr} S_2$. There are $x_1, y_1 \in S_1$ and $x_2, y_2 \in S_2$ such that $z \in [x_1, y_1]$ or $z \in [x_2, y_2]$. By construction since $x_1, y_1, x_2, y_2 \in S_1 \cup S_2$ then $[x_1, y_1], [x_2, y_2] \subseteq \operatorname{thr}(S_1 \cup S_2)$ implies $z \in \operatorname{thr}(S_1 \cup S_2)$. This proves identity (4.10). If $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$ the equalities are evident in both identities. □

In general one can find sets S_1, S_2 such that (4.9) or (4.10) hold with strict inclusion. Consider the simplest example of $H = \mathbb{R}$ equipped with the usual metric $d(r, s) = |r - s|$ for any two real numbers $r, s \in \mathbb{R}$. Take $S_1 := \{0, 2\}, S_2 = \{1\}$. Then $S_1 \cap S_2 = \emptyset$ implies

$$\operatorname{thr}(S_1 \cap S_2) = \emptyset \subsetneq \{1\} = [0, 2] \cap \{1\} = \operatorname{thr} S_1 \cap \operatorname{thr} S_2$$

Similarly for the other identity let $S_2 = \{3\}$ then

$$\operatorname{thr} S_1 \cup \operatorname{thr} S_2 = [0, 2] \cup \{3\} \subsetneq [0, 3] = \operatorname{thr}(S_1 \cup S_2)$$

Lemma 4.13. *Let $S \subseteq H$ and $W(x,y,t)$ given by (2.23) then*

$$\operatorname{thr} S = \{W(x,y,t) \mid x,y \in S, t \in [0,1]\} \tag{4.11}$$

Proof. Denote the right side of equality (4.11) by D. If $z \in D$ then there are $x,y \in S$ and $t \in [0,1]$ such that $z = W(x,y,t)$ i.e. by definition (2.23) $z = (1-t)x \oplus ty$. Hence $z \in [x,y]$ implies $z \in \operatorname{thr} S$. This shows $D \subseteq \operatorname{thr} S$. Now let $z \in \operatorname{thr} S$. By definition of threading there are $x,y \in S$ such that $z \in [x,y]$. Geodesics are parameterized by constant speed therefore there is $t \in [0,1]$ such that $z = (1-t)x \oplus ty$ i.e. $z = W(x,y,t) \in D$ and thus $\operatorname{thr} S \subseteq D$ which completes the proof. □

Proposition 4.14. *Let (H_1, d_1) and (H_2, d_2) be two Hadamard spaces and $\Phi : H_1 \to H_2$ be an isometry isomorphism. Then $\Phi(\operatorname{thr} S) = \operatorname{thr} \Phi(S)$ for any $S \subseteq H_1$.*

Proof. Let $x,y \in S$ then $[x,y] \subseteq \operatorname{thr} S$. Denote by $x_t := (1-t)x \oplus ty$ for $t \in [0,1]$ then

$$d_2(\Phi(x), \Phi(x_t)) = d_1(x, x_t) = td_1(x,y) = td_2(\Phi(x), \Phi(y))$$

implies $\Phi(x_t) = (1-t)\Phi(x) \oplus t\Phi(y)$ for all $t \in [0,1]$. Hence $\Phi([x,y]) = [\Phi(x), \Phi(y)]$. This shows $\Phi(\operatorname{thr} S) \subseteq \operatorname{thr} \Phi(S)$. For the other direction use the inverse mapping Φ^{-1} which is as well an isometry. The same arguments yield $\operatorname{thr} \Phi(S) \subseteq \Phi(\operatorname{thr} S)$. □

Theorem 4.15. *If $S \subseteq H$ is convex then $\operatorname{thr} S = S$. In general for any set $S \subseteq H$ we have the following identity*

$$\bigcup_{n \in \mathbb{N}} \operatorname{thr}^n S = \operatorname{co} S \tag{4.12}$$

Proof. The inclusion $S \subseteq \operatorname{thr} S$ is clear whether S is convex or not. Now let $z \in \operatorname{thr} S$. By definition of threading there are $x,y \in S$ such that $z \in [x,y]$. Assumption S is convex implies $[x,y] \subseteq S$ therefore $z \in S$ which in turn yields $\operatorname{thr} S \subseteq S$. To prove identity (4.12) let $x,y \in \bigcup_{n\in\mathbb{N}} \operatorname{thr}^n S$ then there are $l,m \in \mathbb{N}$ such that $x \in \operatorname{thr}^l S, y \in \operatorname{thr}^m S$. Assume that $m \geqslant l$ then $\operatorname{thr}^l S \subseteq \operatorname{thr}^m S$ implies $x \in \operatorname{thr}^m S$. Therefore $x,y \in \operatorname{thr}^m S$ yields by definition of threading $[x,y] \subseteq \operatorname{thr}^{m+1} S \subseteq \bigcup_{n\in\mathbb{N}} \operatorname{thr}^n S$. Hence $\bigcup_{n\in\mathbb{N}} \operatorname{thr}^n S$ is a convex set. But $\operatorname{co} S$ is the smallest convex set containing S thus

$$\operatorname{co} S \subseteq \bigcup_{n \in \mathbb{N}} \operatorname{thr}^n S$$

On the other hand from (4.9) and the first part of this theorem we have

$$\operatorname{thr} S = \operatorname{thr}(S \cap \operatorname{co} S) \subseteq \operatorname{thr} S \cap \operatorname{thr} \operatorname{co} S \subseteq \operatorname{thr} \operatorname{co} S = \operatorname{co} S$$

therefore iterative threading implies

$$\operatorname{thr}^n S \subseteq \operatorname{co} S, \ \ \forall n \in \mathbb{N} \Rightarrow \bigcup_{n \in \mathbb{N}} \operatorname{thr}^n S \subseteq \operatorname{co} S$$

□

For a given set $S \subseteq H$ the *threading degree* of S denoted by $\deg_{\text{thr}} S$ is defined as the smallest $n \in \mathbb{N}$, if it exists, such that $\text{thr}^n S = \text{co}\, S$. In case such an $n \in \mathbb{N}$ does not exist then we set $\deg_{\text{thr}} S := +\infty$. The threading degree of a Hadamard space H is then the supremum of such n over the set of all subsets of H i.e.

$$\deg_{\text{thr}} H := \sup_{S \subseteq H} \deg_{\text{thr}} S \tag{4.13}$$

If $\deg_{\text{thr}} H$ is finite we say that H is of *finite type.*

Proposition 4.16. *Let (H, d) be a Hadamard space of finite type. For any closed convex set $S \subseteq H$ the subspace (S, d_S) is of finite type, where $d_S := d|_S$.*

Proof. First note that (S, d_S) is itself a Hadamard space. Let $\text{thr}_S S'$ and $\text{thr}_H S'$ denote threading applied to $S' \subseteq S$ in S and H respectively. Since $d_S = d|_S$ then $\text{thr}_S S' = \text{thr}_H S'$. By definition (4.13) it follows

$$\deg_{\text{thr}_S} S = \sup_{S' \subseteq S} \deg_{\text{thr}_S} S' = \sup_{S' \subseteq S} \deg_{\text{thr}_H} S' \leqslant \sup_{S' \subseteq H} \deg_{\text{thr}_H} S' = \deg_{\text{thr}_H} H$$

By assumption $\deg_{\text{thr}_H} H < +\infty$ we get that (S, d_S) is of finite type. □

Proposition 4.17. *If $(H_1, d_1), (H_2, d_2)$ are of finite type then (H, d), where $H := H_1 \times H_2$ and $d(\cdot, \cdot)$ is its canonical metric, is of finite type. Moreover the following holds*

$$\deg_{\text{thr}} H = \max\{\deg_{\text{thr}} H_1, \deg_{\text{thr}} H_2\}$$

Proof. Let $S \subseteq H$ then $S = S_1 \times S_2$ for some $S_1 \subseteq H_1, S_2 \subseteq H_2$. Let $z \in \text{thr}\, S$ then there are $x, y \in S$ such that $z \in [x, y]$. On the other hand $x = (x_1, x_2)$ and $y = (y_1, y_2)$ for some $x_1, y_1 \in S_1$ and $x_2, y_2 \in S_2$. By construction if follows $[x, y] = [x_1, y_1] \times [x_2, y_2]$ therefore $z = (z_1, z_2)$ for some $z_1 \in [x_1, y_1]$ and $z_2 \in [x_2, y_2]$. This means that $z \in \text{thr}\, S_1 \times \text{thr}\, S_2$. The other direction is proved analogously. In general we get that $\text{thr}^n S = \text{thr}^n S_1 \times \text{thr}^n S_2$ for any $n \in \mathbb{N}$. It follows then that $\deg_{\text{thr}} S = \max\{\deg_{\text{thr}} S_1, \deg_{\text{thr}} S_2\}$. On the other hand we have the identity $\{S : S \subseteq H\} = \{S_1 : S_1 \subseteq H_2\} \times \{S_2 : S_2 \subseteq H_2\}$. Taking supremum over $S \subseteq H$

$$\sup_{S \subseteq H} \deg_{\text{thr}} S = \sup_{S_1 \subseteq H_1, S_2 \subseteq H_2} \max\{\deg_{\text{thr}} S_1, \deg_{\text{thr}} S_2\} = \max\{\sup_{S_1 \subseteq H_1} \deg_{\text{thr}} S_1, \sup_{S_2 \subseteq H_2} \deg_{\text{thr}} S_2\}$$

which is equivalent to

$$\deg_{\text{thr}} H = \max\{\deg_{\text{thr}} H_1, \deg_{\text{thr}} H_2\}.$$

□

Notice that in general if $S_1, S_2 \subseteq H$ are such that $S_1 \subseteq S_2$ then neither $\deg_{\text{thr}} S_1 \leqslant \deg_{\text{thr}} S_2$ nor $\deg_{\text{thr}} S_2 \leqslant \deg_{\text{thr}} S_1$ is necessarily true. Consider $H = \mathbb{R}^2$ equipped with the usual Euclidean metric and let S_1 be four corners of a square and S_2 be the square itself. Clearly $S_1 \subseteq S_2$ but $\deg_{\text{thr}} S_1 = 2 > 1 = \deg_{\text{thr}} S_1$. Similarly one can show that if instead S_1 is just two of the corner points and S_2 is the four corner points then $\deg_{\text{thr}} S_1 = 1 < 2 = \deg_{\text{thr}} S_2$. Figure 4.1 depicts an example of threading for three distinct non-collinear points in the Euclidean plane $\mathbb{R}^2$.

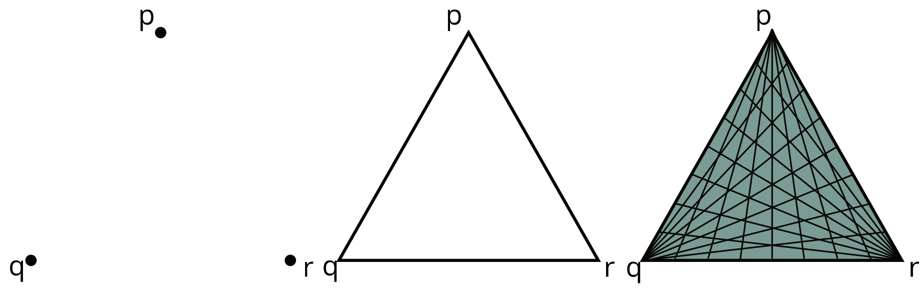


Figure 4.1: Three points $\{p, q, r\}$ in $\mathbb{R}^2$ (left), $\operatorname{thr}\{p, q, r\}$ is the Euclidean triangle (middle) and $\operatorname{thr}^2\{p, q, r\}$ is eventually the solid Euclidean triangle (right).

4.2.2 Threading of compact sets

Lemma 4.18. *If $K \subseteq H$ is compact then $\operatorname{thr}^n K$ is compact for all $n \in \mathbb{N}$.*

Proof. First we prove that $\operatorname{thr} K$ is compact. Let $(x_k)_{k\in\mathbb{N}} \subseteq \operatorname{thr} K$ be some sequence. Then by Lemma (4.13) there are $y_k, z_k \in K$ and $t_k \in [0, 1]$ such that $x_k = W(y_k, z_k, t_k)$ for all k. By assumption K is compact the so is the product space $K \times K \times [0, 1]$. There are convergent subsequences $(y_{k_m})_{m\in\mathbb{N}}, (z_{k_m})_{m\in\mathbb{N}} \subseteq K$ and $(t_k)_{k\in\mathbb{N}} \subseteq [0, 1]$. Let $\lim_m y_{k_m} = y, \lim_m z_{k_m} = z$ and $\lim_m t_{k_m} = t \in [0, 1]$. Then it follows $\lim_m x_{k_m} = \lim_m W(y_m, z_m, t_m) = W(y, z, t) \in \operatorname{thr} K$. Therefore $\operatorname{thr} K$ is sequentially compact and therefore compact. Now suppose $\operatorname{thr}^{n-1} K$ is compact. From the equation $\operatorname{thr}^n K = \operatorname{thr}(\operatorname{thr}^{n-1} K)$ we get that $\operatorname{thr}^n K$ is compact for any $n \in \mathbb{N}$. □

Lemma 4.19. *Let $K \subseteq H$ be compact. If $\deg_{\operatorname{thr}} K$ is finite then $\operatorname{co} K$ is compact.*

Proof. If $\deg_{\operatorname{thr}} K$ is finite then there is some $n \in \mathbb{N}$ such that $\operatorname{thr}^n K = \operatorname{co} K$. By Lemma 4.18 $\operatorname{thr}^n K$ is compact since by assumption K is a compact set. Therefore $\operatorname{co} K$ is compact. □

Proposition 4.20. *If H is of finite type then $\operatorname{co} K$ is compact whenever K is compact.*

Proof. If H is of finite type then $\deg_{\operatorname{thr}} H = n$ for some $n \in \mathbb{N}$. By definition then for any $S \subseteq H$ it holds $\deg_{\operatorname{thr}} S \leqslant n$. Now let $S = K$ be a compact set and let $n_K \leqslant n$ be its degree then $\operatorname{thr}^{n_K} K = \operatorname{co} K$ is compact by Lemma 4.19. □

Note that a Hilbert space $\mathcal{H}$ fails to be of finite type. If $(e_n)_{n\in\mathbb{N}}$ is the standard basis then $\deg_{\operatorname{thr}}\{e_1, ..., e_n\} = n$ (follows from Caratheodory's theorem see Example 1) implies $\deg_{\operatorname{thr}} \mathcal{H} \geqslant \sup_{n\in\mathbb{N}} \deg_{\operatorname{thr}}\{e_1, ..., e_n\} = +\infty$. Now consider the unbounded ray of real numbers with at point $x = n$ an n-dimensional cube attached (1-point union). In this example every compact set is of finite type but their threading degrees are not uniformly bounded. Consequently the space itself is not of finite type [1].

[1] These examples were pointed out by Prof. Thomas Schick.

Consider the limit $\lim_n \operatorname{thr}^n K$ when K is compact. By construction $\operatorname{thr}^{n-1} K \subseteq \operatorname{thr}^n K$ for any set $K \subseteq H$ and any $n \in \mathbb{N}$. On the other hand by Theorem 4.15 we have $\operatorname{co} K = \bigcup_{n\in\mathbb{N}} \operatorname{thr}^n K$. Since $K \subseteq \operatorname{thr} K \subseteq \dots \subseteq \operatorname{thr}^{n-1} K \subseteq \operatorname{thr}^n K \subseteq \dots$ is a sequence of nested nondecreasing sets then by Lemma 4.2 $\lim_n \operatorname{thr}^n K = \operatorname{cl} \bigcup_{n\in\mathbb{N}} \operatorname{thr}^n K$ or equivalently $\lim_n \operatorname{thr}^n K = \operatorname{cl}\operatorname{co} K$. In particular we obtain that when K is compact $\operatorname{cl}\operatorname{co} K$ can be successively approximated in the sense of Painlevé-Kuratowski by a sequence of nondecreasing compact sets. Let $(x_m)_{m\in\mathbb{N}} \subseteq \operatorname{cl}\operatorname{co} K$ then for each x_m there exists $y_m \in \operatorname{co} K$ such that $d(x_m, y_m) < 1/m$. In particular there is $n_m \in \mathbb{N}$ such that $y_m \in \operatorname{thr}^{n_m} K$. First suppose there is some $n \in \mathbb{N}$ such that $\operatorname{thr}^n K$ contains infinitely many terms of the sequence $(y_m)_{m\in\mathbb{N}}$. Let this subsequence be denoted by $(y_{m_k})_{k\in\mathbb{N}}$. By Lemma 4.18 $\operatorname{thr}^n K$ is compact then there is some convergent subsequence $(y_{m_{k_j}})_{j\in\mathbb{N}}$. Let $\lim_j y_{m_{k_j}} = y$. The estimate

$$d(x_{m_{k_j}}, y) \leqslant d(x_{m_{k_j}}, y_{m_{k_j}}) + d(y_{m_{k_j}}, y)$$

implies $\lim_j x_{m_{k_j}} = y$. Therefore $(x_{m_{k_j}})_{j\in\mathbb{N}}$ would be a convergent subsequence of the sequence $(x_m)_{m\in\mathbb{N}}$. The second case is when each $\operatorname{thr}^n K$ contains only finitely many terms of the sequence $(y_m)_{m\in\mathbb{N}}$. It is not clear now whether $(y_m)_{m\in\mathbb{N}}$ posseses a convergent subsequence unless some additional condition is added. From these arguments we make the following conclusion which we present without proof.

Proposition 4.21. *For a given set $K \subseteq H$ define*

$$Y_K := \{(y_m)_{m\in\mathbb{N}} \subseteq \operatorname{co} K : \exists n \in \mathbb{N} \quad \text{such that} \quad y_m \in \operatorname{thr}^n K \quad \text{for infinitely many} \quad m \in \mathbb{N}\}$$

If K is compact and $\operatorname{cl} Y_K = \operatorname{cl}\operatorname{co} K$ then $\operatorname{cl}\operatorname{co} K$ is compact.

Regarding the same topic Kopecká and Reich [74] have made the following observation.

Lemma 4.22. *If a Hadamard space (H, d) has the property that, given a finite set of elements $\{x_1, ..., x_n\} \subset H$, its closed convex hull $\operatorname{cl}\operatorname{co}\{x_1, ..., x_n\}$ is compact then for each compact set $K \subset H$ its closed convex hull $\operatorname{cl}\operatorname{co} K$ is compact.*

This lemma essentially reduces the problem to finite sets only. It suffices to study the threading operation only for finite sets.

4.2.3 Some illustrations

Example 1 (Euclidean spaces)**.** Euclidean spaces are the simplest examples of Hadamard spaces. Let $\mathbb{E}^d$ be a d-dimensional Euclidean space. By the well known Carathéodory Theorem if $x \in \operatorname{co} S$ for some $S \subseteq \mathbb{E}^d$ then x can be expressed as a convex combination of at most $d+1$ points from S. This means that there are $s_1, ..., s_{d+1} \in S$ such that $x \in \operatorname{co}\{s_1, ..., s_{d+1}\}$. Algebraically we have the representation $x = a_1 s_1 + ... + a_{d+1} s_{d+1}$ where $a_i \in [0, 1]$ for all $i = 1, ..., d+1$ and $\sum_i a_i = 1$. By letting $a' := \sum_{i=1}^d a_i$ then we can rewrite $x = a_{d+1} s_{d+1} + (1 - a_{d+1}) \sum_{i=1}^d a'_i s_i$ where $a'_i := a_i / a'$ for all $i = 1, 2, ..., d$.

Hence $x' := \sum_{i=1}^{d} a'_i s_i$ is some point lying in $\text{co}\{s_1, ..., s_d\}$. Following iteratively this method one obtains a sequence of points $x^{(1)}, x^{(2)}, x^{(3)}, ..., x^{(d)}$, where $x^{(1)} = x$, such that $x^{(1)} \in \text{co}\{s_1, ..., s_{d+1}\}, x^{(2)} \in \text{co}\{s_1, ..., s_d\}, ..., x^{(d)} \in \text{co}\{s_1, s_2\}$. By construction we have the following inclusions

$$x^{(d)} \in \text{thr}\{s_1, s_2\}, ..., x^{(2)} \in \text{thr}^{d-1}\{s_1, ..., s_d\}, x^{(1)} \in \text{thr}^{d}\{s_1, ..., s_{d+1}\}$$

By virtue of Carathéodory Theorem this means that $\text{co}\{s_1, ..., s_{d+1}\} \subseteq \text{thr}^d\{s_1, ..., s_{d+1}\}$. Because $x \in S$ is arbitrary then $\text{co}\, S \subseteq \text{thr}^d\, S$ and so $\deg_{\text{thr}} S \leqslant d$. On the other hand $S \subseteq \mathbb{E}^n$ is any subset hence $\deg_{\text{thr}} \mathbb{E}^d \leqslant d$. Therefore any Euclidean space is of finite type. In fact it holds $\deg_{\text{thr}} \mathbb{E}^d \leqslant d$. To see this take $d+1$-linearly independent points $\{s_1, ..., s_{d+1}\}$ in $\mathbb{E}^d$ then if $x \in \text{thr}^d\{s_1, ..., s_d\}$ then it can be easily checked that there are real numbers not all zero $a_1, ..., a_{d+1}$ such that $x = a_1 s_1 + ... a_{d+1} s_{d+1}$ i.e. $x \in \text{co}\{s_1, ..., s_{d+1}\}$. By earlier arguments this means that $\text{co}\{x_1, ..., x_{d+1}\} = \text{thr}^d\{x_1, ..., x_{d+1}\}$. Thus $\deg_{\text{thr}} \mathbb{E}^d = d$.

Example 2 (A non-Euclidean metric space)**.** Consider the set $H := \mathbb{R}^2_+ \cup \mathbb{R}^2_-$ where $\mathbb{R}^2_+ := \{x \in \mathbb{R}^2 : x_1, x_2 \geqslant 0\}$ and $\mathbb{R}^2_- := \{x \in \mathbb{R}^2 : x_1, x_2 \leqslant 0\}$. If H is equipped with a metric d' induced from the length of the shortest path connecting any two points in H then it can be shown that (H, d) is a Hadamard space. Now let $S \subseteq H$. If S is contained entirely in one of the two quadrants of H i.e. $S \subseteq \mathbb{R}^2_+$ or $S \subseteq \mathbb{R}^2_-$, then it is easily seen from the previous example that $\deg_{\text{thr}} S \leqslant 2$. Now suppose that $S = S_1 \cup S_2$ where $S_1 = S \cap \mathbb{R}^2_+$ and $S_2 = S \cap \mathbb{R}^2_-$. Notice that any point in S_1 can be joined with a point in S_2 by a shortest line (possibly broken) only passing through the origin $0 \in \mathbb{R}^2$ since $\mathbb{R}^2_+ \cap \mathbb{R}^2_- = \{0\}$. Therefore $\text{thr}(S_1 \cup \{0\}) \cup \text{thr}(S_2 \cup \{0\}) = \text{thr}\, S$. By same arguments $\text{thr}(S_1 \cup \{0\}) \cap \text{thr}(S_2 \cup \{0\}) = \{0\}$ implies

$$\text{thr}^2(S_1 \cup \{0\}) \cup \text{thr}^2(S_2 \cup \{0\})) = \text{thr}(\text{thr}(S_1 \cup \{0\}) \cup \text{thr}(S_2 \cup \{0\})) = \text{thr}^2 S$$

and in more generality

$$\text{thr}^n(S_1 \cup \{0\}) \cup \text{thr}^n(S_2 \cup \{0\})) = \text{thr}^n S, \ \ \forall n \in \mathbb{N} \tag{4.14}$$

But $S_1 \cup \{0\}$ and $S_2 \cup \{0\}$ are each subsets of Euclidean quadrants. Moreover the metric d' coincides with the usual Euclidean metric. By previous example $\deg_{\text{thr}}(S_i \cup \{0\}) \leqslant 2$ for $i = 1, 2$. By equation (4.14) it follows that $\deg_{\text{thr}} S \leqslant 2$. Since $S \subseteq H$ is arbitrary then $\deg_{\text{thr}} H \leqslant 2$. Therefore (H, d') is of finite type.

4.3. Fréchet Mean

4.3.1 A general convex optimization problem

Let (H, d) be a Hadamard space and $x_1, x_2, ..., x_n \in H$. For a given set of positive numbers $w_1, w_2, ..., w_n \in [0, 1]$ consider the optimization problem

$$\arg\min_{x \in H} F_p(x) \ \text{ where } \ F_p(x) := \sum_{i=1}^{n} w_i d(x, x_i)^p, \ \ p \in [1, +\infty) \tag{4.15}$$

The functional $F_p(x)$ is convex and continuous in x. When $p = 1$ then F_1 becomes the objective function in the *Fermat-Weber* problem for the optimal facility location. A minimizer of F_1 exists and it is known as the *median* of the points $x_1, x_2, ..., x_n$ with respect to the weights $w_1, w_2, ..., w_n$. When $p = 2$ then F_2 is the objective function in the *Fréchet mean* problem. Because $d(x, x_i)^2$ is strongly convex and continuous in x then F_2 has a unique minimizer which we denote by x^*. This unique minimizer is known as the *Fréchet mean* of the points $x_1, x_2, ..., x_n$ with respect to the weights $w_1, w_2, ..., w_n$.

Lemma 4.23. *Let (H, d) be a Hadamard space. The Fréchet mean of any finite set of points in H lies in the closure of the convex hull of the given set.*

Proof. Let $S := \{x_1, x_2, ..., x_n\}$ be some finite set in H. Denote by $\operatorname{cl}\operatorname{co} S$ the closure of the convex hull of S. Suppose that the Fréchet mean $x^* \notin \operatorname{co}\operatorname{cl} S$. Because $\operatorname{cl}\operatorname{co} S$ is closed and convex by construction then the metric projection $P_{\operatorname{cl}\operatorname{co} S}x^*$ of x^* onto $\operatorname{cl}\operatorname{co} S$ exists and it is unique. On the other hand by Corollary 2.10 we have the inequality

$$d(x^*, y)^2 \geqslant d(x^*, P_{\operatorname{cl}\operatorname{co} S}x^*)^2 + d(P_{\operatorname{cl}\operatorname{co} S}x^*, y)^2, \ \ \forall y \in \operatorname{cl}\operatorname{co} S$$

and in particular we must have

$$d(x^*, x_i)^2 \geqslant d(x^*, P_{\operatorname{cl}\operatorname{co} S}x^*)^2 + d(P_{\operatorname{cl}\operatorname{co} S}x^*, x_i)^2, \ \ \forall i = 1, 2, ..., n$$

Therefore we obtain the inequality

$$\sum_{i=1}^{n} w_i d(x^*, x_i)^2 \geqslant \sum_{i=1}^{n} w_i d(x^*, P_{\operatorname{cl}\operatorname{co} S}x^*)^2 + \sum_{i=1}^{n} w_i d(P_{\operatorname{cl}\operatorname{co} S}x^*, x_i)^2 > \sum_{i=1}^{n} w_i d(P_{\operatorname{cl}\operatorname{co} S}x^*, x_i)^2$$

which yields $F_2(P_{\operatorname{cl}\operatorname{co} S}x^*) < F_2(x^*)$. This is a contradiction. □

Corollary 4.24. *Similarly the median of any finite set of points in H lies in the closure of the convex hull of the given set.*

4.3.2 A constructability theorem

In an Euclidean space $\mathbb{E}$ if $x_1, x_2, ..., x_n \in \mathbb{E}$ and x^* denotes their Fréchet mean then it is well known that

$$x^* = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4.16}$$

In fact one gets to x^* iteratively from any starting point in $x_{i_1} \in \{x_1, ..., x_n\}$ in the following way

$$x'_1 = \frac{1}{2}x_{i_1} + \frac{1}{2}x_{i_2}, \quad x_{i_2} \in \{x_1, ..., x_n\} \setminus \{x_{i_1}\}$$
$$x'_2 = \frac{2}{3}x'_1 + \frac{1}{3}x_{i_3}, \quad x_{i_3} \in \{x_1, ..., x_n\} \setminus \{x_{i_1}, x_{i_2}\}$$
......
$$x'_k = \frac{k}{k+1}x'_{k-1} + \frac{1}{k+1}x_{i_{k+1}}, \quad x_{i_{k+1}} \in \{x_1, ..., x_n\} \setminus \{x_{i_1}, ..., x_{i_k}\}$$
......

In particular when $k = n-1$ we obtain

$$x'_{n-1} = \frac{n-1}{n} x'_{n-2} + \frac{1}{n} x_{i_n}$$

which by the means of the recursion is equivalent to formula

$$x'_{n-1} = \frac{1}{n} \sum_{j=1}^{n} x_{i_j}$$

This in turn is equivalent to (4.16). One way to think about this recursion is in terms of threading. Notice that x'_1 lies in the geodesic segment $[x_{i_1}, x_{i_2}]$ joining the elements x_{i_1} and x_{i_2} implying $x'_1 \in \operatorname{thr}\{x_1, x_2, ..., x_n\}$. And in general we have $x'_k \in [x'_{k-1}, x_{i_{k+1}}]$ equivalently $x'_k \in \operatorname{thr}^k\{x_1, x_2, ..., x_n\}$. In general given a set S of finitely many elements and $x \in H$ we say that x is *constructible* from S in n steps whenever $x \in \operatorname{thr}^n S$. Motivated by above recursion and taking advantage of threading we state the following result for Hadamard spaces of finite type.

Theorem 4.25. *Let (H, d) be of finite type and $S \subseteq H$ a finite subset. Then the Fréchet mean x^* of S lies in $\operatorname{co} S$. Moreover x^* is constructible from S in at most $2^{\deg_{\mathrm{thr}} S} - 1$ steps.*

Proof. Let $S = \{x_1, x_2, ..., x_n\}$ where $x_i \in H$ for all i. By Lemma 4.23 the Fréchet mean x^* of S lies in $\operatorname{cl}\operatorname{co} S$. Assumption (H, d) is of finite type implies that S has finite threading degree. If $\deg_{\mathrm{thr}} S = k$ for some $k \in \mathbb{N}$ then by definition this means that $\operatorname{thr}^k S = \operatorname{co} S$. But S is finite and therefore compact. By Lemma 4.18 it follows that $\operatorname{co} S$ is compact and therefore a closed set. Then $\operatorname{co} S = \operatorname{cl}\operatorname{co} S$ implies $x^* \in \operatorname{co} S$. Now $x \in \operatorname{thr}^k S$ means that there are $x_0^{k-1}, x_1^{k-1} \in \operatorname{thr}^{k-1} S$ such that $x \in [x_0^{k-1}, x_1^{k-1}]$. Then there are $x_0^{k-2}, x_1^{k-2}, x_2^{k-2}, x_3^{k-2} \in \operatorname{thr}^{k-2} S$ such that $x_0^{k-1} \in [x_0^{k-2}, x_1^{k-2}]$ and $x_1^{k-1} \in [x_2^{k-2}, x_3^{k-2}]$ and so on. In general for $0 \leqslant m \leqslant k$ there exist $x_0^{k-m}, x_1^{k-m}, ..., x_{2^m-1}^{k-m} \in \operatorname{thr}^{k-m} S$ such that $x_0^{k-m+1} \in [x_0^{k-m}, x_1^{k-m}], x_1^{k-m+1} \in [x_2^{k-m}, x_3^{k-m}], ..., x_{2^{m-1}-1}^{k-m+1} \in [x_{2^m-2}^{k-m}, x_{2^m-1}^{k-m}]$. Hence at each step m we need to compute at most 2^{m-1} elements. This means that in total we have to construct at most

$$\sum_{m=1}^{k} 2^{m-1} = 2^k - 1 = 2^{\deg_{\mathrm{thr}} S} - 1$$

□

Theorem 4.25 will become useful in the next chapter when we apply it to the problem of the average phylogenetic tree.

CHAPTER 5

The Space of Phylogenetic Trees

5.1. Description of Phylogenetic Tree Space

A phylogenetic tree is a diagram which describes the evolutionary relationships among a group of organisms. This diagram is an instance of *graphs* which are mathematical structures consisting of a set of vertices and edges. In a phylogenetic tree, there are no loops, each *interior* vertex has degree at least 3, i.e. an interior vertex connects to at least three other vertices, and there is a one-to-one labelling between the *leaves* (degree 1 vertices) and some set of labels. There are several methods of constructing a phylogenetic tree, from biochemists' methods of using quantitative estimates of variances between substances obtained from different species to the more systematic method of *mutation distances* introduced by Fitch and Margoliash [52]. In principle the mutation

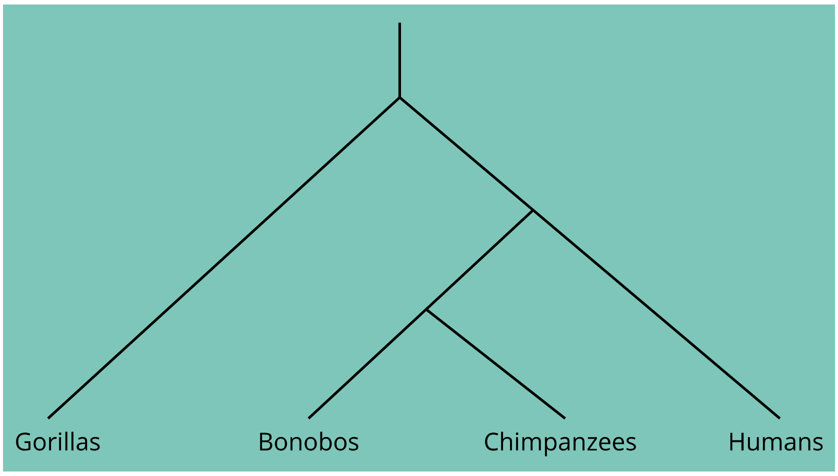


Figure 5.1: The relationship between humans and certain other members of primates.

distance between two DNA sequences is defined as the minimal number of nucleotides that need to be changed so that one DNA sequence transforms completely into the other DNA sequence. Figure 5.1 shows a simple phylogenetic tree about a group of primates based on data of their DNA sequences. According to this diagram it is understood that chimpanzees are closer to bonobos than humans are to bonobos. This means that it takes fewer number of nucleotides in the DNA of chimpanzees to be changed so that it coincides with the DNA sequence of bonobos, than it would take for the DNA sequence of humans to transform into the DNA of bonobos. Similarly using the principle of mutation distance we understand that chimpanzees are closer to humans than gorillas are to humans and so on. The method of mutation distances has more general applicability. Figure 5.2 shows the relationships between different languages in a restricted set of Proto-Indo-European languages based on linguistic data. Here instead of minimal number of nucleotides one looks at the minimal

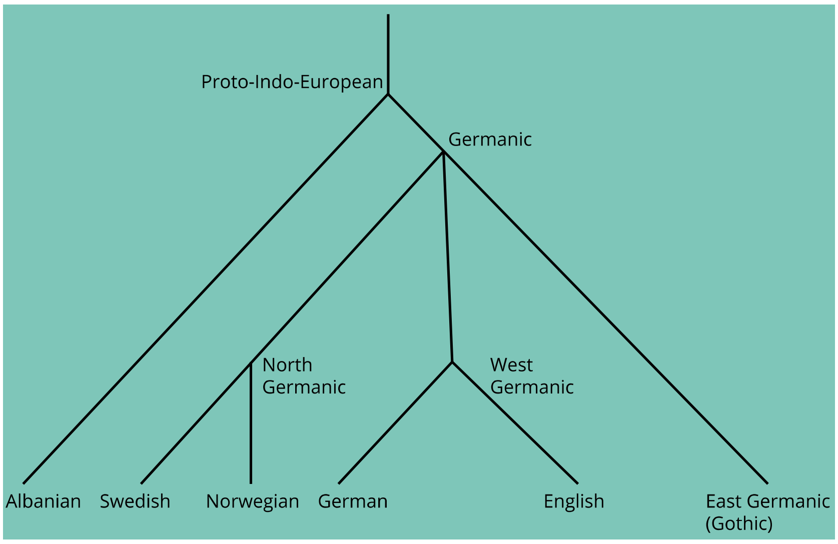


Figure 5.2: The relationship of the Albanian language to Germanic languages within the larger family of Proto-Indo-European languages.

number of certain elementary linguistic structures to be altered. We undertand that English language is closer to Swedish than it is to the Albanian language, though all the indicated languages share a common ancestor, a Proto-Indo-European language, for the complete tree see [30, Bouckaert et al.]. Because of the increasing amount of data on DNA sequences biologists and geneticists in particular use extensively statistical methods to estimate phylogenetic trees like the one depicted in Figure 5.1. Some of these statistical methods include but are not limited to parsimony methods, compatibility approach, least-squares approach, and maximum likelihood method (see Felsenstein [51] and references therein). However from a pure mathematical point of view, apart from these probabilistic approaches, the first successful mathematical space that describes phylogenetic trees was constructed by Billera, Holmes and Vogtmann [26]. They introduce a metric space that admits a natural metric of nonpositive curvature in the sense of Alexandrov. These spaces are known as the BHV spaces and constitute an interesting class of Hadamard spaces. They have been studied by several authors in the last two decades. In particular Owen has investigated BHV spaces from both theoretical and practical point of view with the purpose of studying the Fréchet mean of a finite set of trees [39], [20], [89] computing shortest paths between trees [93], [94], [6] and constructing convex hulls of finite sets of trees [81]. It turns out that BHV spaces and *cubical complexes* in general posses properties which are quite counter-intuitive in nature when compared with their counterparts in the Euclidean spaces, even though the former are essentially unions of Euclidean orthants arranged in a certain combinatorial way. For example Owen [81] has already pointed out that many algorithms and properties of shortest paths and convex hulls in a Euclidean space fail to transfer over to BHV spaces. It is no longer true that for a given finite set of points S: (i) any point on the boundary of the convex hull of S in 2D is on a shortest path between two points of S; (ii) the convex hull of three points is 2-dimensional; (iii) any point inside the convex hull of S can be written as a convex combination of points of S. It is not even known whether the convex hull of S is a closed set. Motivated by these interesting observations our goal in this chapter is to investigate convex hulls and the Fréchet mean. Before doing so we provide an elementary description of the phylogenetic tree space. We denote by $\mathcal{T}_n$ the BHV space describing the set of phylogenetic trees with n leaves. The space $\mathcal{T}_n$ can be described as the union of certain $n-2$ dimensional orthants all lying in a higher dimensional Euclidean space. Each such orthant describes a certain *tree topology* which is determined by the way the inner edges of a given tree *split* the set of leaves. Hence two phylogenetic trees are topologically indistinguishable if and only if their set of splits coincide. If $L_n := \{0, 1, 2, ..., n\}$ denotes the set of leaves (0 corresponds to the root of the tree) then each inner edge of a tree splits the set L_n into two sets, each with at least two elements. There is a one-to-one correspondence between splits and the

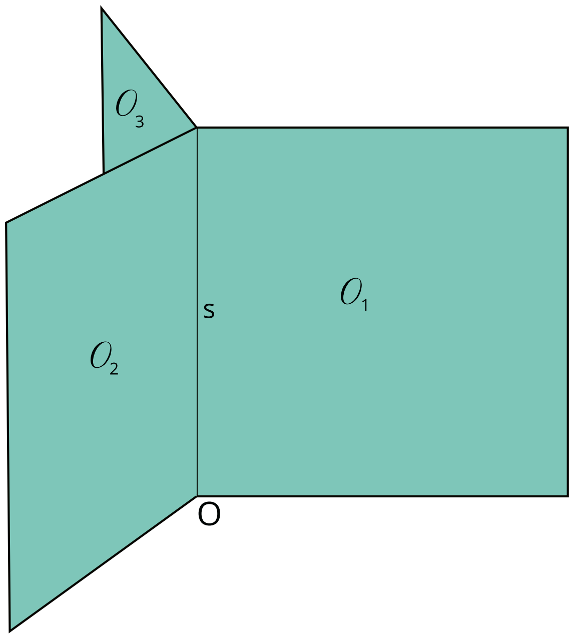


Figure 5.3: Three orthants $\mathcal{O}_1, \mathcal{O}_2$ and $\mathcal{O}_3$ in $\mathcal{T}_4$ sharing a common split s.

set of inner edges of a tree. Let s_1, s_2 be two splits and $\{X_1, Y_1\}, \{X_2, Y_2\}$ their respective partitions of L_n. We say s_1 and s_2 are *compatible* if one of the subsets

$$X_1 \cap X_2, \;\; X_1 \cap Y_2, \;\; X_2 \cap Y_1, \;\; Y_1 \cap Y_2$$

is empty. It is equivalent to say that compatible splits correspond to inner edges of a common tree. Every tree with n-leaves has at most $n - 2$ inner edges and therefore

$n-2$ splits. If one fixes the topology then the axis of the orthant corresponding to this topology represents the splits of the trees having this topology. The length of the inner edges determine how far any given tree in this topology lies along each axis. Two orthants are then glued together along a *boundary face* if and only if the topology of one orthant shares a split or more with the topology of the other orthant (see Figure 5.3 for an example of three orthants in $\mathcal{T}_4$ sharing a common split). If two topologies share no common split then their respective orthants are glued at the origin O. An orthant $\mathcal{O}$ in $\mathcal{T}_n$ is said to be a *maximal* orthant if it has dimension $n-2$. To each orthant $\mathcal{O}$ in $\mathcal{T}_n$ we denote the set of its splits by $\Sigma(\mathcal{O})$. Analogously for any tree $x \in \mathcal{T}_n$ we let $\Sigma(x)$ denote the set of splits of x. It follows in particular that $x \in \mathcal{O}$ if and only if $\Sigma(x) \subseteq \Sigma(\mathcal{O})$. For any $E \subseteq \Sigma(\mathcal{O}_i), F \subseteq \Sigma(\mathcal{O}_j)$ we say E is compatible with F whenever each element of E is compatible with each element of F. Clearly compatibilty is a symmetric relation.

5.2. Preliminaries

5.2.1 Some definitions and notations

By construction the phylogenetic tree space is a collection of maximal orthants glued together along common faces (orthants of lower dimension). Moreover this structure has a common vertex O corresponding to the phylogenetic tree with no inner edges. This special tree is called the *star tree*. Any two trees are connected by a geodesic segment consisting of a finite number of *Euclidean segments*. Notice that any two trees can be connected by a path which passes through the origin O. We call such a path the *cone path* between the two given trees. The cone path may or may not be a geodesic. Following Owen [93] given two trees $x, y \in \mathcal{T}_n$ let $\emptyset = E_k \subset E_{k-1} \subset ... \subset E_1 \subset E_0 = \Sigma(x)$ and $\emptyset = F_0 \subset F_1 \subset ... \subset F_{k-1} \subset F_k = \Sigma(y)$ be such that E_i and F_i are compatible for every $i = 0, 1, ..., k$. Note that for obvious reasons the index k cannot exceed the number of elements in either $\Sigma(x)$ or $\Sigma(y)$. Denote by $\mathcal{O}_i := \text{span}^+\{E_i \cup F_i\}$ the orthant in the *ambient* Euclidean space spanned by the set of splits $E_i \cup F_i$. The collection $\{\mathcal{O}_i\}_{i=0}^k$ is said to be a *path space* between x and y. By construction it follows that $\mathcal{O}_i \cap \mathcal{O}_{i+1} = \text{span}^+\{E_{i+1} \cup F_i\}$. When traversing from orthant $\mathcal{O}_i$ to $\mathcal{O}_{i+1}$ a tree with splits $E_i \cup F_i$ transforms into a tree with splits $E_{i+1} \cup F_{i+1}$. This means at the $i+1$-th step one removes the splits $A_{i+1} := E_i \setminus E_{i+1}$ and adds the splits $B_{i+1} := F_{i+1} \setminus F_i$. Therefore equivalently one can express each orthant $\mathcal{O}_i$ in the path space $\bigcup_i \mathcal{O}_i$ as follows

$$\mathcal{O}_i = B_1 \cup B_2 \cup ... \cup B_i \cup A_{i+1} \cup ... \cup A_k \tag{5.1}$$

If $\Sigma(\mathcal{T}_n)$ denotes the set of all splits in $\mathcal{T}_n$ then a subset $A \subseteq \Sigma(\mathcal{T}_n)$ is called a *set of mutually compatible splits* if any two splits in A are compatible. Let $A, B \subseteq \mathcal{T}_n$ be two sets of mutually compatible splits such that $A \cap B = \emptyset$ then $C_B(A)$ is defined to be the set of splits in B that are compatible with all the splits in A. The set $C_B(A)$ is known as the *compatibility set* of A in B. Define the *crossing set* of A in B, denoted by $X_B(A)$, as the set of splits in B which are not compatible with at least one split in A.

Proposition 5.1. *Let A, B be two sets of mutually compatible splits. If $D \subseteq A$ then $C_B(A) \subseteq C_B(D)$ and $X_B(D) \subseteq X_B(A)$. Moreover $C_B(A)$ and $X_B(A)$ partition the set B.*

Proof. Let $s \in C_B(A)$ then by definition s is compatible with any $s_A \in A$. But $D \subseteq A$ implies that s is compatible with any $s_D \in D$ therefore $s \in C_B(D)$. If $s \in X_B(D)$ then there is $s_D \in D$ such that s and s_D are not compatible. But $D \subseteq A$ implies $s_D \in A$ hence $s \in X_B(A)$. It is clear that the sets $C_B(A)$ and $X_B(A)$ are disjoint by definition and any element in B must be in either of these two sets but never in both of them. Therefore B is the disjoint union of $C_B(A)$ and $X_B(A)$. □

Proposition 5.2. *Let $\emptyset = E_k \subset E_{k-1} \subset ... \subset E_1 \subset E_0 = \Sigma(x)$ and $\emptyset = F_0 \subset F_1 \subset ... \subset F_{k-1} \subset F_k = \Sigma(y)$ be such that E_i and F_i are compatible for every $i = 0, 1, ..., k$. Let $\bigcup_i \mathcal{O}_i$ be the path space between $x, y \in \mathcal{T}_n$ where $\mathcal{O}_i := \text{span}^+\{E_i \cup F_i\}$. Then $E_i \subseteq C_{\Sigma(x)}(F_i)$ and $F_i \subseteq C_{\Sigma(y)}(E_i)$ for all $i = 0, 1, ..., k$.*

Proof. By construction E_i and F_i are compatible for every $i = 0, 1, .., k$. It follows that $C_{E_i}(F_i) = E_i$ for all i. Therefore if $s \in C_{E_i}(F_i)$ then $s \in C_{\Sigma(x)}(F_i)$ since $E_i \subseteq \Sigma(x)$ for every $i = 0, 1, ..., k$. Analogously we can show that $F_i \subseteq C_{\Sigma(y)}(E_i)$ for all $i = 0, 1, ..., k$. □

Proposition 5.3 ([26, Proposition 4.1])**.** *Let $x, y \in \mathcal{T}_n$ such that their cone path is not a geodesic. Then there are nonempty sets $E_1 \supset E_2 \supset ... \supset E_k$ of $\Sigma(x)$, and $F_1 \subset F_2 \subset ... \subset F_k$ of $\Sigma(y)$ such that E_i is compatible with F_i and if $\mathcal{O}_i := \text{span}^+\{E_i \cup F_i\}, \forall i = 1, ..., k$ then the geodesic $[x, y]$ traverses each orthant in the order $\mathcal{O}_1, ..., \mathcal{O}_k$.*

A path space is *maximal* if it is not contained in any other path space. It follows from Proposition 5.3 that a geodesic segment between two trees is contained in a path space. Moreover this path space needs to be a maximal path space.

Theorem 5.4 ([93, Theorem 3.6])**.** *Let $x, y \in \mathcal{T}_n$ be two trees. The maximal path spaces $\bigcup_i \mathcal{O}_i$ between x and y satisfy $E_i = C_{\Sigma(x)}(F_i)$ and $F_i = C_{\Sigma(y)}(E_i)$ for all $i = 0, 1, ..., k$.*

Let $P := \bigcup_{i=0}^k \mathcal{O}_i$ be a path space between two trees $x, y \in \mathcal{T}_n$ and denote by $P(x, y)$ the set of all path spaces between x and y. A *path space geodesic between x and y through P* is the shortest path between x and y that is contained in P. For $0 \leqslant i \leqslant k$ define the orthants

$$V_i := \{(x_1, x_2, ..., x_k) \in \mathbb{R}^k : x_j \leqslant 0 \text{ if } j \leqslant i \text{ and } x_j \geqslant 0 \text{ if } j > i\} \tag{5.2}$$

and denote $V(\mathbb{R}^k) := \bigcup_{i=0}^k V_i$.

Theorem 5.5 ([93, Theorem 4.4])**.** *Let $P \in P(x, y)$ then the path space geodesic between x and y through P is contained in a space isometric to $V(\mathbb{R}^k)$.*

5.2.2 Isometry results

Consider two orthants $\mathcal{O}$ and $\mathcal{O}'$ not necessarily maximal. In the same way as for two distinct trees we can define a path space between orthants $\mathcal{O}$ and $\mathcal{O}'$. Given nonempty subsets $\emptyset = E_k \subset E_{k-1} \subset ... \subset E_1 \subset E_0 = \Sigma(\mathcal{O})$ and $\emptyset = F_0 \subset F_1 \subset ... \subset F_{k-1} \subset F_k = \Sigma(\mathcal{O}')$ such that E_i is compatible with F_i for all $i = 0, 1, ..., k$ then the collection $\{\mathcal{O}_i\}_i$ where $\mathcal{O}_i := \text{span}^+\{E_i \cup F_i\}$ is called an *orthant path space* between $\mathcal{O}$ and $\mathcal{O}'$. Let $P(\mathcal{O}, \mathcal{O}')$ denote the set of orthant path spaces between $\mathcal{O}$ and $\mathcal{O}'$. Since $\Sigma(x) \subseteq \Sigma(\mathcal{O})$ for any $x \in \mathcal{O}$ then any path space between $x \in \mathcal{O}$ and $y \in \mathcal{O}'$ is contained in some orthant path space in $P(\mathcal{O}, \mathcal{O}')$. Define

$$U(\mathcal{O}, \mathcal{O}') := \bigcup_{P(\mathcal{O},\mathcal{O}')} \bigcup_i \mathcal{O}_i \tag{5.3}$$

to be the collection of all possible orthant path spaces between $\mathcal{O}$ and $\mathcal{O}'$.

Theorem 5.6. *Let $\mathcal{O}, \mathcal{O}' \subset \mathcal{T}_n$ be two distinct orthants, not necessarily maximal. Then $U(\mathcal{O}, \mathcal{O}')$ is a closed convex set, in particular it is a Hadamard space.*

Proof. Note that $\Sigma(\mathcal{O})$ is finite for any orthant $\mathcal{O} \subset \mathcal{T}_n$ implies that the power set $\mathcal{P}(\Sigma(\mathcal{O}))$ is finite. In particular the number of possible chains $\{E_i\}$ is also finite and hence for any two orthants $\mathcal{O}$ and $\mathcal{O}'$ there are at most a finite number of distinct orthant path spaces between them, in particular the set $P(\mathcal{O}, \mathcal{O}')$ is finite. Definition (5.3) implies then that the set $U(\mathcal{O}, \mathcal{O}')$ is a union of finitely many orthants in $\mathcal{T}_n$. Because each orthant in $\mathcal{T}_n$ is a closed set then $U(\mathcal{O}, \mathcal{O}')$ is also closed. Now let $x, y \in U(\mathcal{O}, \mathcal{O}')$. By definition (5.3) there are $\mathcal{O}_i, \mathcal{O}_j \in U(\mathcal{O}, \mathcal{O}')$, possibly from different orthant path spaces, such that $x \in \mathcal{O}_i$ and $y \in \mathcal{O}_j$. If $\mathcal{O}_i \subseteq \mathcal{O}_j$ then clearly $[x, y] \subset \mathcal{O}_j$ since any orthant is convex and thus $[x, y] \subset U(\mathcal{O}, \mathcal{O}')$. So without loss of generality assume that $\mathcal{O}_i$ and $\mathcal{O}_j$ are distinct and one is not a subset of the other. By Proposition 5.3 the geodesic segment $[x, y]$ is contained in a path space between x and y. Since any path space between any point $x \in \mathcal{O}_i$ and any point $y \in \mathcal{O}_j$ is contained in some orthant path space in $P(\mathcal{O}_i, \mathcal{O}_j)$ it suffices to show that any element in $P(\mathcal{O}_i, \mathcal{O}_j)$ is part of some element in $P(\mathcal{O}, \mathcal{O}')$. By definition there must be $E_i, E_j \subset \Sigma(\mathcal{O})$ and $F_i, F_j \subset \Sigma(\mathcal{O}')$ such that E_i is compatible with F_i and E_j is compatible with F_j, and moreover $\mathcal{O}_i = \text{span}^+\{E_i \cup F_i\}$ and $\mathcal{O}_j = \text{span}^+\{E_j \cup F_j\}$. By construction we have $\Sigma(\mathcal{O}_i) = E_i \cup F_i$ and $\Sigma(\mathcal{O}_j) = E_j \cup F_j$. It is evident that E_i is compatible with E_j since both $E_i, E_j \subset \Sigma(\mathcal{O})$, likewise F_i is compatible with F_j since both $F_i, F_j \subset \Sigma(\mathcal{O}')$. Let $\emptyset = E'_k \subset E'_{k-1} \subset ... \subset E'_1 \subset E'_0 = \Sigma(\mathcal{O}_i)$ and $\emptyset = F'_0 \subset F'_1 \subset ... \subset F'_{k-1} \subset F'_k = \Sigma(\mathcal{O}_j)$ be such that E'_l is compatible with F'_l for every $l = 0, 1, ..., k$. By construction $E'_l \subset E_i \cup F_i$ and $F'_l \subset E_j \cup F_j$. Therefore $E'_l \cup F'_l \subset (E_i \cup E_j) \cup (F_i \cup F_j)$ for all $l = 0, 1, .., k$. Then there are subsets $E''_l \subset E_i \cup E_j$ and $F''_l \subset F_i \cup F_j$ such that $E'_l \cup F'_l = E''_l \cup F''_l$ for every $l = 0, 1, ..., k$. Hence the path space $\{\text{span}^+\{E'_l \cup F'_l\}\}_{l=0}^k$ is essentially identical to the path space $\{\text{span}^+\{E''_l \cup F''_l\}\}_{l=0}^k$. On the other hand by construction $E''_l \subset \Sigma(\mathcal{O})$ and $F''_l \subset \Sigma(\mathcal{O}')$ for all $l = 0, 1, ..., k$ implies that $\{\text{span}^+\{E''_l \cup F''_l\}\}_{l=0}^k$ would be part of some orthant path space between $\mathcal{O}$ and $\mathcal{O}'$. The inclusions $[x, y] \subset \bigcup_{l=0}^k \text{span}^+\{E''_l \cup F''_l\}$ and $\bigcup_{l=0}^k \text{span}^+\{E''_l \cup F''_l\} \subset \bigcup_i \mathcal{O}_i \in P(\mathcal{O}, \mathcal{O}')$ imply that $[x, y] \subset U(\mathcal{O}, \mathcal{O}')$. Therefore $U(\mathcal{O}, \mathcal{O}')$ is a convex set. Since $(\mathcal{T}_n, d)$ is a Hadamard space then $U(\mathcal{O}, \mathcal{O}')$ is also a Hadamard space with the restricted metric $d_U = d|_U$. □

Let I and I' be the index sets associated with splits in $\Sigma(\mathcal{O})$ and $\Sigma(\mathcal{O}')$ respectively. For $E_i \subset \Sigma(\mathcal{O})$ denote by $I(E_i) \subseteq I$ the set of indices corresponding to elements of E_i. Similarly denote by $I'(F_i)$ for any $F_i \subseteq \Sigma(\mathcal{O}')$. For two compatible set of splits $E_i \subseteq \Sigma(\mathcal{O})$ and $F_i \subseteq \Sigma(\mathcal{O}')$ in $\mathbb{R}^{n-2}$ define the orthant

$$V(\mathcal{O}_i) := \{(x_1, x_2, ..., x_{k_i}) \in \mathbb{R}^{k_i} | x_j \geqslant 0 \text{ if } j \in I(E_i) \text{ and } x_j \leqslant 0 \text{ if } j \in I'(F_i)\}, \tag{5.4}$$

where $k_i := |E_i| + |F_i|$ and $\mathcal{O}_i := \text{span}^+\{E_i \cup F_i\}$. For given two orthants $\mathcal{O}, \mathcal{O}' \subset \mathcal{T}_n$ not necessarily maximal define the following set

$$V(\mathcal{O}, \mathcal{O}') := \bigcup_{P(\mathcal{O},\mathcal{O}')} \bigcup_i V(\mathcal{O}_i) \tag{5.5}$$

Theorem 5.7. *Equip $V(\mathcal{O}, \mathcal{O}')$ with its canonical length metric. Then $U(\mathcal{O}, \mathcal{O}')$ is isometric isomomorphic to $V(\mathcal{O}, \mathcal{O}')$. In particular $V(\mathcal{O}, \mathcal{O}')$ is a Hadamard space.*

Proof. The idea of the proof is similar to the one in Theorem 5.5 though the isometry we use is in a slightly different form. Let $\{s_j\}$ and $\{s'_j\}$ denote the set of splits in $\mathcal{O}$ and $\mathcal{O}'$ respectively. For $x \in U(\mathcal{O}, \mathcal{O}')$ define

$$\Phi(x) := \sum_{j\in I} a_j e_j - \sum_{j\in I'\setminus I} a_j e_j, \;\; \{a_j\} \geqslant 0 \tag{5.6}$$

where $\{e_j\}$ is the standard basis of $\mathbb{R}^{n-2}$. If $\Phi(x) = \Phi(y)$, where say $y = \sum_{j\in I} b_j s_j + \sum_{j\in I'\setminus I} b_j s'_j$ for some $\{b_j\} \geqslant 0$, then we have

$$\sum_{j\in I} a_j e_j - \sum_{j\in I'\setminus I} a_j e_j = \sum_{j\in I} b_j e_j - \sum_{j\in I'\setminus I} b_j e_j$$

implying that

$$\sum_{j\in I} (a_j - b_j) e_j - \sum_{j\in I'\setminus I} (a_j - b_j) e_j = 0$$

which in turn yields $a_j = b_j$ for all $j \in I \cup I'$. Hence $x = y$ and the mapping Φ is injective. Let $x^+ \in V(\mathcal{O}, \mathcal{O}')$ then by definition (5.5) $x^+ \in V(\mathcal{O}_i)$ for some $\mathcal{O}_i$. Therefore there is $\{a_j\} \geqslant 0$ such that $x^+ = \sum_{j\in I(E_i)} a_j e_j - \sum_{j\in I'(F_i)\setminus I(E_i)} a_j e_j$ where $\mathcal{O}_i = \text{span}^+\{E_i \cup F_i\}$. Then the mapping

$$\Psi(x^+) := \sum_{j\in I(E_i)} a_j s_j + \sum_{j\in I'(F_i)\setminus I(E_i)} a_j s'_j \tag{5.7}$$

is well defined for any $x^+ \in V(\mathcal{O}, \mathcal{O}')$. If $x = \sum_{j\in I(E_i)} a_j s_j + \sum_{j\in I'(F_i)\setminus I(E_i)} a_j s'_j$ then this means that for any element $x^+ \in V(\mathcal{O}, \mathcal{O}')$ there is $x \in U(\mathcal{O}, \mathcal{O}')$ such that $\Phi(x) = x^+$ and $\Psi(x^+) = x$. Therefore not only Φ is surjective but also Φ and Ψ are inverses of each other. By construction we have $\Phi(s_j) = e_j$ for any $j \in I$ and $\Phi(s'_j) = -e_j$ for any $j \in I'$. Hence each element of the *basis* $\{s_j\}_{j\in I(E_i)} \cup \{s'_j\}_{j\in I'(F_i)\setminus I(E_i)}$ of $\mathcal{O}_i$ is mapped to a unique element of the basis $\{e_j\}_{j\in I(E_i)} \cup \{-e_j\}_{j\in I'(F_i)\setminus I(E_i)}$ of $V(\mathcal{O}_i)$. Therefore Φ is a *linear* transformation whose matrix is the identity matrix. Since the identity matrix has determinant one then Φ must be an isometry. Analogously $\Psi = \Phi^{-1}$ is an isometry. On the other hand the isometric image of a Hadamard space is a Hadamard space which in turn implies that $V(\mathcal{O}, \mathcal{O}')$ is Hadamard. This completes the proof.

□

As an illustration of the Theorem 5.7 we have the following result.

Proposition 5.8. *Let $\mathcal{O}$ and $\mathcal{O}'$ be two orthants in $\mathcal{T}_n$ such that no split in $\Sigma(\mathcal{O})$ is compatible with any split in $\Sigma(\mathcal{O}')$ then $\mathcal{O} \cup \mathcal{O}'$ is isometricly isomorphic to $\mathbb{R}^k_+ \cup \mathbb{R}^m_-$ equipped with its canonical length metric where $k := |\Sigma(x)|$ and $m := |\Sigma(y)|$. In particular when $\mathcal{O}$ and $\mathcal{O}'$ are maximal then $\mathcal{O} \cup \mathcal{O}'$ is isometricly isomorphic to $\mathbb{R}^{n-2}_+ \cup \mathbb{R}^{n-2}_-$.*

Proof. Let $\mathcal{O} := \text{span}^+\{s_1, s_2, ..., s_k\}$ and $\mathcal{O}' := \text{span}^+\{s'_1, s'_2, ..., s'_m\}$. Clearly $\Sigma(\mathcal{O}) \cap \Sigma(\mathcal{O}') = \emptyset$ since by assumption no split in $\Sigma(\mathcal{O})$ is compatible with any split in $\Sigma(\mathcal{O}')$. Then $s_i \neq s'_j$ for any $i, j \in \{1, 2, ..., k\} \times \{1, 2, ..., m\}$. From definition (5.3) it follows that $U(\mathcal{O}, \mathcal{O}') = \mathcal{O} \cup \mathcal{O}'$. Let $x \in \mathcal{O}$ and $y \in \mathcal{O}'$ where $x = \sum_{i=1}^k a_i s_i$ and $y = \sum_{i=1}^m b_i s'_i$ for some $\{a_i\}, \{b_i\} \geqslant 0$. Since $\Sigma(x) \subseteq \Sigma(\mathcal{O})$ and $\Sigma(y) \subseteq \Sigma(\mathcal{O}')$ then no split in $\Sigma(x)$ is compatible with any split in $\Sigma(y)$. By Proposition 5.3 it follows that the cone path is a geodesic i.e. $[x, y] = [x, O] \cup [O, y]$ and

$$d(x, y) = d(x, O) + d(O, y) = |xO| + |Oy| = \Big(\sum_{i=1}^k a_i^2\Big)^{1/2} + \Big(\sum_{i=1}^m b_i^2\Big)^{1/2} \tag{5.8}$$

On the other hand in view of Theorem 5.7 the isometry Φ reduces to the following mapping

$$\Phi(x) := \begin{cases} \sum_{i=1}^k a_i e_i & x \in \mathcal{O} \\ -\sum_{i=1}^m a_i e_i & x \in \mathcal{O}' \end{cases} \tag{5.9}$$

where $\{e_i\}_{i=1}^{n-2}$ is the standard basis in $\mathbb{R}^{n-2}$. Evidently Φ is a mapping from $\mathcal{O} \cup \mathcal{O}'$ onto $\mathbb{R}^k_+ \cup \mathbb{R}^m_-$. Since $\mathbb{R}^k_+ \cap \mathbb{R}^m_- = \{0\}$ then any point in $\mathbb{R}^k_+$ is connected with a point in $\mathbb{R}^m_-$ by a segment passing through the origin 0. If $d'(\cdot, \cdot)$ denotes the metric of $\mathbb{R}^k_+ \cup \mathbb{R}^m_-$ then by definition of Φ in (5.9) we have

$$d'(\Phi(x), \Phi(y)) = d'(\Phi(x), 0) + d'(0, \Phi(y)) = |\Phi(x)0| + |0\Phi(y)| = \Big(\sum_{i=1}^k a_i^2\Big)^{1/2} + \Big(\sum_{i=1}^m (-b_i)^2\Big)^{1/2} \tag{5.10}$$

It is clear from (5.8) and (5.10) that Φ is an isometry. Moreover the inverse mapping Φ^{-1} exists and it is given by

$$\Phi^{-1}(x) := \begin{cases} \sum_{i=1}^k a_i s_i & x \in \mathbb{R}^k_+ \\ \sum_{i=1}^m a_i s'_i & x \in \mathbb{R}^m_- \end{cases} \tag{5.11}$$

where $\{a_i\} \geqslant 0$. Using similar arguments as above one can show that (5.11) is also an isometry. The second assertion is clear. This completes the proof. □

5.3. The Example of Trees with Four Leaves

5.3.1 Certain lemmas

Let $\mathcal{T}_4$ be the space of phylogenetic trees with four leaves. Since there are at most two inner edges for any tree in $\mathcal{T}_4$ it follows that the set of splits for any tree has at most two elements. Moreover in $\mathcal{T}_4$ there are 10 distinct splits which when combined give rise to 15 different tree topologies therefore 15 different orthants. They are all glued according to the rule described in the first section and no orthant is left loose. The trees having only one inner edge lie in the axis of these orthants and the tree with no inner edge (also known as the *star stree*) is indentified with the origin O. By virtue of Proposition 5.3 the geodesic segment $[x, y]$ between any two points $x, y \in \mathcal{T}_4$ traverses at most one maximal orthant. To see this assume for simplicity that both x and y have *full set of splits*, i.e. $x \in \mathcal{O}, y \in \mathcal{O}'$ where $\mathcal{O}, \mathcal{O}'$ are maximal and $\Sigma(x) = \Sigma(\mathcal{O}), \Sigma(y) = \Sigma(\mathcal{O}')$. Say $\Sigma(\mathcal{O}) = \{s_1, s_2\}$ and $\Sigma(\mathcal{O}') = \{s'_1, s'_2\}$. Then the power sets $\mathcal{P}(\Sigma(x)) = \{\{\emptyset\}, \{s_1\}, \{s_2\}, \{s_1, s_2\}\}$ and $\mathcal{P}(\Sigma(y)) = \{\{\emptyset\}, \{s'_1\}, \{s'_2\}, \{s'_1, s'_2\}\}$ imply that the only possible path spaces between x and y would be $\{\mathcal{O}, \mathcal{O}'\}, \{\mathcal{O}, \{s_i\}, \mathcal{O}'\}, \{\mathcal{O}, \text{span}^+\{s_i, s'_j\}, \mathcal{O}'\}$ for $i, j = 1, 2$. In any case at most one maximal orthant, that is $\text{span}^+\{s_i, s'_j\}$, can be traversed by the geodesic $[x, y]$. A direct consequence of Theorem 5.7 and Proposition 5.8 are the following corollaries.

Corollary 5.9. *Let $\mathcal{O}$ and $\mathcal{O}'$ be two distinct maximal orthants in $\mathcal{T}_4$ such that $\Sigma(\mathcal{O}) \cap \Sigma(\mathcal{O}') = \emptyset$ but some element in $\Sigma(\mathcal{O})$ is compatible with some element in $\Sigma(\mathcal{O}')$. Then $U(\mathcal{O}, \mathcal{O}')$ is isometric isomomorphic to $V(\mathbb{R}^2)$.*

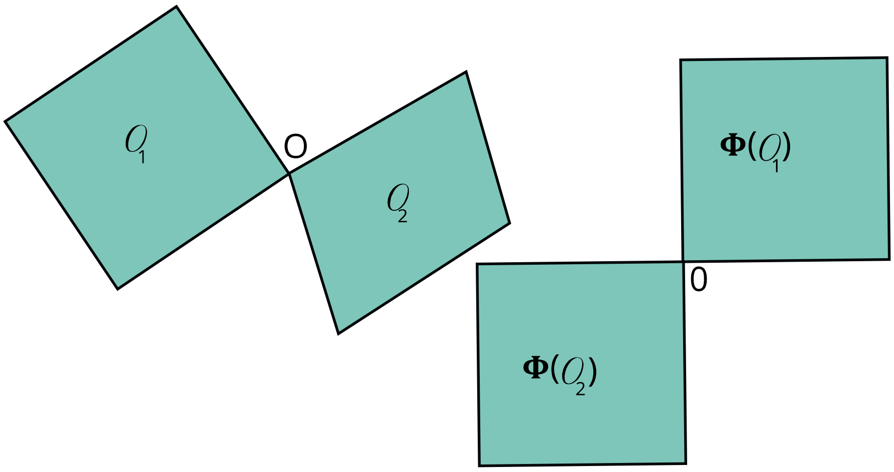


Figure 5.4: Two orthants in $\mathcal{T}_4$ with no compatible splits in general position (left) and their isometric image $\mathbb{R}^2_+ \cup \mathbb{R}^2_-$ (right).

Corollary 5.10. *Let $\mathcal{O}_i$ and $\mathcal{O}_j$ be two distinct orthants in $\mathcal{T}_4$ such that no split in $\Sigma(\mathcal{O}_i)$ is compatible with any split in $\Sigma(\mathcal{O}_j)$. Then $\mathcal{O}_i \cup \mathcal{O}_j$ is isometric isomomorphic to $\mathbb{R}^2_+ \cup \mathbb{R}^2_-$ (see Figure 5.4).*

Lemma 5.11. *Any compact set* $S \subset V(\mathbb{R}^2)$ *satisfies* $\deg_{\text{thr}} S \leqslant 3$.

Proof. Let $S \subseteq V(\mathbb{R}^2)$ be a compact set. By construction $V(\mathbb{R}^2)$ consists of three quadrants in $\mathbb{R}^2$. Write $S_1 = S \cap \mathbb{R}^2_{++}, S_2 = S \cap \mathbb{R}^2_{-+}$ and $S_3 = S \cap \mathbb{R}^2_{--}$. Here the notation $\{-+\}$ means the first coordinate is nonpositive and the second is nonnegative. Analogously for the other two notations. Without loss of generality assume that all S_i are nonempty (other cases follow similarly).

Case 1: Let l be a line in $\mathbb{R}^2 \cap V(\mathbb{R}^2)$ supported at the origin 0. If S is bounded below by l then S lies in a two dimensional Euclidean half-space. There the intrinsic length metric d' coincides with the Euclidean metric, it follows from Example 1 that $\deg_{\text{thr}} S \leqslant 2$.

Case 2: Denote by l_0 the line supported at 0 such that l_0 coincides with the horizontal axis in $\mathbb{R}^2$. If no line l in $\mathbb{R}^2 \cap V(\mathbb{R}^2)$ supported at origin 0 exists such that it bounds from one side the entire set S then there exist points $x, y \in S$ such that the geodesic segment in $V(\mathbb{R}^2)$ joining x with y goes through the origin 0. This means $0 \in \text{thr}\, S$. There are $x_1, x_2 \in S$ such that the segment $[x_1, x_2]$ intersects l_0 at some point y such that $d'(0, y) \geqslant d'(0, y')$ for all $y' \in \text{thr}\, S \cap l_0$ where $d'(\cdot, \cdot)$ is the intrinsic length metric in $V(\mathbb{R}^2)$. This is evident by the continuity of the metric function $d'(0, \cdot)$ since $\text{thr}\, S \cap l_0$ (because S is compact). Define the sets

$$U(S) := (\text{thr}\, S \cap \mathbb{R}^2_{--}) \cup [0, y] \tag{5.12}$$

$$V(S) := (\text{thr}\, S \cap (\mathbb{R}^2_{-+} \cup \mathbb{R}^2_{++})) \cup [0, y] \tag{5.13}$$

We claim that

$$\text{thr}^2 S = \text{thr}\, U \cup \text{thr}\, V \tag{5.14}$$

Notice that by construction $U \cup V = \text{thr}\, S \cup [0, y]$. By the union rule (4.10) it follows $\text{thr}\, U \cup \text{thr}\, V \subseteq \text{thr}(U \cup V) = \text{thr}(\text{thr}\, S \cup [0, y])$. Moreover $[0, y] \subseteq \text{thr}^2 S$.

Case 2.a.1 : If $c \in \text{thr}(\text{thr}\, S \cup [0, y])$ then there are $a, b \in \text{thr}\, S \cup [0, y]$ such that $c \in [a, b]$. If $a, b \in \text{thr}\, S$ or $a, b \in [0, y]$ then $[a, b] \subseteq \text{thr}^2 S \cup \text{thr}[0, y] = \text{thr}^2 S \cup [0, y] = \text{thr}^2 S$ and hence $c \in \text{thr}^2 S$. Therefore $\text{thr}(\text{thr}\, S \cup [0, y]) \subseteq \text{thr}^2 S$.

Case 2.a.2 : Let $a \in \text{thr}\, S$ and $b \in [0, y]$. Denote by l' the line in $\mathbb{R}^2 \cap V(\mathbb{R}^2)$ passing through a and b. Then there is some $c' \in \text{thr}\, S$ such that $[a, b] \subseteq [a, c'] \subseteq l'$. On the other hand by construction we have $[a, c'] \subseteq \text{thr}^2 S$ implying $[a, b] \subseteq \text{thr}^2 S$ hence $c \in \text{thr}^2 S$. Therefore $\text{thr}(\text{thr}\, S \cup [0, y]) \subseteq \text{thr}^2 S$. Both Cases 2.a.1 and 2.a.2 yield $\text{thr}\, U \cup \text{thr}\, V \subseteq \text{thr}^2 S$.

Case 2.b.1 : Now let $c \in \text{thr}^2 S$ then there are $a, b \in \text{thr}\, S$ such that $c \in [a, b]$. If $a, b \in U$ or $a, b \in V$ then $[a, b] \subseteq \text{thr}\, U$ or $[a, b] \subseteq \text{thr}\, V$ respectively implying $[a, b] \subseteq \text{thr}\, U \cup \text{thr}\, V$. Therefore $c \in \text{thr}\, U \cup \text{thr}\, V$.

Case 2.b.2 : Now assume that $a \in U$ and $b \in V$. There are two possibilities for a to connect with b. First if the geodesic segment $[a, b]$ passes through 0 then $[a, 0] \subseteq \text{thr}\, U$ and $[0, b] \subseteq V$ since $0 \in U \cap V$. Hence $[a, b] = [a, 0] \cup [0, b] \subset \text{thr}\, U \cup \text{thr}\, V$ which in turn implies $c \in \text{thr}\, U \cup \text{thr}\, V$. The second case is when a is connected to b by a geodesic segment not passing through the origin 0. Then there exists some $z \in H$ such that $[a,b] \cap l_0 = \{z\}$. On the other hand since $a, b \in \operatorname{thr} S$ then there are $a_1, a_2, b_1, b_2 \in S$ such that $a \in [a_1, a_2]$ and $b \in [b_1, b_2]$. Up to labelling $[a, b]$ lies between geodesic segments $[a_1, b_1]$ and $[a_2, b_2]$. Let $y_i := [a_i, b_i] \cap l_0$ for $i = 1, 2$ then by maximal property of y we have $d'(0, y_1) \leqslant d'(0, z) \leqslant d'(0, y_2) \leqslant d'(0, y)$ implying $z \in [0, y]$. This means that $[a, z] \subseteq \operatorname{thr} U$ and $[z, b] \subseteq \operatorname{thr} V$ and thus $[a, b] \subseteq \operatorname{thr} U \cup \operatorname{thr} V$ which in turn yields $c \in \operatorname{thr} U \cup \operatorname{thr} V$. Therefore from both Cases 2.b.1 and 2.b.2 we have $\operatorname{thr}^2 S \subseteq \operatorname{thr} U \cup \operatorname{thr} V$. This shows that identity (5.14) holds true.

Following same arguments, by using induction on n, one can show that in general we have

$$\operatorname{thr}^n S = \operatorname{thr}^{n-1} U \cup \operatorname{thr}^{n-1} V, \;\; \forall n \in \mathbb{N} \tag{5.15}$$

By construction U, V are subsets of certain Euclidean half-spaces/orthants where the restriction of the length metric d' to U and V coincides with the restriction of the usual Euclidean metric there. By Example 1 we get $\deg_{\text{thr}} U, \deg_{\text{thr}} V \leqslant 2$. Identity (5.15) then implies that $\deg_{\text{thr}} S \leqslant 3$. □

Lemma 5.12. *For a compact set $S \subseteq \mathcal{T}_4$ consider the sets $U^k_{i,n} := \operatorname{thr}^n S \cap \operatorname{span}^+\{s^k_i\}$ for $k = 1, 2$ where $\operatorname{span}^+\{s^1_i, s^2_i\} = \mathcal{O}_i$. For each $i \in \{1, 2, ..., 15\}, k \in \{1, 2\}$ there exist $u^k_{i,n}, v^k_{i,n} \in U^k_{i,n}$ such that $d(O, v^k_{i,n}) \leqslant d(O, u) \leqslant d(O, u^k_{i,n})$ for all $u \in U^k_{i,n}$. Moreover there are $u^k_i, v^k_i \in U^k_{i,1}$ such that $d(O, v^k_i) \leqslant d(O, u) \leqslant d(O, u^k_i)$ for all $u \in U^k_{i,n}$ and all $n \in \mathbb{N}$.*

Proof. Since S is a compact set, then $\operatorname{thr}^n S$ is compact for all $n \in \mathbb{N}$ by Lemma 4.18. On the other hand $\operatorname{span}^+\{s^k_i\}$ is closed for any i and any k. Then $U^k_{i,n}$, as an intersection of a compact set and a closed set, is compact. The distance function $d(\cdot, O)$ is continuous therefore $d(\cdot, O)$ attains its maximum and minimum on each $U^k_{i,n}$ i.e. there exist $u^k_{i,n}, v^k_{i,n} \in U^k_{i,n}$ such that $d(O, v^k_{i,n}) \leqslant d(O, u) \leqslant d(O, u^k_{i,n})$ for all $u \in U^k_{i,n}$.

To prove the second claim note that $\operatorname{thr}^{n-1} S \cap \operatorname{span}^+\{s^k_i\} \subseteq \operatorname{thr}^n S \cap \operatorname{span}^+\{s^k_i\}$ implies $d(O, v^k_{i,n}) \leqslant d(O, v^k_{i,n-1}) \leqslant d(O, u^k_{i,n-1}) \leqslant d(O, u^k_{i,n})$. Therefore it is enough to show

$$d(O, v^k_{i,n-1}) \leqslant d(O, v^k_{i,n}) \leqslant d(O, u^k_{i,n}) \leqslant d(O, u^k_{i,n-1}), \forall n \in \mathbb{N}$$

We consider the case $n = 2$. By induction the general case follows. Let $z \in \operatorname{thr}^2 S \cap \operatorname{span}^+\{s^k_i\}$ then there are $x, y \in \operatorname{thr} S$ such that $z \in [x, y]$. Moreover we can find $x_0, x_1, y_0, y_1 \in S$ such that $x \in [x_0, x_1], y \in [y_0, y_1]$. Without loss of generality let x_0, x_1 be in the same orthant, likewise y_0, y_1 (other cases follow similarly). Consider the quadrangle $x_0 x_1 y_0 y_1$ with sides

$$[x_0, x_1], [x_1, y_1], [y_0, y_1], [x_0, y_0]$$

Up to labelling of the points $\{x_0, x_1, y_0, y_1\}$ the geodesic segment $[x, y]$ lies between $[x_0, y_0]$ and $[x_1, y_1]$. Say $[x_0, y_0]$ is the lower segment and $[x_1, y_1]$ is the upper segment with respect to $[x, y]$. If $a = [x_0, y_0] \cap \operatorname{span}^+\{s^k_i\}$ and $b = [x_1, y_1] \cap \operatorname{span}^+\{s^k_i\}$ then $d(O, a) \leqslant d(O, z) \leqslant d(O, b)$. On the other hand $d(O, v^k_i) \leqslant d(O, a) \leqslant d(O, b) \leqslant d(O, u^k_i)$ implies $d(O, v^k_i) \leqslant d(O, z) \leqslant d(O, u^k_i)$. Since $z \in \operatorname{thr}^2 S \cap \operatorname{span}^+\{s^k_i\}$ was arbitrary then $d(O, v^k_i) \leqslant d(O, v^k_{i,2}) \leqslant d(O, u^k_{i,2}) \leqslant d(O, u^k_i)$. □

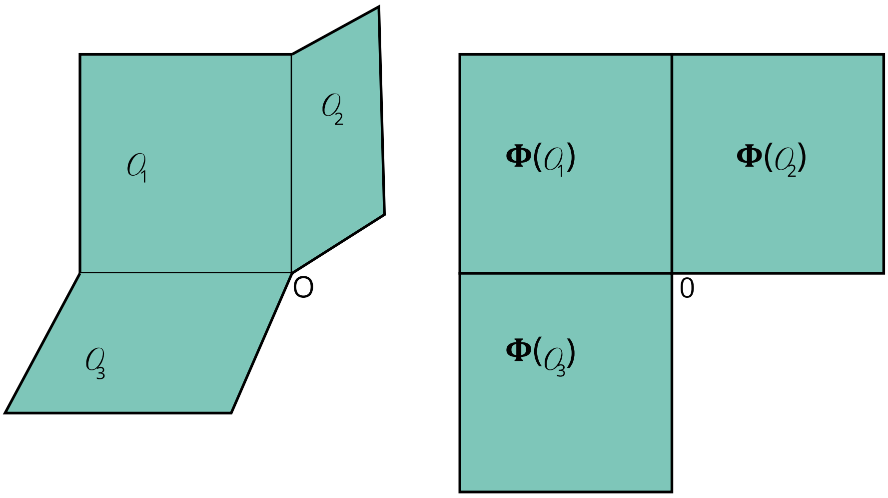


Figure 5.5: A regular tiplet $\{\mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3\}$ in $\mathcal{T}_4$ in general position (left) and its isometric image $V(\mathbb{R}^2)$ (right).

5.3.2 Convex hull of compact sets in $\mathcal{T}_4$

Let $\{\mathcal{O}_i\}_{i=1,2,\dots,15}$ be the collection of the 2-dimensional orthants in $\mathcal{T}_4$. Consider all triples $\{\mathcal{O}_i, \mathcal{O}_j, \mathcal{O}_k\}_{i,j,k}$ such that $\mathcal{O}_i \cup \mathcal{O}_j \cup \mathcal{O}_k$ is isometricly isomomorphic to $V(\mathbb{R}^2)$. Call them *regular triplets*. There are at most $\binom{15}{3}$ such triplets and hence a finite number of them. Let Φ_{ijk} be the isomorphism from $\mathcal{O}_i \cup \mathcal{O}_j \cup \mathcal{O}_k$ onto $V(\mathbb{R}^2)$ (see Figure 5.5). Moreover for each $\mathcal{O}_i$ there are 10 distinct orthants $\mathcal{O}_j$ such that $\mathcal{O}_i \cap \mathcal{O}_j = \{O\}$ and only 2 of these are such that no split in $\Sigma(\mathcal{O}_i)$ is compatible with any split in $\Sigma(\mathcal{O}_j)$. Denote these special orthants by $\mathcal{O}_{i_j}$ where $j = 1, 2$. By Corollary 5.10 each union $\mathcal{O}_i \cup \mathcal{O}_{i_j}$ is isometricly isomomorphic to $\mathbb{R}^2_+ \cup \mathbb{R}^2_-$. Denote these isomorphisms by Φ_{i_j}.

Theorem 5.13. *The convex hull of a compact set in $\mathcal{T}_4$ is compact.*

Proof. Let $S \subseteq \mathcal{T}_4$ be a compact set and denote by $S_i := \mathcal{O}_i \cap \operatorname{thr} S$ for each $i = 1, 2, \dots, 15$. Let $S_{ijk} := S_i \cup S_j \cup S_k$ for all i, j, k. Note that S_i is compact for all $i = 1, 2, \dots, 15$.

We claim

$$\operatorname{co} S = \Big(\bigcup_{i,j,k} \operatorname{co} S_{ijk}\Big) \cup \Big(\bigcup_{i=1}^{i=15} \bigcup_{j=1,2} \operatorname{co}(S_i \cup S_{i_j})\Big). \tag{5.16}$$

By Theorem 4.15 to prove the statement it suffices to show the following identity

$$\operatorname{thr}^n S = \Big(\bigcup_{i,j,k} \operatorname{thr}^{n-1} S_{ijk}\Big) \cup \Big(\bigcup_{i=1}^{i=15} \bigcup_{j=1,2} \operatorname{thr}^{n-1}(S_i \cup S_{i_j})\Big), \quad \forall n \in \mathbb{N}. \tag{5.17}$$

Denote by S_n^* the right side of (5.17). The inclusions $S_{ijk}, S_i \cup S_{i_j} \subseteq \operatorname{thr} S, \forall i, j$ by virtue of identity (4.10) imply $S_n^* \subseteq \operatorname{thr}^n S$. The other direction we use induction. For $n = 1$ the equality in (5.17) is evident. Suppose that it is true for $m \geqslant 1$. Let $n = m + 1$. and $[x, y] \subseteq \operatorname{thr}^{m+1} S$. Then by definition there are $[x_1, x_2], [y_1, y_2] \subseteq \operatorname{thr}^m S$ such that

$x \in [x_1, x_2]$ and $y \in [y_1, y_2]$. By inductive hypothesis we have $[x_1, x_2], [y_1, y_2] \subseteq S_m^*$. In particular $x, y \in S_m^*$. If $\Sigma(x)$ has no compatible split with $\Sigma(y)$ then the cone path is a geodesic. Without loss of generality let $x \in \mathcal{O}_i$ and $y \in \mathcal{O}_p$ for some indices i, p. This means that $[x, y] = [x, O] \cup [O, y]$ where

$$[x, O] \subseteq \operatorname{thr}(\operatorname{thr}^{m-1} S_{ijk} \cap \mathcal{O}_i) \subseteq \operatorname{thr}^m S_{ijk} \cap \mathcal{O}_i \subseteq \operatorname{thr}^m S_{ijk} \tag{5.18}$$

or

$$[x, O] \subseteq \operatorname{thr}(\operatorname{thr}^{m-1}(S_i \cup S_{i_j}) \cap \mathcal{O}_i) \subseteq \operatorname{thr}^m(S_i \cup S_{i_j}) \tag{5.19}$$

and

$$[O, y] \subseteq \operatorname{thr}(\operatorname{thr}^{m-1} S_{pqr} \cap \mathcal{O}_p) \subseteq \operatorname{thr}^m S_{pqr} \cap \mathcal{O}_p \subseteq \operatorname{thr}^m S_{pqr} \tag{5.20}$$

or

$$[O, y] \subseteq \operatorname{thr}(\operatorname{thr}^{m-1}(S_p \cup S_{p_q}) \cap \mathcal{O}_p) \subseteq \operatorname{thr}^m(S_p \cup S_{p_q}) \tag{5.21}$$

If $\Sigma(x)$ has some compatible split with $\Sigma(y)$ then there is a regular triplet $\{\mathcal{O}_i, \mathcal{O}_j, \mathcal{O}_k\}$ such that $x, y \in \mathcal{O}_i \cup \mathcal{O}_j \cup \mathcal{O}_k$. If the cone path is a geodesic then it follows by (5.18)-(5.21) that $[x, y] = [x, O] \cup [O, y] \subseteq S_{m+1}^*$. The other situation is when $[x, y]$ is not a cone path. Without loss of generality let $x \in \operatorname{thr}^{m-1} S_{ijl} \cap \mathcal{O}_i$ and $y \in \operatorname{thr}^{m-1} S_{kjm} \cap \mathcal{O}_k$ for some $l \neq m$. Let $\Sigma(\mathcal{O}_i)$ have no common split with $\Sigma(\mathcal{O}_k)$. Then there exist two points $u \in \operatorname{span}^+\{s_j^1\}$ and $v \in \operatorname{span}^+\{s_j^2\}$ such that $[x, y] = [x, u] \cup [u, v] \cup [v, y]$. Lemma 5.12 then implies that $u \in [v_j^1, u_j^1]$ and $v \in [v_j^2, u_j^2]$. On the other hand $[v_j^1, u_j^1], [v_j^2, u_j^2] \subseteq \operatorname{thr} S_j$ hence we have $[u, v] \in \operatorname{thr}^2 S_j \subseteq \operatorname{thr}^m S_j \subseteq \operatorname{thr}^m S_{ijl}$. Similarly $u \in \operatorname{thr} S_j \subseteq \operatorname{thr} S_{ijl} \subseteq \operatorname{thr}^{m-1} S_{ijl}$ and $v \in \operatorname{thr} S_j \subseteq \operatorname{thr} S_{kjm} \subseteq \operatorname{thr}^{m-1} S_{kjm}$. Therefore $[x, u] \subseteq \operatorname{thr}^m S_{ijl}, [v, y] \subseteq \operatorname{thr}^m S_{kjm}$ implies $[x, y] \subseteq \operatorname{thr}^m S_{ijl} \cup \operatorname{thr}^m S_{kjm}$. In all cases we get $[x, y] \subseteq S_{m+1}^*$ implying $\operatorname{thr}^{m+1} S \subseteq S_{m+1}^*$. This proves identity (5.17).

By virtue of Example 2 and Lemma 5.11 we have that

$$\deg_{\operatorname{thr}} \Phi_{ijk}(S_{ijk}) \leqslant 3, \deg_{\operatorname{thr}} \Phi_{i_j}(S_i \cup S_{i_j}) \leqslant 2, \forall i, j, k$$

which together with $\operatorname{thr}^n S_{ijk} = \Phi_{ijk}^{-1}(\Phi_{ijk}(\operatorname{thr}^n S_{ijk})) = \Phi_{ijk}^{-1}(\operatorname{thr}^n \Phi_{ijk}(S_{ijk}))$ and similarly $\operatorname{thr}^n(S_i \cup S_{i_j}) = \Phi_{i_j}^{-1}(\Phi_{i_j}(\operatorname{thr}^n(S_i \cup S_{i_j}))) = \Phi_{i_j}^{-1}(\operatorname{thr}^n \Phi_{i_j}(S_i \cup S_{i_j}))$ imply

$$\deg_{\operatorname{thr}} S_{ijk} \leqslant 3, \deg_{\operatorname{thr}}(S_i \cup S_{i_j}) \leqslant 2, \forall i, j, k$$

Identity (5.17) yields $\deg_{\operatorname{thr}} S \leqslant 4$. For large enough n, say $n \geqslant 4$, identity (5.17) implies (5.16). But $\operatorname{co} \Phi_{ijk}(S_{ijk}), \operatorname{co} \Phi_{i_j}(S_i \cup S_{i_j})$ are compact for any i, j, k. Because $\Phi_{ijk}^{-1}, \Phi_{i_j}^{-1}$ are isometries, hence continuous maps, for any i, j, k then $\operatorname{co} S_{ijk} = \Phi_{ijk}^{-1}(\operatorname{co} \Phi_{ijk}(S_{ijk}))$ and $\operatorname{co}(S_i \cup S_{i_j}) = \Phi_{i_j}^{-1}(\operatorname{co} \Phi_{i_j}(S_i \cup S_{i_j}))$ are compact. It follows that $\operatorname{co} S$ is a compact set as a finite union of compact sets. This completes the proof. □

A direct consequence of Theorem 5.13 and Theorem 4.25 is the following corollary which we present without proof.

Corollary 5.14. *Let $x_1, x_2, ..., x_n \in \mathcal{T}_4$ be a finite set of trees with four leaves. Then the Fréchet mean x^* exists and lies in $\operatorname{co}\{x_1, x_2, ..., x_n\}$. Moreover x^* is constructible from $\{x_1, x_2, ..., x_n\}$ in at most 15 steps.*

Proof. Follows from Theorem 5.13 and Theorem 4.25. □

Chapter 6

Mosco Convergence and Asymptotic Boundedness

6.1. Mosco Convergence

6.1.1 Mosco convergence of functions

Let (H, d) be a Hadamard space. A sequence of functions $f^n : H \to (-\infty, +\infty]$ is said to be Mosco convergent to $f : H \to (-\infty, +\infty]$ and we write $M - \lim_n f^n = f$ if for each $x \in H$:

(i) $f(x) \leqslant \liminf_n f^n(x_n)$ whenever $x_n \xrightarrow{w} x$

(ii) there exists some sequence $(y_n)_{n\in\mathbb{N}} \subset H$ such that $y_n \to x$ and $f(x) \geqslant \limsup_n f^n(y_n)$.

If (i) is substituted with strong convergence then one gets what is known as Γ-convergence[1]. Therefore Mosco convergence is a stronger type of convergence and subsequently Mosco convergence implies Γ-convergence. The original motivation for introducing Mosco convergence in analysis was to define a special convergence for convex closed sets of a normed space X, in which both the strong and the weak topologies of X are involved (see [91, Definition 1]). Another way to introduce Mosco convergence has been to make the so called *Fenchel conjugate* f^* of a *closed convex proper* function f bicontinous (see [11, pg. 294]). In general if (X, τ) is a *locally convex* topological vector space and $f : X \to (-\infty, +\infty]$ is a convex proper function then the Fenchel conjugate of $f^* : X^* \to (-\infty, +\infty]$ of f is defined as

$$f^*(u) := \sup_{x\in X}(\langle u, x\rangle - f(x)), \ \ \forall u \in X^* \tag{6.1}$$

where X^* is the *topological dual* of X and $\langle u, x\rangle$ denotes the dual pairing of an element $x \in X$ with an element $u \in X^*$. In the case of a Hilbert space this dual pairing is just the canonical inner product. It follows by definition (6.1) that f^* is a proper closed convex function whenever f is. Analogous to (6.1) one can define the Fenchel conjugate f^{**} of f^* also known as the biconjugate of f.

[1]For more on Γ-convergence see [86, G. Dal Maso]).

Theorem 6.1 ([11, Fenchel-Moreau]). *Let (X, τ) be a locally convex topological vector space and $f : X \to (-\infty, +\infty]$ a proper convex function. The biconjugate f^{**} of a function f is equal to $\mathrm{cl}_\tau f$ (the τ-closure of f). In particular if f is τ-closed then $f^{**} = f$.*

One can think of Fenchel conjugation as a transformation from the set of proper closed convex functions on X onto the set of proper closed convex functions on X^*. By virtue of Theorem 6.1 this *Fenchel transform* is a one-to-one mapping.

Theorem 6.2 ([11, Attouch-Fenchel-Moreau]). *Let X be a reflexive Banach space and $(f^n)_{n\in\mathbb{N}}, f$ a sequence of proper closed convex functions from X into $(-\infty, +\infty]$. Then $M - \lim_n f^n = f$ iff $M - \lim_n (f^n)^* = f^*$*

In other words Fenchel transform is bicontinous for the Mosco convergence. While Theorem 6.1 and Theorem 6.2 are applicable to Hilbert spaces, which are a proper subset of Hadamard spaces, we cannot in general extend such results to any Hadamard space. The main reason is that there does not exist yet a good notion for the dual pairing. For an extension of duality theory that generalizes the classical Fenchel conjugation to functions defined on Riemannian manifolds the reader could refer to a recent work of Bergmann et al. [25]. Nevertheless Mosco convergence still plays a crucial role in general Hadamard spaces, in particular in the theory related to the gradient flow (see Bačak [43] for instance).

6.1.2 Mosco convergence of sets

Let ι_S denote the indicator function of a set $S \subseteq H$ i.e. $\iota_S(x) = 0$ if $x \in S$ and $\iota_S(x) = +\infty$ otherwise. A sequence of sets $(S_n)_{n\in\mathbb{N}}$ is said to converge in the sense of Mosco to a set S whenever $(\iota_{S_n})_{n\in\mathbb{N}}$ Mosco converges to ι_S.

Proposition 6.3 ([43, Corollary 5.2.8]). *Let (H, d) be a Hadamard space and $(C_n)_{n\in\mathbb{N}}$ a sequence of closed convex sets. If $M - \lim_n C_n = C$ for some set $C \subseteq H$ then C is closed and convex.*

Proof. By definition $M - \lim_n C_n = C$ means $M - \lim_n \iota_{C_n} = \iota_C$. C_n is convex and closed for all n implies that the indicator function ι_{C_n} is closed convex for all n. But Mosco convergence preserves convexity and lower-semicontinuity therefore ι_C is a closed convex function. This is equivalent to C being a closed convex set. □

Proposition 6.4. *Let $(C_n)_{n\in\mathbb{N}}$ be a sequence of closed convex sets in a Hadamard space H. If $(C_n)_{n\in\mathbb{N}}$ is a nonincreasing sequence of sets then $(C_n)_{n\in\mathbb{N}}$ Mosco converges to its intersection. If $(C_n)_{n\in\mathbb{N}}$ is nondecreasing then it Mosco converges to the closure of its union.*

Proof. The proof follows the lines in [91, Lemma 1.2, Lemma 1.3]. Let $(C_n)_{n\in\mathbb{N}}$ be a nonincreasing sequence of closed convex sets and $C := \bigcap_n C_n$. First note that $C \neq \emptyset$. By definition it suffices to prove that $M - \lim_n \iota_{C_n} = \iota_C$. Let $x \in H$. If $x \notin C$ then clearly $\limsup_n \iota_{C_n}(y_n) = 0 \leqslant \iota_C(x)$ for any sequence $(y_n)_{n\in\mathbb{N}}$ such that $y_n \to x$. Now let $x \in C$. Then we have $x \in C_n$ for every $n \in \mathbb{N}$. Because C_n is a closed set there exists $y_n \in C_n$ such that $d(x, y_n) \leqslant 1/n$, therefore $y_n \to x$. Since $y_n \in C_n$ then $\iota_{C_n}(y_n) = 0$ for all $n \in \mathbb{N}$, hence $\limsup_n \iota_{C_n}(y_n) = 0 = \iota_C(x)$ confirming condition (ii). Now let $(x_n)_{n\in\mathbb{N}}$ be such that $x_n \in C_n$ for all $n \in$ and $w - \lim_n x_n = x$ for some $x \in H$. Assumption $C_n \subseteq C_m$ whenever $m \leqslant n$ implies that $x_n \in C_m$. But C_m is a closed convex set hence by Lemma 3.9 weakly sequentially closed. Therefore $x = w - \lim_n x_n \in C_m$ and this holds for any $m \in \mathbb{N}$ since m was arbitrary. Therefore $x \in C$ implying $\iota_C(x) = 0 \leqslant \liminf_n \iota_{C_n}(x_n)$ which confirms condition (i). Analogue arguments for the second statement. □

6.2. A Theorem of Attouch

6.2.1 Attouch's theorem about Mosco convergence

Let X be a *normed linear* space and $f : X \to (-\infty, +\infty]$ a proper closed convex function. For $\lambda > 0$ the *Moreau-Yosida* envelope of f is defined as follows

$$f_\lambda(x) := \inf_{y\in X} \left\{ f(y) + \frac{1}{2\lambda}\|x - y\|^2 \right\} \tag{6.2}$$

It can be shown that f_λ is a convex continuous function [11, Proposition 3.3]. Moreover $\lim_{\lambda\to 0} f_\lambda(x) = f(x)$ for all $x \in X$. For a given parameter $\lambda > 0$ the *proximal mapping* of f is defined as

$$J_\lambda x := \arg\min_{y\in X} \left\{ f(y) + \frac{1}{2\lambda}\|x - y\|^2 \right\} \tag{6.3}$$

For a function $f : X \to (-\infty, +\infty]$ let $\partial f(x)$ denote the subdifferential of f at $x \in X$ defined as

$$\partial f(x) := \{u \in X^* | f(x) \geqslant f(y) + \langle u, y - x\rangle, \forall y \in X\} \tag{6.4}$$

and we say a pair $(x, u) \in X \times X^*$ lies in ∂f whenever $u \in \partial f(x)$. For more on fundamental concepts in convex analysis in linear spaces refer to the classical book by Rockafellar [99].

Theorem 6.5 ([11, Attouch's Theorem])**.** *Let X be a smooth reflexive Banach space. Let $(f^n)_{n\in\mathbb{N}}, f$ be a sequence of proper closed convex functions from X into $(-\infty, +\infty]$. The following equivalences hold:*

(i) $M - \lim_n f^n = f$

(ii) $\forall\lambda > 0, \forall x \in X$ *it holds* $\lim_n J^n_\lambda x = J_\lambda x$ *and* $\exists(u, v) \in \partial f, \exists(u_n, v_n) \in \partial f^n$ *such that* $\lim_n u_n = u$ *in* X*,* $\lim_n v_n = v$ *in* X^**, and* $\lim_n f^n(u_n) = f(u)$

(iii) $\forall\lambda > 0, \forall x \in X$ *it holds* $\lim_n f^n_\lambda(x) = f_\lambda(x)$*.*

Note that Theorem 6.5 appeared first in [10] for Hilbert spaces and then generalized for any smooth reflexive Banach space in [11].

6.2.2 A theorem of Bačak

Because a norm $\|\cdot\|$ in a linear space X induces a metric $d(x,y) = \|x-y\|$ for any $x, y \in X$ then definitions (6.2) and (6.3) can be accommodated easily in the setting of a Hadamard space using its metric. For a given closed convex function $f : H \to (-\infty, +\infty]$ and parameter $\lambda > 0$ the Moreau envelope f_λ of f is defined as

$$f_\lambda(x) := \inf_{y\in H} \left\{ f(y) + \frac{1}{2\lambda} d(y,x)^2 \right\}, \quad \text{for each} \quad x \in H \tag{6.5}$$

and the proximal mapping of f

$$J_\lambda x := \arg\min_{y\in H} \left\{ f(y) + \frac{1}{2\lambda} d(y,x)^2 \right\}, \quad \text{for each} \quad x \in H \tag{6.6}$$

In his study of the gradient flow in Hadamard spaces [43] Bačak established a result which relates Mosco convergence of a sequence of closed convex functions $(f^n)_{n\in\mathbb{N}}$ to the pointwise convergence of Moreau envelopes $(f^n_\lambda)_{n\in\mathbb{N}}$ and proximal mappings $(J^n_\lambda)_{n\in\mathbb{N}}$.

Theorem 6.6. *(Bačak) Let (H,d) be a Hadamard space and $f^n : H \to (-\infty, +\infty]$ a sequence of closed convex functions. If $M - \lim_n f^n(x) = f(x)$, then $\lim_n f^n_\lambda(x) = f_\lambda(x)$ and $\lim_n J^n_\lambda x = J_\lambda x$ for each $x \in H$.*

This result is the analogue of the implication $(i) \to (iii)$ in Theorem 6.5. Later Bačak et al. [13] proved the following.

Theorem 6.7. *Let (H,d) be a Hadamard space and $f, f^n : H \to (-\infty, +\infty]$ be a sequence of closed convex functions. If $\lim_n f^n_\lambda(x) = f_\lambda(x)$ then $M - \lim_n f^n(x) = f(x)$ for all $x \in H$.*

This result together with Theorem 6.7 imply the equivalence between Mosco convergence and pointwise convergence of Moreau envelopes in Hadamard spaces. This completes the equivalence $(i) \leftrightarrow (iii)$ in Theorem 6.5 for Hadamard spaces. However it is not known whether convergence of proximal mappings imply, under some additional conditions, the Mosco convergence of f^n. This was left an open question by Bačak [43]. That convergence of proximal mappings only is not enough was noted by Bačak in [12]. Indeed consider a sequence of constant functions $0, 1, 0, 1, \ldots$ defined on $\mathbb{R}$. Evidently they are closed and convex but they don't converge in the sense of Mosco to any function f. However their proximal mapping maps $J_\lambda : \mathbb{R} \to \mathbb{R}$ (i.e. $x \mapsto J_\lambda x$) equal the identity map for all $\lambda > 0$. In this note we aim to complete the cycle of equivalences, the analogues of Attouch's theorem. This also answers an open question in [12].

6.3. Asymptotic Boundedness

6.3.1 Some preliminaries

Definition 6.8. *Let $f : H \to (-\infty, +\infty]$ be a closed convex function and $x \in \operatorname{dom} f$. The slope of f at x is defined as*

$$|\partial f|(x) := \limsup_{y \to x} \frac{\max\{f(x) - f(y), 0\}}{d(x, y)} \tag{6.7}$$

and $\operatorname{dom} |\partial f| := \{x \in H : |\partial f|(x) < +\infty\}$*. If $f(x) = +\infty$ we set $|\partial f|(x) := +\infty$.*

It follows that $|\partial f|(x) = 0$ whenever $x \in H$ is a minimizer of f. The following inclusion $\operatorname{dom} |\partial f| \subseteq \operatorname{dom} f$ is obvious. Moreover the followings are true $|\partial(f + g)|(x) \leqslant |\partial f|(x)| + |\partial g|(x)$ and $|\partial(\alpha f)|(x) = \alpha |\partial f|(x)$ for any two functions f, g and any scalar $\alpha > 0$.

Lemma 6.9 ([43, Lemma 5.1.2])**.** *Let $f : H \to (-\infty, +\infty]$ be a closed convex function. Then*

$$|\partial f|(x) = \sup_{y \in H \setminus \{x\}} \frac{\max\{f(x) - f(y), 0\}}{d(x, y)}, \quad x \in \operatorname{dom} f \tag{6.8}$$

Moreover $\operatorname{dom} |\partial f|$ *is dense in* $\operatorname{dom} f$ *and $|\partial f|(x)$ is closed whenever f is closed.*

Lemma 6.10 ([43, Lemma 5.1.3])**.** *Let $f : H \to (-\infty, +\infty]$ be a closed convex function. Then for every $x \in H$ and $\lambda > 0$ we have $J_\lambda x \in \operatorname{dom} |\partial f|$ and*

$$|\partial f|(J_\lambda x) \leqslant \frac{d(J_\lambda x, x)}{\lambda} \tag{6.9}$$

Proposition 6.11 ([43, Proposition 2.2.17])**.** *Let (H, d) be a Hadamard space and let $f : H \to (-\infty, +\infty]$ be a closed strongly convex function with parameter $\mu > 0$ (see definition 2.9). Then f has a unique minimizer $x \in H$ and each minimizing sequence converges to x. Moreover*

$$f(x) + \frac{\mu}{2} d(x, y)^2 \leqslant f(y), \quad \forall y \in H \tag{6.10}$$

Proof. By virtue of [43, Lemma 2.2.14] f is bounded from below. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence. Denote by $x_{mn} := \frac{1}{2} x_m \oplus \frac{1}{2} x_n$. By strong convexity we have

$$f(x_{mn}) \leqslant \frac{1}{2} f(x_m) + \frac{1}{2} f(x_n) - \frac{\mu}{8} d(x_m, x_n)^2$$

which in turn implies

$$\frac{\mu}{8} d(x_m, x_n)^2 \leqslant \frac{1}{2} f(x_m) + \frac{1}{2} f(x_n) - f(x_{mn}).$$

If $(x_n)_{n \in \mathbb{N}}$ is a minimizing sequence of f i.e. $\lim_n f(x_n) = \inf_{y \in H} f(y)$ then so it is the new sequence $(x_{mn})_{m,n \in \mathbb{N}}$. In particular $\lim_{m,n} d(x_m, x_n) = 0$ implies $(x_n)_{n \in \mathbb{N}}$ is Cauchy sequence. Because H is a complete metric space then $(x_n)_{n\in\mathbb{N}}$ converges to some point $x \in H$. Assumption f is closed is equivalent to f being lower-semicontinuous. The inequalities $f(x) \leqslant \liminf_n f(x_n) = \inf_{y\in H} f(y)$ and $f(x) \geqslant \inf_{y\in H} f(y)$ imply that $x \in \arg\min_{y\in H} f(y)$. Uniqueness of minimizer follows immediately from the strong convexity property. Now consider some $y \in H$ and let $\gamma : [0,1] \to H$ be the geodesic emanating from x and ending at y i.e. $\gamma(0) = x, \gamma(1) = y$. Then $f(x) < f(\gamma(t))$ together with the strong convexity imply

$$f(x) < (1-t)f(x) + tf(y) - \frac{\mu}{2}(1-t)td(x,y)^2$$

or equivalently

$$tf(x) < tf(y) - \frac{\mu}{2}(1-t)td(x,y)^2$$

Dividing by t and taking limit $t \downarrow 0$ yields inequality (6.10). □

6.3.2 Asymptotic boundedness for a sequence of functions

Definition 6.12. *A sequence of functions $f^n : H \to (-\infty, +\infty]$ is said to have pointwise asymptotically bounded slope on H whenever $\limsup_n |\partial f^n|(x)$ is finite for all $x \in H$. If additionally for all $x \in H$ we have $\limsup_n |\partial f^n|(x) \leqslant C$ for some $C > 0$ then the sequence of functions f^n is said to have uniform asymptotically bounded slope on H.*

Recall that a set K of a vector space V is a cone (or sometimes called a linear cone) if for each x in K and positive scalars α, the product αx is in K. The set K is a convex cone if and only if any nonnegative combination of elements from K remains in K. Let $F(H)$ denote the vector space of sequences of (extended) real valued functions defined on H and let $A(H) := \{(f_n)_{n\in\mathbb{N}} \in F(H) \mid \limsup_n |\partial f^n|(x) < +\infty, \forall x \in H\}$ denote the set of all sequences that have pointwise asymptotically bounded slope on H.

Proposition 6.13. *$A(H)$ is a convex cone in $F(H)$.*

Proof. It suffices to prove the statement for only two elements. Let $(f^n)_{n\in\mathbb{N}}, (g^n)_{n\in\mathbb{N}} \in F(H)$ and $\alpha, \beta > 0$. Denote by $h^n := \alpha f^n + \beta g^n$ for each $n \in \mathbb{N}$. By definition of the slope (6.7) we have

$$|\partial h^n|(x) = \limsup_{y\to x} \frac{\max\{h^n(x) - h^n(y), 0\}}{d(x,y)}$$

On the other hand

$$\max\{h^n(x) - h^n(y), 0\} \leqslant \alpha \max\{f^n(x) - f^n(y), 0\} + \beta \max\{g^n(x) - g^n(y), 0\}$$

and the fact that the limit superior of the sum is not greater than the sum of limit superior together with $\alpha, \beta > 0$ imply

$$|\partial h^n|(x) \leqslant \alpha \limsup_{y\to x} \frac{\max\{f^n(x) - f^n(y), 0\}}{d(x,y)} + \beta \limsup_{y\to x} \frac{\max\{g^n(x) - g^n(y), 0\}}{d(x,y)}$$

or equivalently

$$|\partial h^n|(x) \leqslant \alpha|\partial f^n|(x) + \beta|\partial g^n|(x), \ \ \forall n \in \mathbb{N}$$

Taking limit superior with respect to n on both sides yields

$$\limsup_n |\partial h^n|(x) \leqslant \limsup_n(\alpha|\partial f^n|(x)+\beta|\partial g^n|(x)) \leqslant \alpha \limsup_n |\partial f^n|(x)+\beta \limsup_n |\partial g^n|(x)$$

Assumption $(f^n)_{n\in\mathbb{N}}, (g^n)_{n\in\mathbb{N}} \in F(x)$ imply $\limsup_n |\partial f^n|(x), \limsup_n |\partial g^n|(x) < +\infty, \forall x \in H$. Hence $\limsup_n |\partial h^n|(x)| < +\infty$ for each $x \in H$ gives $(h^n)_{n\in\mathbb{N}} \in F(H)$ as desired. □

Remark 6.14. *The set $A_0(H)$ of sequences of functions with uniform asymptotic regular slope is also a convex cone.*

Proposition 6.15. *Let $(f^n)_{n\in\mathbb{N}}$ be a sequence of proper closed convex functions defined on a Hadamard space (H, d). Let f be the pointwise limit of $(f^n)_{n\in\mathbb{N}}$ such that $\operatorname{dom}|\partial f| \neq \emptyset$. Assume that $\operatorname{dom} f^n = \operatorname{dom} f = H$ for all n. For a given element $x \in H$ define the sequence of functions (g^n) and g for all $y \in H \setminus \{x\}$*

$$g^n(y;x) := \frac{\max\{f^n(x) - f^n(y), 0\}}{d(x,y)}, \ \ n \in \mathbb{N}$$
$$g(y;x) := \frac{\max\{f(x) - f(y), 0\}}{d(x,y)}$$

Then $(f^n)_{n\in\mathbb{N}}$ has pointwise asymptotically bounded slope on $\operatorname{dom}|\partial f|$ whenever

$$\lim_n \sup_{y\in H\setminus\{x\}} |g^n(y;x) - g(y;x)| = 0 \tag{6.11}$$

If additionally $\sup_{x\in\operatorname{dom}|\partial f|} |\partial f|(x) < +\infty$ then $(f^n)_{n\in\mathbb{N}}$ has uniform asymptotically bounded slope on $\operatorname{dom}|\partial f|$.

Proof. From the elementary reverse triangle inequality

$$\sup_{y\in H\setminus\{x\}} |g^n(y;x) - g(y;x)| \geqslant |\sup_{y\in H\setminus\{x\}} g^n(y;x) - \sup_{y\in H\setminus\{x\}} g(y;x)|$$

assumption (6.11) implies $\lim_n \sup_{y\in H\setminus\{x\}} g^n(y;x) = \sup_{y\in H\setminus\{x\}} g(y;x)$. By virtue of Lemma 6.9 this is equivalent to $\lim_n |\partial f^n|(x) = |\partial f|(x)$. Since $\operatorname{dom}|\partial f| \neq \emptyset$ then $\lim_n |\partial f^n|(x)$ is finite on $\operatorname{dom}|\partial f|$. Therefore $(f^n)_{n\in\mathbb{N}}$ has pointwise asymptotically bounded slope on $\operatorname{dom}|\partial f|$. If additionally $\sup_{x\in\operatorname{dom}|\partial f|} |\partial f|(x) < +\infty$ then $|\partial f|(x) \leqslant C$ for some $C > 0$ for all $x \in \operatorname{dom}|\partial f|$. This implies $\lim_n |\partial f^n|(x) \leqslant C$ for all $x \in \operatorname{dom}|\partial f|$. □

6.3.3 A converse theorem

Theorem 6.16. *Let (H, d) be a Hadamard space and $f^n : H \to (-\infty, +\infty]$ be a sequence of closed convex functions. Suppose*

(i) $\lim_n f^n(x) = f(x)$ *for all* $x \in H$

(ii) $(f^n)_{n\in\mathbb{N}}$ *has pointwise asymptotically bounded slope on* H

If $\lim_n J^n_\lambda x = J_\lambda x$ *then* $\lim_n f^n_\lambda(x) = f_\lambda(x)$ *for each* $x \in H$.

Proof. Note that f^n is convex for each n. Since the metric $d(\cdot, x)^2$ is a strongly convex function then the map

$$y \mapsto f^n(y) + \frac{1}{2\lambda} d(y,x)^2$$

is strongly convex for each $x \in H$. It follows from Proposition 6.11 that the proximal mapping

$$J^n_\lambda x := \arg\min_{y\in H} \left\{ f^n(y) + \frac{1}{2\lambda} d(y,x)^2 \right\}$$

exists and it is unique. Similarly for $J_\lambda x$. By definition for all n we have

$$f^n_\lambda(x) = f^n(J^n_\lambda x) + \frac{1}{2\lambda} d(x, J^n_\lambda x)^2$$

From the elementary triangle inequality $d(x, J^n_\lambda x) \leqslant d(x, J_\lambda x) + d(J_\lambda x, J^n_\lambda x)$ and interchanging the role of $J^n_\lambda x$ with $J_\lambda x$ we obtain the estimate

$$|d(x, J^n_\lambda x) - d(x, J_\lambda x)| \leqslant d(J_\lambda x, J^n_\lambda x)$$

Assumption $\lim_n J^n_\lambda x = J_\lambda x$ implies $\lim_n d(x, J^n_\lambda x) = d(x, J_\lambda x)$ for each $x \in H$. Therefore it is sufficient to prove $\lim_n f^n(J^n_\lambda x) = f(J_\lambda x)$. By Lemma 6.10, $J_\lambda x \in \operatorname{dom} |\partial f|$ for any $x \in H$ yields $J_\lambda x \in \operatorname{dom} f$ since $\operatorname{dom} |\partial f| \subseteq \operatorname{dom} f$. Similarly $J^n_\lambda x \in \operatorname{dom} f^n$. From the definition of Moreau envelope it follows that for all n

$$f^n(J^n_\lambda x) + \frac{1}{2\lambda} d(J^n_\lambda x, x)^2 \leqslant f^n(J_\lambda x) + \frac{1}{2\lambda} d(J_\lambda x, x)^2$$

which in turn together with assumption i and $\lim_n J^n_\lambda x = J_\lambda x$ gives

$$-\infty \leqslant \limsup_n f^n(J^n_\lambda x) \leqslant f(J_\lambda x) < +\infty \tag{6.12}$$

On the other hand assumption ii implies that for some finite valued nonnegative function $C : H \to \mathbb{R}_+$ we have $\limsup_n |\partial f^n|(x) \leqslant C(x)$ for all $x \in H$. In particular $\limsup_n |\partial f^n|(J_\lambda x) \leqslant C(J_\lambda x) < +\infty$ for all $x \in H$. Therefore there exists some $n_0 \in \mathbb{N}$ such that for all $n \geqslant n_0$ we have $J_\lambda x \in \operatorname{dom} |\partial f^n|$ implying that $f^n(J_\lambda x)$ and $|\partial f^n|(J_\lambda x)$ are finite. By virtue of Lemma 6.9 the following inequality holds for all $n \geqslant n_0$

$$f^n(J^n_\lambda x) \geqslant f^n(J_\lambda x) - |\partial f^n|(J_\lambda x) d(J_\lambda x, J^n_\lambda x)$$

This implies

$$+\infty > \liminf_n f^n(J^n_\lambda x) \geqslant f(J_\lambda x) - \limsup_n |\partial f^n|(J_\lambda x) d(J_\lambda x, J^n_\lambda x) \geqslant -\infty \tag{6.13}$$

But $\limsup_n |\partial f^n|(J_\lambda x) \leqslant C(J_\lambda x) < +\infty$ yields

$$\limsup_n |\partial f^n|(J_\lambda x) d(J_\lambda x, J^n_\lambda x) = \limsup_n |\partial f^n|(J_\lambda x) \cdot \lim_n d(J_\lambda x, J^n_\lambda x) \leqslant C(J_\lambda x) \cdot 0 = 0$$

which together with (6.13) gives

$$+\infty > \liminf_n f^n(J_\lambda^n x) \geqslant f(J_\lambda x) > -\infty \tag{6.14}$$

From inequality (6.14) and (6.12) we obtain $f(J_\lambda x) = \lim_n f^n(J_\lambda^n x)$ as required. □

It is natural to ask if, under some additional condition, the pointwise convergence of f^n to f is a necessary condition for pointwise convergence of Moreau envelopes $(f_\lambda^n)_{n\in\mathbb{N}}$ to f_λ.

Theorem 6.17. *Let (H,d) be a Hadamard space and $f, f^n : H \to (-\infty, +\infty]$ be a sequence of closed convex functions on H. Suppose $(f^n)_{n\in\mathbb{N}}$ has pointwise asymptotically bounded slope on H. If for all $x \in H$, $\lim_n f_\lambda^n(x) = f_\lambda(x)$ then*

(i) $\lim_n J_\lambda^n x = J_\lambda x$

(ii) $\lim_n f^n(x) = f(x)$.

Proof. By Theorem 6.6 assumption $\lim_n f_\lambda^n(x) = f_\lambda(x)$ implies $M-\lim_n f^n(x) = f(x)$ for all $x \in H$. Then Theorem 6.7 yields $\lim_n J_\lambda^n x = J_\lambda x$ for all $x \in H$. This proves i which in turn yields

$$f_\lambda(x) = \lim_n f_\lambda^n(x) = \limsup_n f^n(J_\lambda^n x) + \frac{1}{2\lambda}\lim_n d(J_\lambda^n x, x)^2 = \limsup_n f^n(J_\lambda^n x) + \frac{1}{2\lambda} d(J_\lambda x, x)^2.$$

By definition of Moreau-Yosida then it follows $f(J_\lambda x) = \limsup_n f^n(J_\lambda^n x)$. Similarly $f(J_\lambda x) = \liminf_n f^n(J_\lambda^n x)$ hence $f(J_\lambda x) = \lim_n f^n(J_\lambda^n x)$. On the other hand for each $n \in \mathbb{N}$ we have

$$f^n(J_\lambda^n x) \leqslant f^n(J_\lambda^n x) + \frac{1}{2\lambda} d(J_\lambda^n x, x)^2 \leqslant f^n(x) \Rightarrow \lim_n f^n(J_\lambda^n x) \leqslant \liminf_n f^n(x).$$

Therefore $f(J_\lambda x) \leqslant \liminf_n f^n(x)$ for all $x \in H$ and for all $\lambda > 0$. Using $\lim_{\lambda\downarrow 0} J_\lambda x = x$ and the assumption that f is closed we obtain

$$f(x) \leqslant \liminf_{\lambda\downarrow 0} f(J_\lambda x) \leqslant \liminf_n f^n(x). \tag{6.15}$$

By [11, Lemma 1.18] there exists a mapping $n \mapsto \lambda(n)$ such that $\lim_n \lambda(n) = 0$ and

$$\lim_{\lambda\downarrow 0}\lim_n f_\lambda^n(x) = \lim_n f_{\lambda(n)}^n(x).$$

By definition of Moreau envelope we can write

$$f_{\lambda(n)}^n(x) = f^n(J_{\lambda(n)}^n x) + \frac{1}{2\lambda(n)} d(J_{\lambda(n)}^n x, x)^2$$

implying

$$f(x) \geqslant \lim_n \left[f^n(J_{\lambda(n)}^n x) + \frac{1}{2\lambda(n)} d(J_{\lambda(n)}^n x, x)^2 \right] \geqslant \limsup_n f^n(J_{\lambda(n)}^n x). \tag{6.16}$$

By Lemma 6.9 we have the inequalities

$$f^n(J^n_{\lambda(n)}x) + |\partial f^n|(x)d(J^n_{\lambda(n)}x, x) \geqslant f^n(x), \ \ \forall n \in \mathbb{N} \tag{6.17}$$

which then give

$$\limsup_n f^n(J^n_{\lambda(n)}x) + \limsup_n |\partial f^n|(x)d(J^n_{\lambda(n)}x, x) \geqslant \limsup_n f^n(x). \tag{6.18}$$

By assumption $(f^n)_{n\in\mathbb{N}}$ has pointwise asymptotically bounded slope on H, which implies that for some nonnegative finite valued function $C : H \to \mathbb{R}_+$ we have $\limsup_n |\partial f^n|(x) \leqslant C(x)$. Hence

$$0 \leqslant \limsup_n |\partial f^n|(x)d(J^n_{\lambda(n)}x, x) = \limsup_n |\partial f^n|(x) \cdot \lim_n d(J^n_{\lambda(n)}x, x) \leqslant C(x) \cdot 0 = 0.$$

From inequalities (6.16) and (6.18) it follows

$$f(x) \geqslant \limsup_n f^n(J^n_{\lambda(n)}x) \geqslant \limsup_n f^n(x). \tag{6.19}$$

The inequalities (6.15) and (6.19) imply $f(x) = \lim_n f^n(x)$. □

It was pointed out by Bačak[2] that Thoerem 6.17 (ii) can be proved directly by employing the following two key lemmas.

Lemma 6.18 ([43, Proposition 2.2.26])**.** *Let $f : H \to (-\infty, +\infty]$ be a closed convex function and $x \in H$. Then the function $\lambda \mapsto J_\lambda x$ is continuous on $(0, +\infty)$ and*

$$\lim_{\lambda\downarrow 0} J_\lambda x = P_{\mathrm{cl\,dom}\, f}\, x \tag{6.20}$$

In particular if $x \in \mathrm{cl\,dom}\, f$ then $\lambda \mapsto J_\lambda x$ is continuous on $[0, +\infty)$.

Lemma 6.19 ([43, Proposition 5.1.4])**.** *Let $f : H \to (-\infty, +\infty]$ be a closed convex function. Then for any $x \in H$ and $\lambda \in (0, +\infty)$ we have*

$$\frac{f(x) - f_\lambda(x)}{\lambda} \leqslant \frac{|\partial f|^2(x)}{2} \tag{6.21}$$

Without loss of generality let $x \in \mathrm{cl\,dom}\, f$. From triangle inequality for each $n \in \mathbb{N}$ we have the upper estimate

$$|f^n(x) - f(x)| \leqslant |f^n(x) - f^n_\lambda(x)| + |f^n_\lambda(x) - f_\lambda(x)| + |f_\lambda(x) - f(x)| \tag{6.22}$$

By Lemma 6.19 we have $|f^n(x) - f^n_\lambda(x)| \leqslant \lambda|\partial f^n|^2(x)/2$ and for sufficiently large n assumption $(f^n)_{n\in\mathbb{N}} \in A(H)$ implies $|f^n(x) - f^n_\lambda(x)| \leqslant \lambda C(x)$ for some finite valued function $C(x)$. Hence this term vanishes as $\lambda \downarrow 0$. The middle term in (6.22) vanishes by assumption $\lim_n f^n_\lambda(x) = f_\lambda(x)$ for each $x \in H$. On the other hand Lemma 6.18 implies $\lim_{\lambda\downarrow 0} J_\lambda x = x$. The evident chain of inequalities $f(J_\lambda x) \leqslant f_\lambda(x) \leqslant f(x)$ together with lsc of f imply $|f_\lambda(x) - f(x)| \to 0$ as $\lambda \downarrow 0$.

An application of Theorem 6.7 and Theorem 6.6 yield the following.

[2]private correspondence

Theorem 6.20. *Let (H,d) be a Hadamard space and $f, f^n : H \to (-\infty, +\infty]$ be a sequence of closed convex functions. If $(f^n)_{n\in\mathbb{N}}$ has pointwise asymptotically bounded slope on H, then $M - \lim_n f^n = f$ if and only if $\lim_n f^n(x) = f(x)$ and $\lim_n J_\lambda^n x = J_\lambda x$ for each $x \in H$.*

Proof. Assume $(f^n)_{n\in\mathbb{N}} \in A(H)$ and let $\lim_n f^n(x) = f(x)$ for all $x \in H$. Then by Theorem 6.16 $\lim_n J_\lambda^n x = J_\lambda x$ implies $\lim_n f_\lambda^n(x) = f_\lambda(x)$ for all $x \in H$. Theorem 6.7 in turn yields $M - \lim_n f^n(x) = f(x)$. Now suppose $M - \lim_n f^n(x) = f(x)$ then by Theorem 6.6 we get $\lim_n f_\lambda^n(x) = f_\lambda(x)$ for each $x \in H$. Since by assumption $(f^n)_{n\in\mathbb{N}} \in A(H)$ then Theorem 6.17 implies $\lim_n f^n(x) = f(x)$ and $\lim_n J_\lambda^n x = J_\lambda x$ for all $x \in H$. □

6.3.4 A normalization condition

Let $f^n, f : H \to (-\infty, +\infty]$ be a family of proper closed convex functions. We say the sequence of functions $(f^n)_{n\in\mathbb{N}}$ satisfies the *normalization condition* if there exists some sequence $(x_n)_{n\in\mathbb{N}} \subset H$ and $x \in H$ such that $x_n \to x$, $f^n(x_n) \to f(x)$ and $|\partial f^n|(x_n) \to |\partial f|(x)$ as $n \uparrow +\infty$. For a sequence of functions $(f^n)_{n\in\mathbb{N}}$ that Mosco converges to some function f we get the following result.

Lemma 6.21. *A sequence of closed convex functions $(f^n)_{n\in\mathbb{N}}, f : H \to (-\infty, +\infty]$ satisfies the normalization condition whenever $M - \lim_n f^n = f$.*

Proof. Let $x_0 \in H$ then $M - \lim_n f^n = f$ implies by Theorem 6.6 we have $\lim_n J_\lambda^n x_0 = J_\lambda x_0$ for any $\lambda > 0$. Take $x_n := J_\lambda^n x_0$ and $x := J_\lambda x_0$. Then this means $\lim_n x_n = x$. We need to show the other two properties. Note that by definition of the proximal mapping J_λ we have

$$f^n(x_n) + \frac{1}{2\lambda} d(x_0, x_n)^2 \leqslant f^n(y) + \frac{1}{2\lambda} d(x_0, y)^2, \quad \forall y \in H$$

Let $(\xi_n)_{n\in\mathbb{N}} \subset H$ be a sequence strongly converging to x. From the last inequality we obtain in particular that

$$f^n(x_n) + \frac{1}{2\lambda} d(x_0, x_n)^2 \leqslant f^n(\xi_n) + \frac{1}{2\lambda} d(x_0, \xi_n)^2, \quad \forall n \in \mathbb{N}$$

implying $\limsup_n f^n(x_n) \leqslant \limsup_n f^n(\xi_n)_{n\in\mathbb{N}}$. On the other hand by definition of Mosco convergence we can have $(\xi_n)_{n\in\mathbb{N}}$ such that $\limsup_n f^n(\xi_n) \leqslant f(x)$. Hence $\limsup_n f^n(x_n) \leqslant f(x)$. Moreover $\lim_n x_n = x$ implies in particular that $x_n \xrightarrow{w} x$. Again by definition of Mosco convergence we obtain $f(x) \leqslant \liminf_n f^n(x_n)$. Therefore $f(x) = \lim_n f^n(x_n)$ as desired. Next we need to show the property about the slopes. Note that by Lemma 6.9 we have

$$\frac{\max\{f^n(x_n) - f^n(y), 0\}}{d(x_n, y)} \leqslant |\partial f^n|(x_n), \quad \forall y \in H, \forall n \in \mathbb{N}.$$

Again by Mosco convergence for each $y \in H$ there is a sequence $(\xi_n)_{n\in\mathbb{N}}$ strongly converging to y such that $\limsup_n f^n(\xi_n) \leqslant f(y)$. Applying the last inequality for ξ_n we have

$$\frac{\max\{f^n(x_n) - f^n(\xi_n), 0\}}{d(x_n, y)} \leqslant |\partial f^n|(x_n), \quad \forall n \in \mathbb{N}$$

which in turn yields

$$\frac{\max\{f(x) - \limsup_n f^n(\xi_n), 0\}}{d(x,y)} \leqslant \liminf_n |\partial f^n|(x_n).$$

Using $\limsup_n f^n(\xi_n) \leqslant f(y)$ we get

$$\frac{\max\{f(x) - f(y), 0\}}{d(x,y)} \leqslant \liminf_n |\partial f^n|(x_n)$$

Because the last inequality holds for any $y \in H$ then $|\partial f|(x) \leqslant \liminf_n |\partial f^n|(x_n)$. Now by definition (6.7) we obtain

$$|\partial f^n|(x_n) \leqslant \frac{\max\{f^n(x_n) - f^n(y_n), 0\}}{d(x_n, y_n)} + \varepsilon_n, \quad \forall n \in \mathbb{N}$$

for sufficiently small $\varepsilon_n > 0$ and y_n sufficiently close to x_n. Note that strong convergence of x_n to x implies that for any $\delta > 0$ all but finitely many of the terms $y_n \in \mathbb{B}(x, \delta)$. In particular (y_n) is a bounded sequence hence by Lemma 3.8 it has a weakly convergent subsequence (y_{n_k}). Moreover $y_{n_k} \xrightarrow{w} y \in \operatorname{cl}\mathbb{B}(x,\delta)$. One can choose (ε_n) such that $\lim_k \varepsilon_{n_k} = 0$. Since $d(x, \cdot)$ is weakly lsc then we get

$$\limsup_k |\partial f^{n_k}|(x_{n_k}) \leqslant \frac{\max\{f(x) - \liminf_k f^{n_k}(y_{n_k}), 0\}}{d(x,y)}.$$

By definition of Mosco convergence follows $\liminf_n f^n(y_n) \geqslant f(y)$. Hence

$$\limsup_n |\partial f^n|(x_n) \leqslant \limsup_k |\partial f^{n_k}|(x_{n_k}) \leqslant \frac{\max\{f(x) - f(y), 0\}}{d(x,y)}.$$

The last inequality implies $\limsup_n |\partial f^n|(x_n) \leqslant |\partial f|(x)$. □

A family of functions $f^n : H \to (-\infty, +\infty]$ is said to be *equi locally Lipschitz* if for any bounded set $K \subseteq H$ there is a constant $C_K > 0$ such that

$$|f^n(x) - f^n(y)| \leqslant C_K d(x,y), \quad \forall x, y \in K, \forall n \in \mathbb{N} \tag{6.23}$$

Lemma 6.22. *Let $f^n : H \to (-\infty, +\infty]$ be a sequence of closed convex functions such that $\lim_n f^n_\lambda(x_0) = \alpha_0 \in \mathbb{R}$ for some $x_0 \in H$ and some $\lambda > 0$. Then $(f^n_\lambda)_{n\in\mathbb{N}}$ are equi locally Lipschitz functions.*

Proof. By virtue of [11, Theorem 2.64 (ii)] it suffices to show that there is $r > 0$ and $x_0 \in H$ such that $f^n(x) + r(d(x,x_0)^2 + 1) \geqslant 0$ for all $x \in H$ and all $n \in \mathbb{N}$. Let $x_0 \in H$ be such that $\lim_n f^n_\lambda(x_0) = \alpha_0 \in \mathbb{R}$. Notice that by definition of Moreau envelope we have

$$f^n(x) \geqslant f^n_\lambda(x_0) - \frac{1}{2\lambda} d(x_0, x)^2 \geqslant \alpha_0 - \delta - \frac{1}{2\lambda} d(x_0, x)^2$$

for some $\delta > 0$ and sufficiently large n. If one takes $\delta = \alpha_0 + 1/2\lambda$ then one gets

$$f^n(x) \geqslant -\frac{1}{2\lambda}(d(x_0,x)^2 + 1), \quad \forall x \in H$$

For any $r \geqslant 1/2\lambda$ we obtain $f^n(x) + r(d(x_0,x)^2 + 1) \geqslant 0$ for all $x \in H$ and all $n \in \mathbb{N}$. □

Let $f : H \to (-\infty, +\infty]$. The *geodesic lower directional derivative* of f at $x \in H$ along a geodesic $\gamma \in \Gamma_x(H)$ is defined as

$$f'_-(x;\gamma) := \liminf_{y \overset{\gamma}{\to} x} \frac{f(y) - f(x)}{d(y,x)} \tag{6.24}$$

Analogously the *geodesic upper directional derivative*, denoted by $f'_+(x;\gamma)$, is defined with liminf replaced by limsup. If both limits exist and coincide then we say f is geodesically differentiable at x along $\gamma \in \Gamma_x(H)$ and denote it by $f'(x;\gamma)$ (compare with (3.19)).

Theorem 6.23. *Let $f^n, f : H \to (-\infty, +\infty]$ be a sequence of closed convex functions such that*

(i) $\forall \lambda > 0, \forall x \in H$ it holds $\lim_n J^n_\lambda x = J_\lambda x$

(ii) $(f^n)_{n\in\mathbb{N}}$ satisfies the normalization condition with $(x_n)_{n\in\mathbb{N}}$ such that $x_n \to x_0 \subset H$

(iii) $\lim_n f'_{n,\lambda}(x_t;\gamma) = f'_\lambda(x_t;\gamma)$ for all $\gamma \in \Gamma_{x_0}(H)$ and $x_t \in \gamma$ where $t \in [0,1]$.

Then $\forall \lambda > 0, \forall x \in H$ it holds $\lim_n f^n_\lambda(x) = f_\lambda(x)$.

Proof. Let $(f^n)_{n\in\mathbb{N}}, f$ satisfy the normalization condition. Then there exists $(x_n), x_0 \subset H$ such that $\lim_n x_n = x_0, \lim_n f^n(x_n) = f(x_0)$ and $\lim_n |\partial f^n|(x_n) = |\partial f|(x_0)$. Let $\lambda > 0$. First we claim that $\lim_n f^n_\lambda(x_0) = f_\lambda(x_0)$. Introduce the variables $u_n := J^n_\lambda x_n$ for each $n \in \mathbb{N}$ and $u_0 := J_\lambda x_0$. Note that by assumption (i) for each fixed $m \in \mathbb{N}$ we have $\lim_n J^n_\lambda x_m = J_\lambda x_m$. Since the mapping $x \mapsto J_\lambda x$ is continuous (it is firmly nonexpansive) then $\lim_m J_\lambda x_m = J_\lambda x_0$. By triangle inequality $d(J^n_\lambda x_n, J_\lambda x_0) \leqslant d(J^n_\lambda x_n, J^n_\lambda x_m) + d(J^n_\lambda x_m, J_\lambda x_0)$ and nonexpansiveness of J^n_λ we have

$$d(J^n_\lambda x_n, J_\lambda x_0) \leqslant d(x_n, x_m) + d(J^n_\lambda x_m, J_\lambda x_0).$$

Passing in the limit as $m, n \uparrow +\infty$ we obtain $\lim_n u_n = \lim_n J^n_\lambda x_n = J_\lambda x_0 = u_0$. On the other hand

$$|f^n(u_n) - f(u_0)| \leqslant |f^n(u_n) - f^n(x_n)| + |f^n(x_n) - f(x_0)| + |f(x_0) - f(u_0)|.$$

By normalization condition and using $\lim_{\lambda\downarrow 0} u_n = \lim_{\lambda\downarrow 0} J^n_\lambda x_n = x_n, \lim_{\lambda\downarrow 0} u_0 = \lim_{\lambda\downarrow 0} J_\lambda x_0 = x_0$ and lsc of f^n and f implies in the limit as $\lambda \downarrow 0$ and $n \uparrow +\infty$ that $\lim_n f^n(u_n) = f(u_0)$. Again by definition of Moreau envelope

$$f^n_\lambda(x_n) = f^n(u_n) + \frac{1}{2\lambda} d(x_n, u_n)^2 \to f(u_0) + \frac{1}{2\lambda} d(x_0, u_0)^2 := f_\lambda(x_0), \quad \text{as} \quad n \uparrow +\infty.$$

Note that

$$f^n_\lambda(x_0) \leqslant f^n(x_n) + \frac{1}{2\lambda} d(x_0, x_n)^2 \to f(x_0) \quad \text{as} \quad n \uparrow +\infty.$$

On the other hand we have

$$\begin{aligned} f^n_\lambda(x_0) \geqslant f^n(J^n_\lambda x_0) &\geqslant f^n(x_n) - |\partial f^n|(x_n) d(J^n_\lambda x_0, x_n) \\ &\to f(x_0) - |\partial f|(x_0) d(J_\lambda x_0, x_0) > -\infty \quad \text{as} \quad n \uparrow +\infty. \end{aligned}$$

In particular we obtain that $-\infty < \liminf_n f_\lambda^n(x_0) \leqslant \limsup_n f_\lambda^n(x_0) < +\infty$ (one can assume that $x_0 \in \operatorname{dom} f$ else there is nothing to show). By Lemma 6.22 we get that $(f_\lambda^n)_{n\in\mathbb{N}}$ is equi locally Lipschitz in H. This means that for any bounded domain $K \subseteq H$ there is $C_K > 0$ such that

$$|f_\lambda^n(x) - f_\lambda^n(y)| \leqslant C_K d(x,y), \quad \forall x, y \in K, \forall n \in \mathbb{N}$$

From this and the estimate

$$|f_\lambda^n(x_0) - f_\lambda(x_0)| \leqslant |f_\lambda^n(x_0) - f_\lambda^n(x_n)| + |f_\lambda^n(x_n) - f_\lambda(x_0)| \leqslant C_K d(x_n, x_0) + |f_\lambda^n(x_n) - f_\lambda(x_0)|$$

follows $\lim_n f_\lambda^n(x_0) = f_\lambda(x_0)$. Now define $g_{n,\lambda}(t) := f_\lambda^n(x_t)$ where $x_t := W(x, x_0, t)$ and $x \in H$ is arbitrary. Consider

$$g'_{n,\lambda}(t) := \lim_{s\to 0} \frac{g_{n,\lambda}(t+s) - g_{n,\lambda}(s)}{s}$$

Since f_λ^n is convex for each $n \in \mathbb{N}$ then it is absolutely continuous on every geodesic segment. In particular $g'_{n,\lambda}(t)$ exists almost everywhere on $[0,1]$, it is Lebesgue integrable on the interval $[0,1]$ and satifies

$$f_\lambda^n(x) = f_\lambda^n(x_0) + \int_0^1 g'_{n,\lambda}(t)\, dt \tag{6.25}$$

On the other hand $g'_{n,\lambda}(t) = f'_{n,\lambda}(x_t;\gamma) d(x_0, x)$ where $\gamma \in \Gamma_{x_0}(H)$ connects x_0 with x and $x_t \in \gamma$. Assumption (iii) implies $\lim_n g'_{n,\lambda}(t) = g'_\lambda(t)$ for all $t \in [0,1]$. Moreover equi locally Lipschitz property of $(f_\lambda^n)_{n\in\mathbb{N}}$ implies that $\sup_n g'_{n,\lambda}(t) \leqslant C_K d(x_0, x)$ for any bounded domain K around x_0 and $x \in K$. By Lebesgue dominated convergence theorem we obtain in the limit

$$\lim_n f_\lambda^n(x) = f_\lambda(x_0) + \int_0^1 \lim_n g'_{n,\lambda}(t)\, dt = f_\lambda(x_0) + \int_0^1 g'_\lambda(t)\, dt = f_\lambda(x)$$

□

CHAPTER 7

Firmly Nonexpansive Operators in Hadamard Spaces

7.1. Fixed Point Theory

7.1.1 Some history

Many mathematical problems can be formulated in the following form:

$$\text{solve } Tx = x \tag{7.1}$$

When T is an operator from one space X to itself often the solution x^* to problem (7.1) is called a *fixed point* of T. An old and celebrated theorem of Banach [16], also known as *Banach fixed point theorem*, ensures under certain conditions the existence and uniqueness of a fixed point. This theorem also provides a method for obtaining an approximation to the fixed point. The two main assumptions in Banach's theorem are

(i) the mapping T is *contractive*;

(ii) the mapping T is an operator from a complete metric space X into itself.

In the study of fixed point iterations in metric spaces, both linear and nonlinear settings, it is often assumed that the *fixed point mapping* is contractive. The obvious advantage of such an assumption is that fixed point iterations with such operators *converge globally linearly* to a fixed point. However many interesting operators in several applications do not satisfy assumption i. There are ways getting around this problem. For instance if the operator T is merely only *nonexpansive* then there are at least two effective methods which guarantee that the iterations converge either in a strong or a weak sense to a fixed point. The first method is *Krasnoselski-Mann iteration*, attributed to Krasnoselski [77] and Mann [85], which arises from the following operator acting in a linear metric space

$$\widehat{T}x := (1-\alpha)Tx + \alpha x, \;\; x \in X, \alpha \in (0,1) \tag{7.2}$$

The second method is *Halpern's iteration*, attributed to Halpern [58], which for a given point $x_0 \in X$ is defined as

$$\widehat{T}x := (1-\alpha)Tx + \alpha x_0, \;\; x \in X, \alpha \in (0,1) \tag{7.3}$$

Methods (7.2) and (7.3) were successfully employed in a seminal paper by Browder and Petryshin [38], where they studied fixed points for nonlinear mappings in Hilbert spaces. During the same time in a series of papers [33], [34], [35], [36] Browder used (7.2) and (7.3) to investigate fixed points of noncompact nonlinear operators in Hilbert and Banach spaces and their relations to *variational inequalities*. Note that (7.2) and (7.3) are *convex combinations* of the operator T with identity and a constant operator respectively. In principle both methods are usable in any metric space X that admits a convex structure W. So (7.2) takes the form

$$\widehat{T}x := W(Tx, x, \alpha), \;\; x \in X, \alpha \in (0,1) \tag{7.4}$$

and (7.3) becomes

$$\widehat{T}x := W(Tx, x_0, \alpha), \;\; x \in X, \alpha \in (0,1) \tag{7.5}$$

In particular in a Hadamard space (H, d) both methods are meaningful and applicable where the convex structure W is given by $W(x, y, \alpha) = (1-\alpha)x \oplus \alpha y$ for $x, y \in H$ and $\alpha \in [0, 1]$. Analogue theorems as in linear metric spaces can be obtained also for the case of a Hadamard space (see [43, Theorem 6.2.1, Theorem 6.2.2]). While our list of references about fixed point theory is by no means exhaustive the interested reader can refer to some early classical works both in the setting of linear and nonlinear spaces. To mention a few Halpern [57] studies nonexpansive maps from the unit ball of a real Hilbert space into itself; in a series of papers Kannan [65], [66], [67], [68] extends Banach fixed point theorem to metric spaces not necessarily complete; Takahashi [112], [104] and Shimizu [103] study fixed points of nonexpansive operators in convex metric spaces; Kirk [71], Kirk and Panyanak [70] investigated fixed points in CAT(0) spaces and R-trees.

7.1.2 Departure in Hadamard spaces

There is another way around the problem when an operator T defined on some metric space (X, d) does not satisfy assumption i. We are motivated by the following observation in a Euclidean space $\mathbb{E}$. If $T : \mathbb{E} \to \mathbb{E}$ is some mapping, possibly *expansive*, if the set of its fixed points $\operatorname{Fix} T := \{x \in \mathbb{E} \mid Tx = x\}$ is sufficiently *attractive* as characterized by *metric subregularity* of the operator T with respect to its fixed points $\operatorname{Fix} T$, then for any initial point close enough to $\operatorname{Fix} T$, the sequence $\{Tx_n\}_{n\in\mathbb{N}}$ can be shown to converge linearly to a fixed point (see for instance [84] or [83]). The purpose of the present study is to lay the foundations for the extension of these methods to Hadamard spaces. We follow the framework established in [84] which is built on only two fundamental elements in the Euclidean setting:

(i) pointwise almost α-averaging [84, Definition 2.2];

(ii) and metric regularity [60, Definition 2.1.b].

Almost averaged mappings are, in general, set-valued. In Hadamard spaces, there are several difficulties that arise: first, there is no straight-forward generalization of the averaging property since addition is not defined on Hadamard spaces; and second, multivaluedness, which comes with allowing mappings to be expansive. The issue of multivaluedness introduces technical overhead, but does not, at this early stage, seem to present any conceptual difficulties. We therefore restrict our attention to an appropriate generalization of single-valued, pointwise α-averaged mappings. The main contribution is establishing a calculus for these mappings in Hadamard spaces, showing in particular how the property is preserved under compositions and convex combinations. This is of central importance to splitting algorithms that are built by such convex combinations and compositions. We then apply this theory in the study of cyclic projections, averaged projections, projected proximal mapping and projected gradient flow.

7.2. Hilbert Space Review

Let $(\mathcal{H}, \|\cdot\|)$ be a Hilbert space equipped with its canonical norm $\|\cdot\|$. An operator $T : \mathcal{H} \to \mathcal{H}$ is called *nonexpansive* whenever

$$\|Tx - Ty\| \leqslant \|x - y\|, \qquad \forall x, y \in \mathcal{H}. \tag{7.6}$$

A pointwise localization of this notion was used heavily in [84]: $T : \mathcal{H} \to \mathcal{H}$ is *pointwise nonexpansive* at $y \in D \subset \mathcal{H}$ on D whenever

$$\|Tx - Ty\| \leqslant \|x - y\|, \qquad \forall x \in D. \tag{7.7}$$

This is just pointwise local Lipschitz continuity of T at y with constant 1. This is variously referred to as *one-sided* Lipschitz continuity or *calmness*, respectively with constant 1.

An operator $T : \mathcal{H} \to \mathcal{H}$ is called *firmly nonexpansive* whenever

$$\|Tx - Ty\|^2 \leqslant \langle x - y, Tx - Ty\rangle, \qquad \forall x, y \in \mathcal{H}. \tag{7.8}$$

This too can be relaxed to the pointwise local version whenever the inequality above holds only at $y \in D \subset \mathcal{H}$ for all $x \in D$, in which case we say that T is *pointwise firmly nonexpansive* at y on D. It is clear that (7.8) implies (7.7) as a direct consequence of Cauchy-Schwarz inequality. A mapping T is firmly nonexpansive if and only if the mapping $A = T^{-1} - \mathrm{Id}$ is *monotone* [49, Theorem 3.6]:

$$\langle x - y, Ax - Ay\rangle \geqslant 0 \;\; \forall x, y \in \mathcal{H}. \tag{7.9}$$

In other words, firmly nonexpansive mappings are the *proximal mappings* of monotone mappings, where the proximal mapping of A is defined by $J_A \equiv (A + \mathrm{Id})^{-1}$.

Another important class of mappings are α-*averaged* operators. An operator $T : \mathcal{H} \to \mathcal{H}$ is α-averaged with averaging constant $\alpha \in (0,1)$ whenever the operator $A : \mathcal{H} \to \mathcal{H}$ defined by

$$T := (1-\alpha)\operatorname{Id} + \alpha A \tag{7.10}$$

is nonexpansive. Equivalently, an operator $T : \mathcal{H} \to \mathcal{H}$ is α-averaged on a Hilbert space $\mathcal{H}$ with constant $\alpha \in (0,1)$ if and only if the following inequality holds

$$\|Tx - Ty\|^2 \leqslant \|x-y\|^2 - \frac{1-\alpha}{\alpha}\|(\operatorname{Id} - T)x - (\operatorname{Id} - T)y\|^2, \qquad \forall x, y \in \mathcal{H}. \tag{7.11}$$

Rearranging terms yields the equivalent characterization

$$\|Tx - Ty\|^2 + (1-2\alpha)\|x-y\|^2 \leqslant 2(1-\alpha)\langle x-y, Tx-Ty\rangle, \qquad \forall x, y \in \mathcal{H}. \tag{7.12}$$

The mapping T is only *pointwise* α-averaged at y on $D \subset \mathcal{H}$ when the corresponding operator A in (7.10) is only pointwise nonexpansive at y on D, or, equivalently, when inequalities (7.11) and (7.12) hold only at $y \in D$ for all $x \in D$.

Note that when $\alpha = 1/2$, (7.12) is just the inequality defining firmly nonexpansive mappings. Denoting the set of firmly nonexpansive operators by $F(\mathcal{H})$ and the set of α-averaged operators with constant α by $A_\alpha(\mathcal{H})$ we thus have

$$A_{1/2}(\mathcal{H}) = F(\mathcal{H}), \tag{7.13}$$

i.e. an operator T is firmly nonexpansive when it is α-averaged with parameter $\alpha = 1/2$.

In a Hilbert space it is known that convex combinations of α-averaged operators are also α-averaged, though not necessarily with the same constant. In particular, if $\{T_1, T_2, \dots, T_I\}$ is a finite collection of α-averaged operators from a subset of $\mathcal{H}$ to $\mathcal{H}$, each with constant $\alpha_i \in (0,1)$, then the convex combination $T \equiv \sum_{i=1}^I \lambda_i T_i$ ($\lambda_i \in [0,1]$, and $\sum_{i=1}^I \lambda_i = 1$) is α-averaged with constant $\alpha = \max_i\{\alpha_i\}$ (see, for instance, [44, Proposition 4.30]). Similarly, compositions of finite collections of α-averaged operators are α-averaged, though not with the same constant [44, Proposition 4.32] : $T \equiv T_1 \circ T_2 \circ \cdots \circ T_I$ is α-averaged with constant

$$\alpha = \frac{I}{I - 1 + \frac{1}{\max_{i \in I}\{\alpha_i\}}}. \tag{7.14}$$

7.3. Firmly Nonexpansive Operators

7.3.1 Discrepancy of an operator

The notion of nonexpansiveness (7.7) applies equally well in Hadamard spaces as for linear spaces. Averagedness is not so straight-forward. To generalize the notion of averaged operators, the expression (7.10) would be written as

$$T = W(A, \operatorname{Id}, \alpha) \ \text{ for some } \alpha \in (0,1) \tag{7.15}$$

where $W(A, \mathrm{Id}, \alpha) := (1-\alpha)\,\mathrm{Id} \oplus \alpha A$ and A is a nonexpansive operator as defined by (7.7) with the norm replaced by the metric d. However this definition in a Hadamard space is ambiguous since for each $x \in H$ it requires extension of geodesics beyond point Tx and, moreover, even if such an extension exists, it is not generally unique. The characterizations of firmly nonexpansive mappings (7.8) and α-averaged mappings (7.12) are specialized to linear spaces, though there are notions of an 'inner product' for nonlinear spaces which formally capture the idea. Note, however, that given four points x, y, u, v in a Hilbert space $\mathcal{H}$ we have the identity

$$\langle x-y, u-v\rangle = \frac{1}{2}(\|x-v\|^2 + \|y-u\|^2 - \|x-u\|^2 - \|y-v\|^2). \tag{7.16}$$

The expression on the right side of (7.16) does not require any linear structure, and so makes sense for a Hadamard space when written in terms of the metric. Indeed, given four points $x, y, u, v \in H$, we define the mapping $\Delta : H \times H \times H \times H \to \mathbb{R}$ by

$$\Delta(x, y, u, v) := \frac{1}{2}(d(x,v)^2 + d(y,u)^2 - d(x,u)^2 - d(y,v)^2). \tag{7.17}$$

This object was used in a quasilinearization of the metric $d(\cdot,\cdot)$ in a metric space (M, d) by Berg and Nikolaev [24]. In the context of characterizing the regularity of a mapping T, $u = Tx$ and $v = Ty$; for convenience we denote $\Delta_T(x, y) := \Delta(x, y, Tx, Ty)$ and refer to this as the *discrepancy of the operator T at $x, y \in H \times H$*. It follows by definition (7.17) that Δ_T is symmetric i.e. $\Delta_T(x, y) = \Delta_T(y, x)$. Moreover by continuity of the metric Δ_T is continuous on $H \times H$ whenever T is a continuous mapping. Also straightforward calculations show that Δ_T vanishes whenever T is a constant map and equals $d(x, y)^2$ when T coincides with the identity operator Id.

7.3.2 α-firmly nonexpansive operators

A natural way to define firm nonexpansiveness of an operator T in a Hadamard space (H, d) is in terms of the discrepancy of T:

$$d(Tx, Ty)^2 \leqslant \Delta_T(x, y) \qquad \forall x, y \in H. \tag{7.18}$$

When this holds only pointwise at y on a neighborhood $D \subset H$ of y we write

$$d(Tx, Ty)^2 \leqslant \Delta_T(x, y) \qquad \forall x \in D. \tag{7.19}$$

This provides for a natural extension of monotonicity: an operator $T : H \to H$ is monotone whenever

$$\Delta_T(x, y) \geqslant 0, \qquad \forall x, y \in H. \tag{7.20}$$

If this inequality holds strictly then we say T is strictly monotone. From these definitions it follows that if T is firmly nonexpansive then T is a monotone operator.

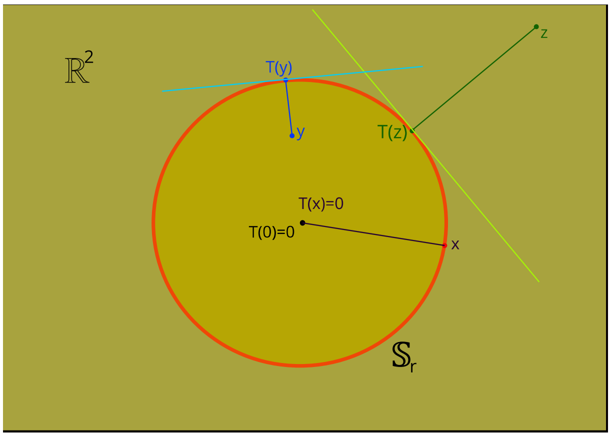


Figure 7.1: Let $\mathbb{S}_r$ and $\mathbb{B}(0,r)$ denote the circle and the open ball of radius $r > 0$ respectively centered at the origin 0 in the Euclidean plane $\mathbb{R}^2$. The operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined as $T(x) = 0$ for $x \in \mathbb{S}_r \cup \{0\}$, $T(x) = P_{\mathbb{S}_r} x$ for $x \in \mathbb{R}^2 \setminus \operatorname{cl} \mathbb{B}(0,r)$, and $T(x) = rx/\|x\|$ otherwise, is firmly nonexpansive at 0 on the complement of the open ball $\mathbb{B}(0,r)$, i.e. $D := \{0\}$ and $E := \mathbb{R}^2 \setminus \mathbb{B}(0,r)$.

In a Hilbert space it is an immediate consequence of the Cauchy-Schwarz inequality that firmly nonexpansive operators are nonexpansive. In the context of metric spaces the Cauchy-Schwarz inequality corresponds to the inequality

$$\Delta(x, y, u, v) \leqslant d(x, y)d(u, v). \tag{7.21}$$

However (7.21) does not hold for general metric spaces. Restricting ourselves to metric spaces where (7.21) holds – call these Cauchy-Schwarz metric spaces – then thanks to [24, Theorem 1, Corollary 3], a metric space is Cauchy-Schwarz if and only if it is a CAT(0) space. Since a Hadamard space is a complete CAT(0) space then (7.21) holds. It follows, then that firm nonexpansiveness implies nonexpansiveness in Hadamard spaces.

The characterization of α-averaged mappings (7.12) can be extended to a general Hadamard space (H, d) by using its metric and the discrepancy Δ_T.

Definition 7.1. *Let (H, d) be a Hadamard space and $D, E \subseteq H$. An operator $T : H \to H$ is pointwise α-firmly nonexpansive at $y \in D$ on E if there exists an $\alpha \in (0, 1)$, possibly depending on y, such that*

$$d(Tx, Ty)^2 + (1 - 2\alpha)d(x, y)^2 \leqslant 2(1 - \alpha)\Delta_T(x, y), \qquad \forall x \in E \tag{7.22}$$

If (7.22) holds for all $y \in D$ with the same α and $D = E$ then T is α-firmly nonexpansive on D. If $D = E = H$ the mapping T is simply said to be α-firmly nonexpansive. If the set of fixed points $\operatorname{Fix} T \neq \emptyset$ and $D = \operatorname{Fix} T, E = H$ then T is said to be a quasi α-firmly nonexpansive operator.

The sets D and E need not have nonempty intersection. Figure 7.1 shows such an example. In the special case when $D \subseteq E$ then the operator T is nonexpansive on D and therefore continuous on D. The relation (7.13) holds in Hadamard spaces as well. Indeed, from

(7.22) it follows that an operator $T : H \to H$ is α-firmly nonexpansive with $\alpha = 1/2$ if and only if it is firmly nonexpansive. Hence $F(H) = A_{1/2}(H)$.

On the other hand (7.22) can be equivalently written as

$$d(Tx, Ty)^2 \leqslant d(x, y)^2 - \frac{1-\alpha}{\alpha}(d(x, y)^2 - 2\Delta_T(x, y) + d(Tx, Ty)^2). \tag{7.23}$$

By Cauchy-Schwartz inequality it follows that

$$d(x, y)^2 - 2\Delta_T(x, y) + d(Tx, Ty)^2 \geqslant (d(x, y) - d(Tx, Ty))^2 \geqslant 0.$$

In the particular case when T is a quasi α-firmly nonexpansive operator inequality (7.23) reduces to

$$d(Tx, y)^2 \leqslant d(x, y)^2 - \frac{1-\alpha}{\alpha}d(x, Tx)^2, \quad \forall x \in H. \tag{7.24}$$

With these observations we summarize the above discussion with the following lemma and corollary.

Lemma 7.2. *An α-firmly nonexpansive mapping $T : H \to H$ on a Hadamard space (H, d) is also nonexpansive.*

Corollary 7.3. *An operator $T : H \to H$ is quasi α-firmly nonexpansive if and only if inequality (7.24) holds true.*

7.4. Calculus of α-Firmly Nonexpansive Operators

7.4.1 Preliminary results

Nonexpansiveness is preserved under compositions and convex combinations, as the next result shows.

Proposition 7.4. *The set of all nonexpansive operators in (H, d) is closed under compositions and convex combinations.*

Proof. Let $T, S : H \to H$ be two nonexpansive operators. Define $R := TS$. Then for any $x, y \in H$ we have

$$d(Rx, Ry) = d(TSx, TSy) \leqslant d(Sx, Sy) \leqslant d(x, y)$$

Now let $\lambda \in (0, 1)$ and define $R := (1-\lambda)T \oplus \lambda S$. By Theorem 2.8 in a Hadamard space the metric is jointly convex therefore for any $x, y \in H$ we have

$$d(Rx, Ry) \leqslant (1-\lambda)d(Tx, Ty) + \lambda d(Sx, Sy)$$

On the other hand assumption T and S are nonexpansive yields

$$d(Rx, Ry) \leqslant (1-\lambda)d(x, y) + \lambda d(x, y) = d(x, y)$$

hence $d(Rx, Ry) \leqslant d(x, y)$ as desired. This completes the proof. □

An operator $T : H \to H$ is called *contractive* if there exists some positive number $\delta < 1$ such that

$$d(Tx, Ty) \leqslant \delta d(x, y), \ \ \forall x, y \in H \tag{7.25}$$

It is clear from the definition that a contractive operator is nonexpansive. Moreover routine calculations like those in Proposition 7.4 imply that contractive operators are closed under compositions and convex combinations. More specifically if $T, S : H \to H$ are contractive with contraction parameters δ_1, δ_2 then $R = TS$ (likewise $R = ST$) is a contraction with parameter $\delta = \delta_1\delta_2$. And if $R = W(S, T, \lambda)$ for $\lambda \in (0, 1)$ then R is a contraction with parameter $\delta = (1 - \lambda)\delta_1 + \lambda\delta_2$.

For an operator $T : H \to H$ and $x, y \in H$ define $\phi_T(t) = d(u_t, v_t)$ for $t \in [0, 1]$ where $u_t := W(Tx, x, t)$ and $v_t := W(Ty, y, t)$. There is a notion of firm nonexpansiveness in the literature that employs the function $\phi_T(t)$ according to which an operator $T : H \to H$ is firmly nonexpansive whenever ϕ_T is nonincreasing on $[0, 1]$, see for example Bačak [43] or for its appications in the so called *W-hyperbolic spaces* in Lopez et al. [8]. However this notion of firmly nonexpansive mappings was originally introduced by Bruck [62] in the context of Banach spaces and by Browder [37], under the name firmly contractive, in the setting of Hilbert spaces. To distinguish between Bruck's notion and our definition of firm nonexpansiveness we refer to the former as strong firm nonexpansiveness. This is motivated by the following elementary observation.

Proposition 7.5. *If $\phi_T(t)$ is a nonincreasing function on $[0, 1]$ for all $x, y \in H$ then T is a firmly nonexpansive operator in the sense of* (7.18).

Proof. Assumption $\phi_T(t)$ is a nonincreasing function on $[0, 1]$ for all $x, y \in H$ implies that $\phi_T(1) \leqslant \phi_T(t)$ for all $t \in [0, 1]$. On the other hand strong convexity of the metric implies

$$\begin{aligned}\phi_T^2(t) = d(u_t, v_t)^2 \leqslant (1 - t)^2 d(x, y)^2 + t^2 d(Tx, Ty)^2 \\ + t(1 - t)[d(x, Ty)^2 + d(y, Tx)^2 - d(x, Tx)^2 - d(y, Ty)^2]\end{aligned}$$

Hence we have

$$\phi_T^2(1) = d(Tx, Ty)^2 \leqslant (1 - t)^2 d(x, y)^2 + t^2 d(Tx, Ty)^2 + 2t(1 - t)\Delta_T(x, y)$$

equivalently

$$(1 - t^2)d(Tx, Ty)^2 \leqslant (1 - t)^2 d(x, y)^2 + 2t(1 - t)\Delta_T(x, y)$$

Dividing by $1 - t$ and letting $t \uparrow 1$ we obtain $d(Tx, Ty)^2 \leqslant \Delta_T(x, y)$. □

Let $T : H \to H$ be an operator such that $\operatorname{Fix} T \neq \emptyset$. We say T is quasi nonexpansive if

$$d(Tx, y) \leqslant d(x, y), \ \forall y \in \operatorname{Fix} T, \forall x \in H \tag{7.26}$$

Lemma 7.6. *Let $T : H \to H$ be a quasi nonexpansive operator. Then* $\operatorname{Fix} T$ *is a closed convex set.*

Proof. Clearly $\operatorname{Fix} T$ is nonempty by definition of quasi nonexpansiveness. Let $(x_n)_{n\in\mathbb{N}} \subseteq \operatorname{Fix} T$ be a sequence of fixed points of T converging to some element $x \in X$. Then by triangle inequality and quasi nonexpansiveness we obtain

$$d(Tx, x) \leqslant d(Tx, x_n) + d(x_n, x) \leqslant 2d(x_n, x).$$

In the limit as $n \uparrow +\infty$ we obtain $d(Tx, x) = 0$ and thus $x \in \operatorname{Fix} T$. Therefore $\operatorname{Fix} T$ is a closed set. Now let $x_1, x_2 \in \operatorname{Fix} T$ and $s \in [0, 1]$. Consider the element $x_s := (1-s)x_1 \oplus sx_2$. Then by strong convexity of the metric we get

$$d(Tx_s, x_s)^2 \leqslant (1-s)d(Tx_s, x_1)^2 + sd(Tx_s, x_2)^2 - s(1-s)d(x_1, x_2)^2.$$

Because T is quasi nonexpansive then $d(Tx_s, x_1) \leqslant d(x_s, x_1)$ and $d(Tx_s, x_2) \leqslant d(x_s, x_2)$. Applying once more strong convexity we finally get the upper estimate

$$\begin{aligned} d(Tx_s, x_s)^2 \leqslant{}& (1-s)[sd(x_1, x_2)^2 - s(1-s)d(x_1, x_2)^2] \\ &+ s[(1-s)d(x_1, x_2)^2 - s(1-s)d(x_1, x_2)^2] - s(1-s)d(x_1, x_2)^2 = 0. \end{aligned}$$

Therefore $x_s \in \operatorname{Fix} T$. Since $x_1, x_2 \in \operatorname{Fix} T$ and $s \in [0, 1]$ are arbitrary then $\operatorname{Fix} T$ is a convex set. □

7.4.2 Compositions of α-firmly nonexpansive operators

In this section we show how the composition of two α-firmly nonexpansive operators is again α-firmly nonexpansive. In general this does not hold, but the property does hold pointwise at fixed points of the composite operator, and for many applications this is all that is needed. The next result requires the following quantities:

$$L(x, y) \equiv d(x, y)^2 - 2\Delta_S(x, y) + d(Sx, Sy)^2; \tag{7.27a}$$

$$M(x, y) := d(Sx, Sy)^2 - 2\Delta_T(Sx, Sy) + d(TSx, TSy)^2; \tag{7.27b}$$

$$U(x, y) := \Delta_{TS}(x, y) + d(Sx, Sy)^2 - \Delta_S(x, y) - \Delta_T(Sx, Sy); \tag{7.27c}$$

$$V(x, y) := d(x, y)^2 - 2\Delta_{TS}(x, y) + d(TSx, TSy)^2. \tag{7.27d}$$

Lemma 7.7. *Let $S : H \to H$ be pointwise α-firmly nonexpansive at y on H with constant α_S and let $T : H \to H$ be pointwise α-firmly nonexpansive at Sy on H with constant α_T. Then the composition TS is pointwise α-firmly nonexpansive at y on H with constant*

$$\alpha_{TS} \equiv \frac{\alpha_S + \alpha_T - 2\alpha_S\alpha_T}{1 - \alpha_S\alpha_T} \tag{7.28}$$

whenever

$$\left(\frac{1-\alpha_S}{\tau\alpha_S}\right)^2 L(x, y) + \left(\frac{1-\alpha_T}{\tau\alpha_T}\right)^2 M(x, y) + 2\left(\frac{1-\alpha_S}{\tau\alpha_S}\right)\left(\frac{1-\alpha_T}{\tau\alpha_T}\right)U(x, y) \geqslant 0, \qquad \forall x \in H, \tag{7.29}$$

where

$$\tau \equiv \frac{1-\alpha_S}{\alpha_S} + \frac{1-\alpha_T}{\alpha_T}. \tag{7.30}$$

Proof. Since T is pointwise α-firmly nonexpansive at Sy with constant α_T on H we have

$$d(TSx,TSy)^2 + (1-2\alpha_T)d(Sx,Sy)^2 \leqslant 2(1-\alpha_T)\Delta_T(Sx,Sy) \qquad \forall Sx \in H.$$

This is equivalent to

$$d(TSx,TSy)^2 \leqslant d(Sx,Sy)^2 - \frac{1-\alpha_T}{\alpha_T}[d(Sx,Sy)^2 - 2\Delta_T(Sx,Sy) + d(TSx,TSy)^2)]$$

for all $Sx \in H$. On the other hand, since S is pointwise α-firmly nonexpansive at y on H with constant α_S we have

$$\begin{aligned} d(TSx,TSy)^2 \leqslant d(x,y)^2 &- \frac{1-\alpha_S}{\alpha_S}[d(x,y)^2 - 2\Delta_S(x,y) + d(Sx,Sy)^2)] \\ &- \frac{1-\alpha_T}{\alpha_T}[d(Sx,Sy)^2 - 2\Delta_T(Sx,Sy) + d(TSx,TSy)^2] \end{aligned}$$

for all $x \in H$. A short calculation yields

$$\begin{aligned} \frac{1-\alpha_S}{\tau\alpha_S}L + \frac{1-\alpha_T}{\tau\alpha_T}M = {}& \left(\frac{1-\alpha_S}{\tau\alpha_S}\right)^2 L + \left(\frac{1-\alpha_T}{\tau\alpha_T}\right)^2 M + \\ & 2\left(\frac{1-\alpha_S}{\tau\alpha_S}\right)\left(\frac{1-\alpha_T}{\tau\alpha_T}\right)U + \left(\frac{1-\alpha_S}{\tau\alpha_S}\right)\left(\frac{1-\alpha_T}{\tau\alpha_T}\right)V, \end{aligned}$$

where τ is given by (7.30) and L, M, U and V are given by (7.27). By the Cauchy-Schwarz inequality we have $L, M \geqslant 0$. If inequality (7.29) holds, then

$$\frac{1-\alpha_S}{\tau\alpha_S}L + \frac{1-\alpha_T}{\tau\alpha_T}M \geqslant \left(\frac{1-\alpha_S}{\tau\alpha_S}\right)\left(\frac{1-\alpha_T}{\tau\alpha_T}\right)V.$$

Therefore

$$d(TSx,TSy)^2 \leqslant d(x,y)^2 - \frac{1-\alpha_S}{\alpha_S}L - \frac{1-\alpha_T}{\alpha_T}M \leqslant d(x,y)^2 - \frac{1}{\tau}\left(\frac{1-\alpha_S}{\alpha_S}\right)\left(\frac{1-\alpha_T}{\alpha_T}\right)V.$$

Subsituting for V, τ and α_{TS} we obtain

$$d(TSx,TSy)^2 \leqslant d(x,y)^2 - \frac{1-\alpha_{TS}}{\alpha_{TS}}[d(x,y)^2 - 2\Delta_{TS}(x,y) + d(TSx,TSy)^2].$$

Rearranging terms the last inequality is equivalent to

$$d(TSx,TSy)^2 + (1-2\alpha_{TS})d(x,y)^2 \leqslant 2(1-\alpha_{TS})\Delta_{TS}(x,y) \quad \forall x \in H,$$

as claimed. □

Lemma 7.8. *Let S be pointwise α-firmly nonexpansive and let T be pointwise nonexpansive at all $y \in \operatorname{Fix} T \cap \operatorname{Fix} S \neq \emptyset$ on $\operatorname{Fix} TS$. Then $\operatorname{Fix} TS = \operatorname{Fix} T \cap \operatorname{Fix} S$.*

Proof. The inclusion $\operatorname{Fix} T \cap \operatorname{Fix} S \subseteq \operatorname{Fix} TS$ is obvious. Now let $x \in \operatorname{Fix} TS$ and $y \in \operatorname{Fix} T \cap \operatorname{Fix} S$. There are three mutually exclusive cases. First let $Sx \in \operatorname{Fix} T$ then $Sx = TSx = x$ implies $x \in \operatorname{Fix} T \cap \operatorname{Fix} S$. Second let $x \in \operatorname{Fix} S$ then $x = TSx = Tx$ implies $x \in \operatorname{Fix} T \cap \operatorname{Fix} S$. Finally, let $x \notin \operatorname{Fix} S$ and $Sx \notin \operatorname{Fix} T$. This yields

$$\begin{aligned} d(x,y)^2 = d(TSx,Ty)^2 \leqslant d(Sx,y)^2 = d(Sx,Sy)^2 \\ \leqslant d(x,y)^2 - \frac{1-\alpha}{\alpha}[d(x,y)^2 - 2\Delta_S(x,y) + d(Sx,Sy)^2] \end{aligned}$$

where the first inequality follows from pointwise nonexpansiveness of T, and the second inequality follows from pointwise α-firm nonexpansiveness of S at y with some constant α on $\operatorname{Fix} TS$. Assumption $y \in \operatorname{Fix} S$ and $x \notin \operatorname{Fix} S$ imply

$$d(Sx,Sy)^2 \leqslant d(x,y)^2 - \frac{1-\alpha}{\alpha} d(x,Sx)^2 < d(x,y)^2,$$

but $d(x,y)^2 < d(x,y)^2$ is impossible. Therefore $\operatorname{Fix} TS \subseteq \operatorname{Fix} T \cap \operatorname{Fix} S$. □

Theorem 7.9. *Let S be quasi α-firmly nonexpansive with constant α_S, let T be quasi α-firmly nonexpansive with constant α_T, and let $\operatorname{Fix} T \cap \operatorname{Fix} S \neq \emptyset$. Then the composite operator TS is quasi α-firmly nonexpansive with constant α_{TS} given by* (7.28).

Proof. By Lemma 7.7, it suffices to show that inequality (7.29) holds at all points $y \in \operatorname{Fix} TS$. Note assumption $\operatorname{Fix} T \cap \operatorname{Fix} S \neq \emptyset$ implies by Lemma 7.8 that $\operatorname{Fix} TS = \operatorname{Fix} T \cap \operatorname{Fix} S$. Then for $y \in \operatorname{Fix} TS$, we have $L(x,y) = d(x,Sx)^2, M(x,y) = d(Sx,TSx)^2$ and $2U(x,y) = d(x,Sx)^2 + d(Sx,TSx)^2 - d(x,TSx)^2$, where L, M and U are defined in (7.27). Then from (7.29) it suffices to show that

$$\begin{aligned} \left(\frac{1-\alpha_S}{\tau\alpha_S}\right)^2 d(x,Sx)^2 + \left(\frac{1-\alpha_T}{\tau\alpha_T}\right)^2 d(Sx,TSx)^2 \\ + \left(\frac{1-\alpha_S}{\tau\alpha_S}\right)\left(\frac{1-\alpha_T}{\tau\alpha_T}\right)[d(x,Sx)^2 + d(Sx,TSx)^2 - d(x,TSx)^2] \geqslant 0 \end{aligned}$$

for all $x \in H$, where τ is given by (7.30). If we let $\kappa := \frac{1-\alpha_S}{\alpha_S} / \frac{1-\alpha_T}{\alpha_T}$ then it is equivalent to prove that

$$(\kappa+1) d(x,Sx)^2 + \frac{\kappa+1}{\kappa} d(Sx,TSx)^2 - d(x,TSx)^2 \geqslant 0$$

for all $x \in H$. On the other hand we have the elementary inequality

$$\kappa d(x,Sx)^2 + \frac{1}{\kappa} d(Sx,TSx)^2 \geqslant 2d(x,Sx)d(Sx,TSx), \qquad \forall \kappa > 0$$

which together with the triangle inequality $d(x,Sx) + d(Sx,TSx) \geqslant d(x,TSx)$ imply

$$(\kappa+1) d(x,Sx)^2 + \frac{\kappa+1}{\kappa} d(Sx,TSx)^2 \geqslant (d(x,Sx) + d(Sx,TSx))^2 \geqslant d(x,TSx)^2,$$

for all $x \in H$, which completes the proof. □

7.4.3 Constructing α-firmly nonexpansive operators from nonexpansive maps

Let $S : H \to H$ be a nonexpansive operator and $\lambda \in (0,1)$. Consider the operator $T : H \to H$ as a convex combination of the identity Id and S with parameter λ i.e. $T = W(S, \mathrm{Id}, \lambda)$. For $x, y \in H$ by strong convexity we have the inequality

$$d(Tx,Ty)^2 \leqslant (1-\lambda)^2 d(x,y)^2 + \lambda^2 d(Sx,Sy)^2 \\ - \lambda(1-\lambda)(d(x,Sx)^2 + d(y,Sy)^2 - d(x,Sy)^2 - d(y,Sx)^2)$$

Using the definition for $\Delta_S(x,y)$ and proper rearrangement of terms the last inequality is equivalent to

$$d(Tx,Ty)^2 \leqslant (1-\lambda) d(x,y)^2 + \lambda d(Sx,Sy)^2 \\ - \lambda(1-\lambda)(d(x,y)^2 - 2\Delta_S(x,y) + d(Sx,Sy)^2).$$

By assumption S is a nonexpansive map, so $d(Sx,Sy) \leqslant d(x,y)$ and

$$d(Tx,Ty)^2 \leqslant d(x,y)^2 - \lambda(1-\lambda)(d(x,y)^2 - 2\Delta_S(x,y) + d(Sx,Sy)^2). \tag{7.31}$$

A short argument shows that $\operatorname{Fix} T = \operatorname{Fix} S$. The inclusion $\operatorname{Fix} S \subseteq \operatorname{Fix} T$ is obvious. For the other inclusion, let $x \in \operatorname{Fix} T$ then $0 = d(x,Tx) = \lambda d(x,Sx)$ implies $d(x,Sx) = 0$ whenever $\lambda > 0$, hence $x \in \operatorname{Fix} S$.

Now, for $y \in \operatorname{Fix} S$ the inequality (7.31) reduces to

$$d(Tx,Ty)^2 \leqslant d(x,y)^2 - \lambda(1-\lambda) d(x,Sx)^2 = d(x,y)^2 - \frac{1-\lambda}{\lambda} d(x,Tx)^2. \tag{7.32}$$

But $y \in \operatorname{Fix} S$ means $y \in \operatorname{Fix} T$ and the algebraic expression $d(x,y)^2 - 2\Delta_T(x,y) + d(Tx,Ty)^2$ reduces to $d(x,Tx)^2$. Therefore

$$d(Tx,Ty)^2 \leqslant d(x,y)^2 - \frac{1-\lambda}{\lambda}(d(x,y)^2 - 2\Delta_T(x,y) + d(Tx,Ty)^2), \qquad \forall x \in H, \forall y \in \operatorname{Fix} T$$

implies that T is quasi α-firmly nonexpansive with constant λ on H. Altogether these observations yield

Theorem 7.10. *For any nonexpansive map $S : H \to H$ and all $\lambda \in (0,1)$, the convex combination $T = W(S, \mathrm{Id}, \lambda)$ is quasi α-firmly nonexpansive with constant λ.*

The assumptions on S in Theorem 7.10 could be relaxed to pointwise nonexpansiveness at $y \in \operatorname{Fix} T$, but the salient point here is that the stronger assumption of nonexpansiveness of S does not yield a stronger result for the convex relaxation.

7.4.4 Convex combinations of α-firmly nonexpansive operators

7.4.4.1 Convex combinations of elements

While for two elements $x, y \in H$ their convex combination is well defined by the convex structure W, the lack in general of an additive structure in H makes the concept of *convex combination* of more than two elements somewhat ambiguous. However one way to define convex combinations of more than two elements would be the following. To keep arguments simple say we are given three points $x, y, z \in H$ and three numbers $w_1, w_2, w_3 \in [0, 1]$ such that $w_1 + w_2 + w_3 = 1$. Then the following expressions

$$(w_1 + w_2)\Big(\frac{w_1}{w_1 + w_2}x \oplus \frac{w_2}{w_1 + w_2}y\Big) \oplus w_3 z$$
$$(w_2 + w_3)\Big(\frac{w_2}{w_2 + w_3}y \oplus \frac{w_3}{w_2 + w_3}z\Big) \oplus w_1 x$$
$$(w_1 + w_3)\Big(\frac{w_1}{w_1 + w_3}x \oplus \frac{w_3}{w_1 + w_3}z\Big) \oplus w_2 y$$

seem all reasonable choices for a convex combination of x, y, z. But they are not guaranteed to be equal unless H admits an additive structure like in the case of a Hilbert space. Notice that by construction all three expressions are in $\text{thr}^2\{x, y, z\}$ and hence by Theorem 4.15 all three are in the convex hull $\text{co}\{x, y, z\}$. Now which one to choose as a representative for the convex combination of x, y, z is not obvious. However there is another way to define convex combinations. This method is based on the *barycenter method* or equivalently the Fréchet mean. To be more precise let $x_1, x_2, ..., x_n \in H$ and $w_1, w_2, ..., w_n \in [0, 1]$ such that $\sum_i w_i = 1$. Then we define the convex combination x^* as the solution to the minimization problem

$$\min_{x \in H} F(x) = \sum_{i=1}^{n} w_i d(x, x_i)^2 \tag{7.33}$$

The existence and uniqueness of x^* follows from (7.33) being strongly convex. With some abuse of notation we denote

$$x^* := w_1 x_1 \oplus w_2 x_2 \oplus ... \oplus w_n x_n \tag{7.34}$$

It follows from (7.33) and definition (7.34) that $x^* = W(x_1, x_2, w)$ when only two points x_1, x_2 are to be considered and $w_1 = w, w_2 = 1-w$. The geometric interpretation of (7.34) is as the unique point lying in $\text{cl}\,\text{co}\{x_1, x_2, ..., x_n\}$ which minimizes the functional F (see Proposition 4.23). There is a clear trade off when choosing (7.33), (7.34) for definition of convex combination. While it guarantees uniqueness it takes us to the closure of the convex hull instead of the convex hull itself. However for Hadamard spaces of finite type this is not a problem since the convex hull of any compact set, and in particular of any finite set, is compact. And hence x^* would lie in the convex hull.

7.4.4.2 Convex combinations of operators

Having defined convex combinations of an arbitrary finite set of elements in H then it is easy to define the convex combinations of operators. Let $T_1, T_2, ..., T_n : H \to H$ be a family of operators. An operator $T : H \to H$ is said to be a *convex combination* of $T_1, T_2, ..., T_n$ for a given set of weights $w_1, w_2, ..., w_n \in [0,1]$ if

$$Tx = \arg\min_{y \in H} \sum_{i=1}^{n} w_i d(y, T_i x)^2 \tag{7.35}$$

and we denote it by

$$T := w_1 T_1 \oplus w_2 T_2 \oplus ... \oplus w_n T_n \tag{7.36}$$

Lemma 7.11. *Let T_i be quasi α-firmly nonexpansive with $\alpha_i \in (0,1)$ for $i = 1, 2, ..., n$ and T be given by (7.36) then $\bigcap_i \operatorname{Fix} T_i = \operatorname{Fix} T$ whenever $\bigcap_i \operatorname{Fix} T_i \neq \emptyset$.*

Proof. Assume $\bigcap_i \operatorname{Fix} T_i \neq \emptyset$ and let $x \in \bigcap_i \operatorname{Fix} T_i$ then by (7.35) follows

$$Tx = \arg\min_{y \in H} \sum_{i=1}^{n} w_i d(y, T_i x)^2 = \arg\min_{y \in H} \sum_{i=1}^{n} w_i d(y, x)^2 = \arg\min_{y \in H} d(y, x)^2 = x$$

This shows $\bigcap_i \operatorname{Fix} T_i \subseteq \operatorname{Fix} T$. Now let $x \in \operatorname{Fix} T$ and $y \in \bigcap_i \operatorname{Fix} T_i$. Note that $d(\cdot, T_i x)^2$ is strongly convex with parameter $\mu = 2$ for every $i = 1, 2, ..., n$. Therefore the functional $F(\cdot) := \sum_i w_i d(\cdot, T_i x)^2$ is strongly convex as a finite sum of strongly convex functions with parameter $\sum_i 2w_i = 2$. By virtue of Proposition 6.11 the following inequality holds

$$F(Tx) + d(Tx, y)^2 \leqslant F(y), \ \ \forall y \in H \tag{7.37}$$

In particular since $x \in \operatorname{Fix} T$ then

$$F(x) + d(x, y)^2 \leqslant F(y), \ \ \forall y \in H \tag{7.38}$$

By assumption T_i is quasi α-firmly nonexpansive for every $i = 1, 2, ..., n$. Then Corollary 7.3 implies

$$d(T_i x, y)^2 \leqslant d(x, y)^2 - \frac{1 - \alpha_i}{\alpha_i} d(x, T_i x)^2, \ \ \forall x \in H, \forall y \in \operatorname{Fix} T_i$$

Therefore in aggregate we obtain

$$\sum_{i=1}^{n} w_i d(T_i x, y)^2 \leqslant d(x, y)^2 - \sum_{i=1}^{n} w_i \frac{1 - \alpha_i}{\alpha_i} d(x, T_i x)^2, \ \ \forall x \in H, \forall y \in \bigcap_i \operatorname{Fix} T_i$$

By inequality (7.38) and definition of F we get

$$0 \leqslant F(x) \leqslant - \sum_{i=1}^{n} w_i \frac{1 - \alpha_i}{\alpha_i} d(x, T_i x)^2 \leqslant 0, \ \ \forall x \in \operatorname{Fix} T$$

Therefore $T_i x = x$ for all i implies $\operatorname{Fix} T \subseteq \bigcap_i \operatorname{Fix} T_i$. □

Remark 7.12. *Note that the above lemma still holds under slightly milder condition that the operator T_i be strictly quasi-nonexpansive for every $i = 1, 2, ..., n$, i.e.*

$$d(T_i x, y) < d(x, y), \ \ \forall y \in \operatorname{Fix} T_i, \forall x \in H \setminus \operatorname{Fix} T_i \tag{7.39}$$

Similarly for the operator S in Lemma 7.8 the condition can be relaxed to just strictly quasi-nonexpansive.

Suppose that T_i is quasi α-firmly nonexpansive with some constant $\alpha_i \in (0, 1)$ for $i = 1, 2, ..., n$. We would like to know under what conditions is $T = w_1 T_1 \oplus ... \oplus w_n T_n$ a quasi α-firmly nonexpansive operator.

Theorem 7.13. *Let $T_i : H \to H$ be a family of quasi α-firmly nonexpansive operators with parameters α_i for $i = 1, 2, ..., n$. If $\bigcap_i \operatorname{Fix} T_i \neq \emptyset$ then the operator $T : H \to H$ defined as in* (7.36) *is quasi α-firmly nonexpansive with parameter $\alpha = \max_i\{\alpha_i\}$.*

Proof. Inequality (7.37) implies $d(Tx, y)^2 \leqslant F(y)$ for all $y \in H$. Using the definition of the functional F and assumption T_i is quasi α-firmly for all $i = 1, 2, ..., n$ by Corollary 7.3 we obtain

$$d(Tx, y)^2 \leqslant d(x, y)^2 - \sum_{i=1}^{n} w_i \frac{1 - \alpha_i}{\alpha_i} d(T_i x, x)^2, \ \ \forall y \in \bigcap_i \operatorname{Fix} T_i \tag{7.40}$$

This inequality is meaningful since by assumption $\bigcap_i \operatorname{Fix} T_i \neq \emptyset$. Moreover by Lemma 7.11 we have $\operatorname{Fix} T = \bigcap_i \operatorname{Fix} T_i$. Therefore (7.40) is equivalent to

$$d(Tx, y)^2 \leqslant d(x, y)^2 - \sum_{i=1}^{n} w_i \frac{1 - \alpha_i}{\alpha_i} d(T_i x, x)^2, \ \ \forall y \in \operatorname{Fix} T \tag{7.41}$$

From (7.41) we also get

$$d(Tx, y)^2 \leqslant d(x, y)^2 - \frac{1 - \alpha}{\alpha} \sum_{i=1}^{n} w_i d(T_i x, x)^2, \ \ \forall y \in \operatorname{Fix} T \tag{7.42}$$

where $\alpha := \max_i\{\alpha_i\}$. By definition of F this is the same as

$$d(Tx, y)^2 \leqslant d(x, y)^2 - \frac{1 - \alpha}{\alpha} F(x), \ \ \forall y \in \operatorname{Fix} T \tag{7.43}$$

In (7.37) we have in particular $d(Tx, x)^2 \leqslant F(x)$. Therefore

$$d(Tx, y)^2 \leqslant d(x, y)^2 - \frac{1 - \alpha}{\alpha} d(Tx, x)^2, \ \ \forall y \in \operatorname{Fix} T \tag{7.44}$$

□

7.4.5 Computing the averaged operator

If $F : H \to H$ is some mapping denote by $F^{(n)} := \underbrace{F \circ F \circ ... \circ F}_{n-\text{times}}$ for $n \in \mathbb{N}$. Given a lsc convex function $f : H \to (-\infty, +\infty]$ and $x_0 \in \operatorname{cl}\operatorname{dom} f$ then by virtue of [43, Theorem 5.1.6] the following limit

$$S_t x_0 := \lim_k J^{(k)}_{\frac{t}{k}} x_0 \tag{7.45}$$

exists for every $t \geqslant 0$. The mapping $S_t : \operatorname{cl}\operatorname{dom} f \to \operatorname{cl}\operatorname{dom} f$ is nonexpansive and it is known as the *gradient flow*. In the special case when the function f is of the form

$$f := \sum_{i=1}^{n} f_i \tag{7.46}$$

where each f_i is a lsc convex function then by Lie-Trotter-Kato formula (see [43, Theorem 5.3.7]) the gradient flow S_t satisfies the following limit

$$S_t x_0 = \lim_k (J^n_{\frac{t}{k}} \circ J^{n-1}_{\frac{t}{k}} \circ ... \circ J^1_{\frac{t}{k}})^{(k)} x_0 \tag{7.47}$$

for every $t \geqslant 0$ and $x_0 \in \operatorname{cl}\operatorname{dom} f$. Here $J^i_{\frac{t}{k}}$ denotes the proximal mapping of the function f_i with parameter t/k. Formula (7.47) can be applied in particular to the functional $F(\cdot) = \sum_{i=1}^n w_i d(\cdot, T_i x)^2$ where $\{T_i\}_{i=1}^n$ is a family of quasi α-firmly nonexpansive operators. But the functional F satisfies even stronger conditions, it is continuous and strongly convex with strong convexity parameter $\mu = 2$. An application of [43, Proposition 5.1.15] for the gradient flow of F yields the following inequality

$$d(S_t x, S_t y) \leqslant \exp(-t/2) d(x, y), \ \ \forall x, y \in \operatorname{cl}\operatorname{dom} F, \forall t \geqslant 0 \tag{7.48}$$

In particular inequality (7.48) implies that the mapping S_t is contractive and if $y = Tx$ is the minimizer of the functional F, hence a fixed point of the mapping S_t for every $t \geqslant 0$, then for any $x_0 \in \operatorname{cl}\operatorname{dom} F$ the limit $\lim_t S_t x_0 = Tx$ is valid.

7.5. Cyclic Projections and Other Methods

7.5.1 Cyclic projections

Let $(C_i)_{i=1}^N$ be a family of closed convex sets in a Hadamard space H. Denote by $P_i := P_{C_i}$ the metric projection onto C_i for all $i = 1, 2, ..., N$. Suppose that $\bigcap_i C_i \neq \emptyset$ and let $C := \bigcap_i C_i$. Denote by P_C the metric projection onto C. For a given arbitrary point $x \in H$ the *cyclic projection method* is defined as follows

$$x_0 := x \quad \text{and} \quad x_n := P_{[n]} x_{n-1}, n = 1, 2, 3, ... \tag{7.49}$$

where $[n] := n(\mod N) + 1 \in \{1, 2, ..., N\}$. In particular $x_{nN} = (P_N P_{N-1} ... P_1)^n x_0$. Given a set $S \subseteq H$ and a sequence $(x_n)_{n\in\mathbb{N}}$ in H we say $(x_n)_{n\in\mathbb{N}}$ is *Fejér monotone with respect to* S whenever

$$d(x_n, y) \leqslant d(x_{n-1}, y) \qquad \forall y \in S \tag{7.50}$$

Lemma 7.14. *The sequence $(x_n)_{n\in\mathbb{N}}$ in (7.49) is Fejér monotone with respect to C. In particular it is bounded.*

Proof. Let $y \in C$ then $y \in C_i$ for all $i = 1, 2, ..., N$. Since P_i is nonexpansive for all $i = 1, 2, ..., N$ then

$$d(x_n, y) = d(P_{[n]}x_{n-1}, y) = d(P_{[n]}x_{n-1}, P_{[n]}y) \leqslant d(x_{n-1}, y)$$

From last inequality it follows that $d(x_n, y_0) \leqslant d(x_0, y_0) < +\infty$ for some $y_0 \in C$. Hence $(x_n)_{n\in\mathbb{N}}$ is bounded. □

Proposition 7.15. *The following holds* $\operatorname{Fix} P = \operatorname{Fix} P_C$ *where* $P := P_N P_{N-1} ... P_1$.

Proof. For every $i = 1, 2, ..., N$ from Corollary 2.10 holds

$$d(x, P_i x)^2 + d(P_i x, y)^2 \leqslant d(x, y)^2, \qquad \forall y \in C_i$$

On the other hand, since $\operatorname{Fix} P_i = C_i$, short calculations show that the last inequality is equivalent to

$$d(P_i x, y)^2 \leqslant \Delta_{P_i}(x, y), \qquad \forall x \in H, \forall y \in \operatorname{Fix} P_i$$

Therefore P_i is quasi-firmly nonexpansive operator for every $i = 1, 2, ..., n$. By definition the operator P is a composition of quasi-firmly nonexpansive operators. An iterative application of Theorem 7.9 implies that P is also quasi-firmly nonexpansive and

$$\operatorname{Fix} P = \bigcap_i \operatorname{Fix} P_i = \bigcap_i C_i = C = \operatorname{Fix} P_C$$

□

Theorem 7.16. *Let $(x_n)_{n\in\mathbb{N}}$ be generated by* (7.49)*. Then x_n converges weakly to some element $x^* \in C$. Moreover* $\lim_n d(P_C x_n, x^*) = 0$.

Proof. It suffices to show for the subsequence $(\hat{x}_n)_{n\in\mathbb{N}}$ where $\hat{x}_n := x_{nN}$ for each $n \in \mathbb{N}$. By definition $\hat{x}_n = P^n x_0$ where $P = P_N P_{N-1} ... P_1$. By Proposition 7.15 we have

$$d(P\hat{x}_{n-1}, y)^2 \leqslant d(\hat{x}_{n-1}, y)^2 - d(\hat{x}_{n-1}, P\hat{x}_{n-1})^2, \qquad \forall y \in \operatorname{Fix} P$$

Since $\operatorname{Fix} P = C$ then

$$d(\hat{x}_{n-1}, P\hat{x}_{n-1})^2 \leqslant d(\hat{x}_{n-1}, y)^2 - d(P\hat{x}_{n-1}, y)^2, \qquad \forall y \in C$$

Moreover by Lemma 7.14 $\lim_n d(\hat{x}_n, C) = l$ for some $l \geqslant 0$ and using $\hat{x}_n = P\hat{x}_{n-1}$ yield

$$\lim_n d(\hat{x}_{n-1}, P\hat{x}_{n-1})^2 \leqslant \lim_n d(\hat{x}_{n-1}, y)^2 - \lim_n d(\hat{x}_n, y)^2 = 0$$

hence $\lim_n d(\hat{x}_{n-1}, P\hat{x}_{n-1}) = 0$. On the other hand the sequence $(x_n)_{n\in\mathbb{N}}$ is bounded and in particular so is the sequence $(\hat{x}_n)_{n\in\mathbb{N}}$. Therefore $(\hat{x}_n)_{n\in\mathbb{N}}$ has a weakly convergent subsequence $(\hat{x}_{n_k})_{k\in\mathbb{N}}$, say $\hat{x}_{n_k} \xrightarrow{w} x^*$. Then

$$\limsup_n d(Px^*, \hat{x}_{n_k}) \leqslant \limsup_n (d(Px^*, P\hat{x}_{n_k}) + d(P\hat{x}_{n_k}, \hat{x}_{n_k})) \leqslant \limsup_n d(x^*, \hat{x}_{n_k})$$

The last inequality follows since P is nonexpansive operator and the limit of the second term in the middle vanishes. Moreover this inequality shows that Px^* is an asymptotic center for the subsequence $(\hat{x}_{n_k})_{k\in\mathbb{N}}$. Since asymptotic centers are unique and x^* is also an asymptotic center then $Px^* = x^*$. Therefore $x^* \in \operatorname{Fix} P = C$. Now if $(\hat{x}_{n_m})_{m\in\mathbb{N}}$ is another weakly convergent subsequence, say $\hat{x}_{n_m} \xrightarrow{w} x^{**}$ then similar arguments show that $x^{**} \in C$. But C is a closed convex set and $(\hat{x}_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to C, it follows by [43, Proposition 3.2.6 (iii)] that $x^{**} = x^*$ and consequently the whole sequence satisfies $\hat{x}_n \xrightarrow{w} x^*$. Now by Corollary 2.10 we have

$$d(\hat{x}_m, P_C\hat{x}_m)^2 + d(P_C\hat{x}_m, P_C\hat{x}_n)^2 \leqslant d(\hat{x}_m, P_C\hat{x}_n)^2 \tag{7.51}$$

which together with Fejér monotonicity implies

$$d(P_C\hat{x}_m, P_C\hat{x}_n)^2 \leqslant d(\hat{x}_m, P_C\hat{x}_n)^2 - d(\hat{x}_m, P_C\hat{x}_m)^2 \leqslant d(\hat{x}_n, P_C\hat{x}_n)^2 - d(\hat{x}_m, P_C\hat{x}_m)^2 \tag{7.52}$$

whenever $m \geqslant n$. Taking limit in (7.52) as $m, n \to +\infty$ gives $\lim_{m,n} d(P_C\hat{x}_m, P_C\hat{x}_n) = 0$ hence $(P_C\hat{x}_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in C. Because the set C is closed and hence complete then $P_C\hat{x}_n \to \bar{x}$ for some $\bar{x} \in C$. Again by Corollary 2.10 we have

$$d(\hat{x}_n, P_C\hat{x}_n)^2 + d(P_C\hat{x}_n, x^*)^2 \leqslant d(\hat{x}_n, x^*)^2 \quad \forall n \in \mathbb{N}.$$

This implies $\liminf_{n\to+\infty} d(\hat{x}_n, \bar{x}) \leqslant \liminf_{n\to+\infty} d(\hat{x}_n, x^*)$. Since the sequence $(\hat{x}_n)_{n\in\mathbb{N}}$ is bounded then by Opial's property it follows that $\bar{x} = x^*$. This completes the proof. □

Corollary 7.17. *The sequence $(x_n)_{n\in\mathbb{N}}$ generated by (7.49) converges strongly to some element $x^* \in C$ whenever the underlying Hadamard space (H, d) is locally compact.*

Historical remarks: Cyclic projections algorithm has a long history in mathematics. The basic theorem for the case of two intersecting subspaces in a Hilbert space is due to von Neumann [113], thereafter it was generalized by Halpern [56] for an arbitrary finite number of intersecting subspaces. However this generalization was in fact discovered earlier in 1937 independently by S. Kaczmarz for solving a system of linear equations [63]. This iterative method is also known as *Kaczmarzs algorithm* or *Kaczmarzs method*. The first statement about cyclic projections for a finite number of intersecting closed convex sets in a Hilbert space was proved by Bregman [31]. Baillon and Brezis [14] showed that $(P_C(x_n))_{n\in\mathbb{N}}$ converged in norm to some element of C. Bauschke [21, Theorem 6.2.2(iii)] established in his PhD Thesis that the norm limit of $(P_C(x_n))_{n\in\mathbb{N}}$ was in fact the weak limit of $(x_n)_{n\in\mathbb{N}}$. Later in a series of papers Deutsch and Hundal studied angles between the convex sets [45], the norm of nonlinear operators [46], and regularity of convex sets [47] all these in relation to the rate of convergence for cyclic projections method. In this historical context our Theorem 7.16 can be regarded as an extension of the classical result in Hilbert spaces to the broader class of Hadamard spaces.

7.5.2 Averaged projections

Let $(C_i)_{i=1}^N$ be a family of closed convex sets in a Hadamard space (H, d) with nonempty intersection $C \neq \emptyset$. Let P_i denote the metric projection onto C_i as before and P_C the metric projection onto C. Let

$$P := \frac{1}{N}P_1 \oplus \frac{1}{N}P_2 \oplus ... \oplus \frac{1}{N}P_N \tag{7.53}$$

For a given element $x \in H$ the *averaged projections method* is defined as

$$x_0 := x \quad \text{and} \quad x_n := Px_{n-1}, n = 1, 2, 3, ... \tag{7.54}$$

Note that by definition (7.53) the element Px_{n-1} solves the problem

$$\min_{y \in H} \sum_{i=1}^{N} \frac{1}{N} d(y, P_i x_{n-1})^2, \qquad n = 1, 2, 3, ... \tag{7.55}$$

For each $n \in \mathbb{N}$ the element Px_{n-1} exists and it is unique.

Proposition 7.18. *The operator P defined by* (7.53) *is a mapping from H onto H that is quasi-firmly nonexpansive. Moreover* $\operatorname{Fix} P = C$.

Proof. It is evident by definition that P is a mapping from H onto H. By Theorem 7.13 P is quasi-firmly nonexpansive as a convex combination of quasi-firmly nonexpansive operators with weights $w_i = 1/N$ for all $i = 1, 2, ..., N$. Since by assumptions $C \neq \emptyset$ and P_i is quasi-firmly nonexpansive for all $i = 1, 2, ..., N$ then Lemma 7.11 implies

$$\operatorname{Fix} P = \bigcap_i \operatorname{Fix} P_i = \bigcap_i C_i = C$$

This completes the proof. ☐

Lemma 7.19. *The sequence $(x_n)_{n \in \mathbb{N}}$ in* (7.54) *is Fejér monotone with respect to C. In particular it is bounded.*

Proof. By Proposition 7.18 if $y \in C$ then $y \in \operatorname{Fix} P$. Therefore

$$d(x_n, y) = d(Px_{n-1}, Py) \leqslant d(x_{n-1}, y), \quad \forall n \in \mathbb{N}, \forall y \in C$$

The last inequality follows from P being nonexpansive on $\operatorname{Fix} P$. In particular it follows $d(x_n, y) \leqslant d(x_0, y) < +\infty$ and hence $(x_n)_{n \in \mathbb{N}}$ is bounded. ☐

Theorem 7.20. *Let $(x_n)_{n \in \mathbb{N}}$ be generated by* (7.54). *Then x_n converges weakly to some element $x^* \in C$. Moreover* $\lim_n d(P_C x_n, x^*) = 0$.

Proof. Similar to the proof of Theorem 7.16 but using Proposition 7.18 and Lemma 7.19 instead. ☐

Corollary 7.21. *The sequence $(x_n)_{n \in \mathbb{N}}$ generated by* (7.54) *converges strongly to some element $x^* \in C$ whenever the underlying Hadamard space (H, d) is locally compact.*

Note that in literature (7.53)-(7.54) is also known as the *Cimmino's method* (see for instance Bauschke and Borwein's review on this subject [55] and references therein).

7.5.3 Projected proximal mappings

Let $f : H \to H$ be a proper lsc convex function and $C \subseteq H$ a nonempty closed convex set. Consider the problem

$$\inf_{x \in C} f(x) \tag{7.56}$$

If ι_C denotes the usual indicator function then problem (7.56) could be written as an unconstrained optimization problem

$$\inf_{x \in H} f(x) + \iota_C(x) \tag{7.57}$$

It is known that ι_C is proper lsc convex function whenever C is a nonempty closed convex set. The objective function in (7.57) is a special case of functions f of the form

$$f(x) := \sum_{i=1}^{n} f_i(x) \tag{7.58}$$

where each f_i is a proper lsc convex function. We can therefore apply the splitting method of Bačak with $f_1 = f$ and $f_2 = \iota_C$. Given $x \in H$ and $\lambda > 0$ this leads to the following iteration

$$x_0 := x \tag{7.59}$$
$$y_i := J_\lambda x_{i-1} \tag{7.60}$$
$$x_i := P_C y_i \quad \text{for} \quad i = 1, 2, 3, ... \tag{7.61}$$

This splitting method is also known as *backward-backward method*. For a general case of this method in Hadamard spaces refer to Banert [18]. Note that J_λ is a firmly non-expansive operator whenever f is a convex function. Also projection P_C onto a closed convex set is a firmly nonexpansive operator. It is clear that $\operatorname{Fix} P_C = C$. We claim that $\operatorname{Fix} J_\lambda = \arg\min_{z \in H} f$. Let $x \in \operatorname{Fix} J_\lambda$. By [43, Lemma 2.2.23] the inequality

$$\frac{1}{2\lambda} d(J_\lambda x, y)^2 - \frac{1}{2\lambda} d(x, y)^2 \leqslant f(y) - f_\lambda(x), \quad \forall x, y \in H$$

whenever f is a closed convex function implies $f(x) \leqslant f(y)$ for all $y \in H$ which together with the evident relation $f(x) \geqslant \inf_{y \in H} f(y)$ yields $f(x) = \inf_{y \in H} f(y)$. Hence $x \in \arg\min_{z \in H} f$. Similarly let $x \in \arg\min_{z \in H} f$ then $f(x) \leqslant f(y)$ for all $y \in H$. In particular $f(x) \leqslant f(J_\lambda x)$ which together with the inequality

$$f(J_\lambda x) + \frac{1}{2\lambda} d(J_\lambda x, x)^2 \leqslant f(x)$$

imply $x = J_\lambda x$. Hence $x \in \operatorname{Fix} J_\lambda$. If additionally $\operatorname{Fix} P_C \cap \operatorname{Fix} J_\lambda \neq \emptyset$ then by Lemma 7.8 $\operatorname{Fix} P_C J_\lambda = \operatorname{Fix} P_C \cap \operatorname{Fix} J_\lambda$. Moreover by Theorem 7.9 the operator $P_C J_\lambda$ is pointwise α-firmly nonexpansive on $\operatorname{Fix} P_C J_\lambda$ with $\alpha = 2/3$.

7.5.4 Projected flow

Consider the same problem (7.56). However this time we follow a different method. Let $x_0 \in H$ (assume without loss of generality that $\operatorname{dom} f = H$). For a given $t > 0$ and $n \in \mathbb{N}$ let $J_{\frac{t}{n}} x_0$ be the proximal mapping of f with steplength $\lambda := t/n$. Let

$$J^{(n)}_{\frac{t}{n}} := \underbrace{J_{\frac{t}{n}} \circ J_{\frac{t}{n}} \circ \ldots \circ J_{\frac{t}{n}}}_{\text{n-times}}$$

It is shown in Bačak [43] that the limit

$$S_t x_0 := \lim_n J^{(n)}_{\frac{t}{n}} x_0 \tag{7.62}$$

exists for each $t \geqslant 0$ and every $x_0 \in H$. Moreover the mapping S_t is nonexpansive in H for all $t \geqslant 0$. Let $t > 0$, for a given $x \in H$ define

$$x_0 := x \tag{7.63}$$

$$y_i := S_t x_{i-1} \tag{7.64}$$

$$z_i := (1-\lambda) x_{i-1} \oplus \lambda y_i \tag{7.65}$$

$$x_i := P_C z_i \quad \text{for} \quad i = 1, 2, 3, \ldots \tag{7.66}$$

It is known that $\operatorname{Fix} S_t = \arg\min_{z \in H} f$ for all $t \geqslant 0$. Denote by $S_{\lambda,t} := (1-\lambda)\operatorname{Id} \oplus \lambda S_t$ and $T := P_C S_{\lambda,t}$. By Theorem 7.10 $\operatorname{Fix} S_{\lambda,t} = \operatorname{Fix} S_t$ and $S_{\lambda,t}$ is α-firmly nonexpansive on $\operatorname{Fix} S_t$ with $\alpha = \lambda$. If $\operatorname{Fix} S_t \cap \operatorname{Fix} P_C \neq \emptyset$ then by Lemma 7.8 we have $\operatorname{Fix} T = \operatorname{Fix} S_t \cap \operatorname{Fix} P_C$. Therefore by Theorem 7.9 the operator T is pointwise α-firmly nonexpansive on $\operatorname{Fix} T$ with $\alpha = 1/(2-\lambda)$.

7.6. Metric Regularity

Definition 7.22. *Let (X, d_X) and (Y, d_Y) be two metric spaces. A mapping $\Phi : X \to Y$ is said to be metrically regular on $U \times V$, where $U \subseteq X, V \subseteq Y$, if there exists $\kappa > 0$ such that*

$$d_X(x, \Phi^{-1}(y)) \leqslant \kappa d_Y(y, \Phi(x)), \quad \forall x \in U, \forall y \in V \tag{7.67}$$

Essentially the concept of metric regularity on a set characterizes the stability of mappings at points in their image and has played a central role, implicitly and explicitly, in the convergence analysis of fixed point iterations (see for instance [83], [110], [59] and [1]).

Let $T : H \to H$ be some operator and $\Phi_T : H \to \mathbb{R}_+$ be the *displacement function* defined as $\Phi_T(x) := d(x, Tx)$ for all $x \in H$. Note that $\Phi_T(x) = 0$ if and only if $x \in \operatorname{Fix} T$. When $V = \{0\}$ we have in view of Definition 7.22 that

$$d(x, \operatorname{Fix} T) \leqslant \kappa \Phi_T(x), \quad \forall x \in U \tag{7.68}$$

for some $\kappa > 0$ whenever Φ_T is metrically regular on $U \times \{0\}$. A sequence $(x_k)_{k\in\mathbb{N}}$ is said to be *linearly monotone* relative to $S \subseteq H$ with constant c whenever $d(x_{k+1}, S) \leqslant cd(x_k, S)$ for all $k \in \mathbb{N}$. A sequence $(x_k)_{k\in\mathbb{N}}$ is said to converge *R-linearly* to some element $x^* \in H$ with rate $c \in [0, 1)$ if there exists some $a > 0$ such that $d(x_k, x^*) \leqslant ac^k$ for all $k \in \mathbb{N}$.

Proposition 7.23. *Let $T : H \to H$ be an operator such that $\operatorname{Fix} T$ is nonempty and closed. Let $U \subset H$. If for every $x_0 \in U$ the sequence $x_{k+1} = Tx_k \subset U$ is linearly monotone relative to $\operatorname{Fix} T$ with constant $c < 1$, then the displacement function Φ_T is metrically regular on $U \times \{0\}$ with constant $\kappa = 1/(1 - c)$.*

Proof. By assumption (Tx_k) is linearly monotone relative to $\operatorname{Fix} T$ with constant $c < 1$ whenever $x_0 \in U$. By triangle inequality we have

$$d(x_{k+1}, x_k) \geqslant d(x_k, \operatorname{Fix} T) - d(x_{k+1}, \operatorname{Fix} T) = (1 - c)d(x_k, \operatorname{Fix} T), \quad \forall k \in \mathbb{N}.$$

Rearranging terms yields

$$d(x_k, \operatorname{Fix} T) \leqslant \frac{1}{1-c} d(Tx_k, x_k) = \frac{1}{1-c} \Phi_T(x_k), \quad \forall k \in \mathbb{N}.$$

□

For $\delta > 0$ define the set

$$D_\delta := \bigcup_{x \in \operatorname{Fix} T} \{y \in H : d(x, y) \leqslant \delta\}.$$

The next result shows the interplay between metric regularity, quasi α-firmly nonexpansiveness and local linear convergence of an operator T. It extends to Hadamard spaces a quantitative convergence theorem in the Euclidean settings ([83, Theorem 1]).

Theorem 7.24. *Let $T : H \to H$ be a quasi α-firmly nonexpansive mapping with constant $\alpha \in (0, 1)$ on D_δ. Assume that Φ_T is metrically regular on $(D_\delta \setminus \operatorname{Fix} T) \times \{0\}$ with constant $\kappa > 0$. Then it holds*

$$d(Tx, \operatorname{Fix} T) \leqslant cd(x, \operatorname{Fix} T), \quad \forall x \in D_\delta \tag{7.69}$$

where

$$c := \sqrt{1 - \frac{1-\alpha}{\alpha\kappa^2}}. \tag{7.70}$$

Moreover if $c < 1$ then any sequence $x_{k+1} = Tx_k$ with $x_0 \in D_\delta$ converges R-linearly to $\operatorname{Fix} T$ with rate c.

Proof. By assumption Φ_T is metrically regular on $(D_\delta \setminus \operatorname{Fix} T) \times \{0\}$ with constant $\kappa > 0$. Then (7.68) implies

$$d(x, \operatorname{Fix} T) \leqslant \kappa\Phi_T(x), \quad \forall x \in D_\delta \setminus \operatorname{Fix} T.$$

On the other hand by assumption T is quasi α-firmly nonexpansive with constant $\alpha \in (0,1)$ on D_δ. Then Corollary 7.3 implies

$$d(Tx,y)^2 \leqslant d(x,y)^2 - \frac{1-\alpha}{\alpha} d(x,Tx)^2, \quad \forall x \in D_\delta, \forall y \in \operatorname{Fix} T$$

In particular the last inequality holds for all $x \in D_\delta \backslash \operatorname{Fix} T$. Moreover T is quasi nonexpansive on D_δ. Therefore by Lemma (7.6) $\operatorname{Fix} T$ is a closed convex set. Given $x \in D_\delta \setminus \operatorname{Fix} T$ let $\bar{x} \in P_{\operatorname{Fix} T} x$. From the last inequality we obtain for $y = \bar{x}$

$$d(Tx,\bar{x})^2 \leqslant d(x,\bar{x})^2 - \frac{1-\alpha}{\alpha} d(x,Tx)^2. \tag{7.71}$$

On the other side we have $d(x,\bar{x}) = d(x, \operatorname{Fix} T) \leqslant \kappa \Phi_T(x) = \kappa d(x,Tx)$. From inequality (7.71) it follows

$$d(Tx,\bar{x})^2 \leqslant \left(1 - \frac{1-\alpha}{\alpha\kappa^2}\right) d(x,\bar{x})^2 \Rightarrow d(Tx,\bar{x}) \leqslant c d(x,\bar{x}) \quad c := \sqrt{1 - \frac{1-\alpha}{\alpha\kappa^2}}.$$

Using $d(Tx,\bar{x}) \geqslant d(Tx, \operatorname{Fix} T)$ and $d(x,\bar{x}) = d(x, \operatorname{Fix} T)$ yields relation (7.69). Now let $x_0 \in D_\delta$. From (7.69) the sequence $x_{k+1} = Tx_k$ is linearly monotone relative to $\operatorname{Fix} T$. Suppose that $c < 1$. Denote by $\bar{x}_k \in P_{\operatorname{Fix} T} x_k$ for each $k \in \mathbb{N}$. Then applying inequality (7.71) for each $k \in \mathbb{N}$ yields

$$d(x_{k+1},\bar{x}_k)^2 \leqslant d(x_k,\bar{x}_k)^2 - \frac{1-\alpha}{\alpha} d(x_k,x_{k+1})^2$$

This in turn implies the inequality

$$\sqrt{\frac{1-\alpha}{\alpha}} d(x_k,x_{k+1}) \leqslant d(x_k,\bar{x}_k)$$

On the other hand $d(x_k,\bar{x}_k) = d(x_k, \operatorname{Fix} T) \leqslant c d(x_{k-1}, \operatorname{Fix} T)$ by (7.69). Therefore an iterative application of linear monotonicity and $d(x_0, \operatorname{Fix} T) \leqslant \delta$ yield

$$\sqrt{\frac{1-\alpha}{\alpha}} d(x_k,x_{k+1}) \leqslant \delta c^k, \quad \forall k \in \mathbb{N}$$

For any given natural numbers k, l with $k < l$ an iterative application of the triangle inequality gives the upper estimate

$$\begin{aligned} d(x_k,x_l) &\leqslant d(x_k,x_{k+1}) + d(x_{k+1},x_{k+2}) + ... + d(x_{l-1},x_l) \\ &\leqslant \sqrt{\frac{\alpha}{1-\alpha}} \delta (c^k + c^{k+1} + ... + c^{l-1}) < \sqrt{\frac{\alpha}{1-\alpha}} \delta c^k \sum_{m=1}^{\infty} c^{m-1} = \sqrt{\frac{\alpha}{1-\alpha}} \frac{\delta c^k}{1-c} \end{aligned}$$

Letting $k, l \to +\infty$ one obtains $\lim_{k,l} d(x_k,x_l) = 0$ hence $(x_k)_{k\in\mathbb{N}}$ is a Cauchy sequence. Because H is a complete metric space then $x_k \to x^*$ for some $x^* \in H$. We need to show that $x^* \in \operatorname{Fix} T$. Note that for each $k \in \mathbb{N}$ we have

$$d(x_k,\bar{x}_k) = d(x_k, \operatorname{Fix} T) \leqslant \delta c^k$$

which passing in the limit as $k \to +\infty$ gives $\lim_k d(x_k, \bar{x}_k) = 0$. From the triangle inequality we get

$$d(\bar{x}_k, x^*) \leqslant d(x_k, \bar{x}_k) + d(x_k, x^*)$$

hence $\lim_k d(\bar{x}_k, x^*) = 0$. By construction $(\bar{x}_k)_{k\in\mathbb{N}} \subseteq \operatorname{Fix} T$ and assumption $\operatorname{Fix} T$ is a closed set imply that $x^* \in \operatorname{Fix} T$. Letting $l \to +\infty$ gives

$$\lim_{l\to+\infty} d(x_k, x_l) = d(x_k, x^*) \leqslant ac^k, \quad a := \sqrt{\frac{\alpha}{1-\alpha}\frac{\delta}{1-c}}$$

Therefore $(x_k)_{k\in\mathbb{N}}$ converges R-linearly to $x^* \in \operatorname{Fix} T$ with rate c and constant a. □

Remark 7.25. *In view of Definition 7.1 one can extend the notion of α-firmly nonexpansiveness to any metric space (X, d). However in general α-firmly nonexpansiveness would not imply nonexpansiveness. If (X, d) is a* CAT(0) *space then the implication would hold. Moreover quasi α-firmly nonexpansiveness implies nonexpansiveness in any metric space. As a result Theorem 7.24 holds true in any complete metric space. By the same argument Lemma 7.8 holds in general for any metric space.*

Bibliography

[1] Ioffe A.D., *Metric regularity and subdifferential calculus*, Russ. Math. Surv. **55** (2000), no. 3, 501.

[2] S. Alexander, V. Kapovitch, and A. Petrunin, *An Invitation to Alexandrov Geometry*, 1 ed., Springer, 2019.

[3] A. D. Alexandrov, *A theorem on triangles in a metric space and some of its applications*, Trudy Mat. Inst. Steklova **38** (1951), 5–23.

[4] ———, *Uber eine Verallgemeinerung der Riemannschen Geometrie*, Schriftreihe des Forschinstituts für Mathematik **1** (1957), 33–84.

[5] Ch. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis*, 3 ed., A Hitchhiker's Guide, Springer-Verlag, 2006.

[6] F. Ardila, M. Owen, and S. Sullivant, *Geodesics in CAT(0) cubical complexes*, Adv. Appl. Math. **48** (2012), 142–163.

[7] F. Ardila-Mantilla, *CAT(0) geometry, robotics and society*, arXiv: 0902.0885 (2019).

[8] D. Ariza-Ruiz, L. I. L. Tean, and G. López-Acedo, *Firmly nonexpansive mappings in classes of geodesic spaces*, Trans. Am. Math. Soc. **366** (2014), 4299–4322.

[9] N. Aronszajn, *Neuer Beweis der Streckenverbundenheit vollstandiger konvexer Raumer*, Ergebnisse eines mat. Kolloquiums **4** (1932), 45–56.

[10] H. Attouch, *Famille d'opérateurs maximaux monotones et mésurabilité*, Ann. Mat. Pura Appl. **120** (1979), no. 4, 35–111.

[11] ———, *Variational Convergence for Functions and Operators*, Applicable Mathematics Series, Pitman (Advanced Publishing Program), Boston, MA, 1984.

[12] M. Baćak, *Old and new challenges in hadamard spaces*, arXiv:1807.01355 (2018).

[13] M. Baćak, M. Montag, and G. Steidl, *Convergence of functions and their moreau envelopes on hadamard spaces*, J. Approx. Theory, (C) **224** (2017), 1–12.

[14] J.B. Baillon and H. Brezis, *Une remarque sur le comportement asymptotique des semigroupes non lineaires*, Houston J. Math. **2** (1976), no. 1, 5–7.

[15] W. Ballmann, *Lectures on Spaces of Nonpositive Curvature*, Birkhäuser, 1995.

[16] S. Banach, *Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales*, Fund. Math. **3** (1922), 133–181.

[17] Jósef Banaś and Antonio Martinón, *Some properties of the Hausdorff distance in metric spaces*, Bull. Aust. Math. Soc. **42** (1990), 511–516.

[18] S. Banert, *Backward–backward splitting in Hadamard spaces*, J. Math. Anal. Appl. **414** (2014), no. 2, 656–665.

[19] S. Baratella and Ng. Siu-Ah, *A Nonstandard Proof of the Eberlein-Šmulian Theorem*, Proc. Am. Math. Soc. **131** (2003), no. 10, 3177–3180.

[20] D. Barden, H. Le, and M. Owen, *Central limit theorems for Fréchet means in the space of phylogenetic trees*, Electron. J. Probab. **18** (2013), no. 25, 1–25.

[21] H.H. Bauschke, *Projection algorithms and monotone operators*, Ph.D. thesis, Simon Fraser University, BC, Canada, 1996.

[22] G. Beer, *Topologies on Closed and Closed Convex Sets*, Mathematics and Its Applications, vol. 268, Kluwer Academic Publishers, 1993.

[23] I. Beg and A. Azam, *Fixed points on star-shaped subsets of convex metric spaces*, Indian J. Pure Appl. Math. **18** (1987), no. 7, 594–596.

[24] I. D. Berg and I. G. Nikolaev, *Quasilinearization and curvature of Aleksandrov spaces*, Geom. Dedicata **133** (2008), 195–218.

[25] R. Bergmann, R. Herzog, D. Tenbrinck, and J. Vidal-Núñez, *Fenchel duality theory and a primal-dual algorithm on Riemannian manifolds*, arXiv:1908.02022 (2019).

[26] L. J. Billera, S. P. Holmes, and K. Vogtmann, *Geometry of the space of phylogenetic trees*, Adv. Appl. Math. **27** (2001), 733–767.

[27] L. M. Blumenthal, *Theory and Applications of Distance Geometry*, Oxford University Press, Oxford, 1953.

[28] V.G. Boltyanskii and I.M. Yaglom, *Convex figures and bodies*, Encyclopaedia of elementary mathematics (in Russian) **55** (1966), no. 5, 181–269.

[29] K. Borsuk, *On a metrization of polytopes*, Fund. Math. **47** (1959), 325–341.

[30] R. Bouckaert, Ph. Lemey, M. Dunn, S. J. Greenhill, A. V. Alekseyenko, A. J. Drummond, R. D. Gray, M. A. Suchard, and Q. D. Atkinson, *Mapping the Origins and Expansion of the Indo-European Language Family*, Science **337** (2012), 957–960.

[31] L.M. Bregman, *The method of successive projection for finding a common point of convex sets*, Soviet Math. Dokl. **6** (1965), 688–692.

[32] M. R. Bridson and A. Haefliger, *Metric Spaces of Nonpositive Curvature*, A Series of Comprehensive Studies in Mathematics, vol. 319, Birkhäuser Boston Inc., Boston, 1999.

[33] F. E. Browder, *Fixed-point theorems for noncompact mappings in Hilbert space*, Proc. Natl. Acad. Sci. U.S.A. **53** (1965), 1272–1276.

[34] ______, *Mapping theorems for noncompact nonlinear operators in Banach spaces*, Proc. Natl. Acad. Sci. U.S.A. **54** (1965), no. 2, 1272–1276.

[35] ______, *Mapping theorems for noncompact nonlinear operators in Banach spaces*, Proc. Natl. Acad. Sci. U.S.A. **54** (1965), 1041–1044.

[36] ______, *Existence and approximations of solutions of nonlinear variational inequalities*, Proc. Natl. Acad. Sci. U.S.A. **56** (1966), 1080–1086.

[37] ______, *Convergence theorems for sequences of nonlinear operators in Banach spaces*, Math. Z. **100** (1967), 201–225.

[38] F. E. Browder and W. V. Petryshin, *Construction of fixed points of nonlinear mappings in Hilbert space*, J. Math. Anal. Appl. **20** (1967), 197–228.

[39] D. G. Brown and M. Owen, *Mean and variance of phylogenetic trees*, Syst. Biol. **69** (2020), no. 1, 139–154.

[40] F. Bruhat and J. Tits, *Groupes réductifs sur un corps local. i. donneés radicielles valueés*, Inst. Hautes Études Sci. Publ. Math. **41** (1972), 5–251.

[41] D. Burago, Y. Burago, and S. Ivanov, *A Course in Metric Geometry*, Graduate Studies in Mathematics, vol. 33, Birkhäuser Boston Inc., Boston, 2001.

[42] H. Busemann, *The Geometry of Geodesics*, New York, 1955.

[43] M. Baćak, *Convex Analysis and Optimization in Hadamard Spaces*, vol. 22 of De Gruyter Series in Nonlinear Analysis and Applications, De Gruyter, Berlin, 2014.

[44] P. L. Combettes and I. Yamada, *Compositions and convex combinations of averaged nonexpansive operators*, J. Math. Anal. Appl. **425** (2015), no. 1, 55–70.

[45] F. Deutsch and H. Hundal, *The rate of convergence for the cyclic projections algorithm I: Angles between convex sets*, J. Approx. Theory **142** (2006), 36–55.

[46] ______, *The rate of convergence for the cyclic projections algorithm II: Norms of nonlinear operators*, J. Approx. Theory **142** (2006), 56–82.

[47] ______, *The rate of convergence for the cyclic projections algorithm III: Regularity of convex sets*, J. Approx. Theory **155** (2008), 155–184.

[48] M. P. do Carmo, *Riemannian Geometry*, Birkhäuser, Boston, 1992.

[49] J. Eckstein, *Splitting methods for monotone operators with applications to parallel optimization*, Ph.D. thesis, MIT, Cambridge, MA, 1989.

[50] R. Espínola and A. Fernández-León, *CAT(κ)-spaces, weak convergence and fixed points*, J. Math. Anal. Appl. **353** (2009), 410–427.

[51] J. Felsenstein, *Evolutionary Trees from DNA Sequences:A Maximum Likelyhood Approach*, J. Mol. Evol. **17** (1981), 368–376.

[52] W. M. Fitch and E. Margoliash, *Construction of Phylogenetic Trees*, Science **155** (1967), no. 3760, 279–284.

[53] M. Gromov, *Hyperbolic groups*, vol. 8, Essays in Group Theory. Math. Sci. Res. Inst. Publ., Springer, New York, NY, 1987.

[54] ______, *Metric Structures for Riemannian and non-Riemannian Spaces*, vol. 152, Birkhäuser Boston Inc., Boston, 1999 (Based on the 1981 French original).

[55] J. M. Borwein H. H. Bauschke, *On projection algorithms for solving convex feasibility problems*, SIAM Review **38** (1996), no. 3, 367–426.

[56] I. Halperin, *The product of projection operators*, Acta Sci. Math. (Szeged) **23** (1962), 96–99.

[57] B. Halpern, *Fixed points of nonexpanding maps*, Bull. Am. Math. Soc. **73** (1967), 957–961.

[58] ______, *The kernel of a starshaped subset of the plane*, Proc. Am. Math. Soc. **123** (1969), 692–696.

[59] Bauschke H.H. and Borwein J.M., *On the convergence of von Neumann's alternating projection algorithm for two sets*, Set-Valued Anal. **1** (1993), no. 2, 185–212.

[60] A. D. Ioffe, *Regularity on a fixed set*, SIAM J. Optim. **21** (2011), no. 4, 1345–1370.

[61] J. Jost, *Equilibrium maps between metric spaces*, Calc. Var. Partial Diff. Equations **2** (1994), 173–204.

[62] R. E. Bruck Jr., *Nonexpansive projections on subsets of Banach spaces*, Pac. J. Math. **47** (1973), 341–355.

[63] S. Kaczmarz, *Angenäherte Auflosung von Systemen linearer Gleichungen*, Bull. Int. Acad. Pol. Sci. Let. **A35** (1937), 355–357.

[64] B. A. Kakavandi, *Weak topologies in complete CAT(0) metric spaces*, Proc. Am. Math. Soc. **141** (2012), 1029–1039.

[65] R. Kannan, *Some results in fixed points*, Bull. Cal. Math. Soc. **60** (1968), 71–76.

[66] ______, *Some results in fixed points II*, Am. Math. Mon. **76** (1970), 71–76.

[67] ______, *Some results in fixed points III*, Fund. Math. Mon. **70** (1971), 169–177.

[68] ______, *Some results in fixed points IV*, Fund. Math. Mon. **74** (1972), 181–187.

[69] M. Kell, *Uniformly convex metric spaces*, Anal. Geom. Metr. Spaces **2** (2014), 359–380.

[70] W. A. Kirk, *Krasnoselskii's iteration process in hyperbolic space*, Numer. Funct. Anal. Optim. **4** (1981/82), 371–381.

[71] ______, *Fixed point theorems in CAT(0) spaces and R-trees*, J. Fixed Point Theory Appl. (2004), 309–316.

[72] W.A. Kirk and B. Panyanak, *A concept of convergence in geodesic spaces*, Nonlinear Anal. **68** (2008), 3689–3696.

[73] U. Kohlenbach, *Some logical metatheorems with applications in functional analysis*, Trans. Am. Math. Soc. **357** (2005), 89–128.

[74] E. Kopecká and S. Reich, *Nonexpansive retracts in Banach spaces*, J. Fixed Point Theory Appl. **77** (2007), 161–174.

[75] B. Krakus, *Some properties of convex metric spaces*, Fund. Math. **73** (1972), 193–198.

[76] ______, *An example of a strongly convex metric space*, Ark. Mat. (1974), no. 1-2.

[77] M. A. Krasnoselskij, *Two remarks on the method of successive approximations*, Uspehi. Mat. Nauk (N.S.) **10** (1955), 123–127.

[78] K. Kuratowski, *Topology*, Academic Press, 1966.

[79] A. Lalek and W. Nitka, *On convex metric spaces I*, Fund. Math. **49** (1961), 183–204.

[80] T.C. Lim, *Remarks on some fixed point theorems*, Proc. Am. Math. Soc. **60** (1976), 179–182.

[81] A. Lubiw, D. Maftuleac, and M. Owen, *Shortest paths and convex hulls in 2D complexes with non-positive curvature*, Comput. Geom. **89** (2020), 51–91.

[82] D. R. Luke, *Numerical Variational Analysis*, Georg-August-Universität Göttingen, 2019.

[83] D. R. Luke, M. Teboulle, and N. H. Thao, *Necessary conditions for linear convergence of iterated expansive, set-valued mappings*, Math. Program **180** (2020), 1–31.

[84] D. R. Luke, N. H. Thao, and M. K. Tam, *Quantitative convergence analysis of iterated expansive, set-valued mappings*, Math. Oper. Res. **43** (2018), no. 4, 1143–1176.

[85] W. R. Mann, *Mean value methods in iteration*, Proc. Am. Math. Soc. **4** (1953), 506–510.

[86] G. Dal Maso, *An Introduction to* Γ*-convergence*, Birkhäuser, Basel, 1993.

[87] B. Mendelson, *Introduction to Topology*, 3rd ed., Allyn and Bacon, Inc. Boston, 1975.

[88] K. Menger, *Untersuchen über allgemeine metric*, Math. Ann. **100** (1928), 75–163.

[89] E. Miller, M. Owen, and J. S. Provan, *Averaging metric phylogenetic trees*, Adv. Appl. Math. **68** (2012), 51–91.

[90] N. Monod, *Superrigidity for irreducible lattices and geometric splitting*, J. Am. Math. Soc. **19** (2008), 781–814.

[91] U. Mosco, *Convergence of convex sets and of solutions of variational inequalities*, Adv. Math. **3** (1969), no. 4, 510–585.

[92] J. R. Munkres, *Topology*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2000.

[93] M. Owen, *Computing geodesic distances in tree space*, SIAM J. Discret. Math. **25** (2011), 1506–1529.

[94] M. Owen and S. Provan, *A fast algorithm for computing geodesic distance in tree space*, IEEE/ACM Trans. Comput. Biol. Bioinform **8** (2011), 2–13.

[95] S. Reich, *The asymptotic behavior of a class of nonlinear semigroups in the Hilbert Ball*, J. Math. Anal. Appl. **157** (1991), 237–242.

[96] ______, *The alternating algorithm of von Neumann in the Hilbert ball*, Dyn. Syst. Appl. **2** (1993), 21–25.

[97] S. Reich and I. Shafrir, *Nonexpansive iterations in hyperbolic spaces*, Nonlinear Anal. **15** (1990), 537–558.

[98] J. G. Reshetnyak, *Non-expansive maps in a space of curvature no greater than K*, Sibirsk. Mat. Z. **9** (1968), 918–927.

[99] R. T. Rockafellar, *Convex Analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, N.J., 1970.

[100] R. T. Rockafellar and R. J. Wets, *Variational Analysis*, Grundlehren Math. Wiss., Springer-Verlag, Berlin, 1998.

[101] W. Rudin, *Principles of Mathematical Aanalysis*, 3 ed., Springer, 1976.

[102] T. Sakai, *Riemannian Geometry*, Translations of mathematical monographs, vol. 149, American Mathematical Society, Providence, R.I., 1996.

[103] T. Shimizu, *Fixed points of multivalued nonexpansive mappings in certain convex metric space*, Keynotes of the Institute of Mathematics, Kyoto **1031** (1998), 75–84.

[104] T. Shimizu and W. Takahashi, *Fixed point theorems in certain convex metric spaces*, Math. Japon. **37** (1992), 855–859.

[105] E.N. Sosov, *On analogues of weak convergence in a special metric space*, Russian Math. (Iz. VUZ) **48** (2004), no. 5, 79–83.

[106] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, 3rd ed., vol. 1, Publish or Perish, Houston, TX, 1999.

[107] A. H. Stone, *Paracompactness and product spaces*, Bull. Amer. Math. Soc. **54** (1948), 977–982.

[108] K.T. Sturm, *Nonlinear martingale theory for processes with values in metric spaces of nonpositive curvature*, Ann. Probab. **30** (2002), 1195–1222.

[109] ______, *Probability Measures on Metric Spaces of Nonpositive Curvature*, Contemporary Mathematics, 2003.

[110] Aspelmeier T., Charitha C., and Luke D.R., *Local linear convergence of the ADMM/Douglas–Rachford algorithms without strong convexity and application to statistical imaging*, SIAM J. Imaging Sci. **9** (2016), no. 2, 842–868.

[111] W. Takahashi, *Fixed point theorem for amenable semigroup of nonexpansive mappings*, Kodai Math. Sem. Rep. **21** (1969), 383–386.

[112] ______, *A convexity in metric space and nonexpansive mappings, I*, Kodai Math. Sem. Rep. **22** (1970), 142–149.

[113] J. von Neumann, *Functional Operators, The Geometry of Orthogonal Spaces, Annals of Mathematical Studies*, vol. 2, Princeton University Press, Princeton, NJ, 1950 (This is a reprint of mimeographed lecture notes first distributed in 1933.).

[114] R.J. Whitley, *An elementary proof of the Eberlein-Šmulian theorem*, Math. Ann. **172** (1967), no. 2, 116–118.

[115] ______, *The Krein-Šmulian Theorem*, Proc. Am. Math. Soc. **97** (1986), no. 2, 376–377.

[116] K. Yosida, *Functional Analysis*, 4 ed., Die Grundlehren der mathematischen Wissenschaften, vol. 123, Springer-Verlag, 1974.

Curriculum Vitae

Arian Bërdëllima

Albrecht-Thaer-Weg 12a, 37075 Göttingen
Email: arian.berdellima@mathematik.uni-goettingen.de

Personal Information

Birth date	05.12.1988
Birth place	Elbasan, Albania
Nationality	Albanian

Education

Since 09/2017	**Doctoral studies in Mathematical Sciences** Georg-August Universität Göttingen Promotionsschule GAUSS
11/2016-08/2017	**Pre-Doctoral studies in Mathematical Sciences** Georg-August Universität Göttingen Institute of Numerical and Applied Mathematics
10/2012-06/2014	**M.Phil. in Economics** University of Oxford Major field of study: Economics and Game Theory Grade: Pass
09/2007-05/2011	**B.A. in Mathematics** American University in Bulgaria Major field of study: Mathematics Grade: 3.96
09/2007 – 05/2011	**B.A. in Economics** American University in Bulgaria Major field of study: Economics Grade: 3.70
09/2003-06/2007	**Highschool** Mehmet Akif College Advanced courses: mathematics, physics, chemistry, english Grade: 9.96

Work experience

Sep. 2017-Current:	PhD candidate at Georg-August Universität Göttingen Institute of Numerical and Applied Mathematics;
Mar. 2015-Jun. 2016:	Adjunct Lecturer in Mathematics at Epoka University Department of Electronics and Computer Engineering Tirane, Albania;
Winter 2015:	Intern at Bank of Albania, Department of Monetary Policy Tirane, Albania;
Sep. 2013-Jun. 2014:	Guest Teacher in Economics at London School of Economics and Political Science, Department of Economics London, UK;
Summer 2013/2014:	Maths and Engineering Teacher at Oxbridge Academic Programs Oxford, UK;

Research interests

Convex and Variational Analysis and their Applications;
Metric Geometry Theory and its Applications;
Non-linear Analysis;
Theory of Functions.

Publications

A. Bërdëllima, F. Lauster and D. R. Luke. Firmly nonexpansive operators in nonlinear spaces. *preprint* (2020)

A. Bërdëllima, D. R. Luke and M. Wardetzky. Weak topology in Hadamard spaces. *preprint*, (2020)

A. Bërdëllima. On a theorem about Mosco convergence. *preprint*, (2020)

A. Bërdëllima. On a notion of averaged operators in CAT(0) spaces. *preprint*, (2020)

A. Bërdëllima. A note on convex hulls and the Fréchet mean in phylogenetic tree spaces. *preprint*, (2020)

A. Bërdëllima. On Khabibullin's conjecture about pair of integral inequalities. *Ufimsk. Mat. Zh.*, 10:3 (2018), 121–134

A. Bërdëllima. A note on a conjecture by Khabibullin. *Zap. Nauchn. Sem. POMI*, 1467 (2018), 7–20; *J. Math. Sci. (N. Y.)*, 243:6 (2019), 825–834

A. Bërdëllima. About a conjecture regarding plurisubharmonic functions. *Ufimsk. Mat. Zh.*, 1:4 (2015), 160–171

Participation in Workshops and Lecture Series

Winter 2018/2019:	Fourth Central-European Set-Valued and Variational Analysis Meeting (CESVVAM 2018), Philipps-Universität Marburg;

Jan. 2019-Feb.2019:	Lecture Series-“Modern Maximal Monotone Operator Theory: From Nonsmooth Optimization to Differential Inclusions”, Erwin Schrödinger Institute (ESI), Vienna, Austria;
17-18 July 2019:	RTG Workshop-“Discovering structure in complex data: Statistics meets Optimization and Inverse Problems”, Göttingen.