Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

Alternating projections CAT(0) spaces

by

Miroslav Bačák, Ian Searston, and Brailey Sims

Preprint no.: 51 2011

ALTERNATING PROJECTIONS IN CAT(0) SPACES

MIROSLAV BAČÁK, IAN SEARSTON, AND BRAILEY SIMS

Dedicated to Jon Borwein on the occasion of his 60th birthday.

Abstract. By using recently developed theory which extends the idea of weak convergence into CAT(0) space we prove the convergence of the alternating projection method for convex closed subsets of a CAT(0) space. Given the right notion of weak convergence it turns out that the generalization of the well-known results in Hilbert spaces is straightforward and allows the use of the method in a nonlinear setting. As an application, we use the alternating projection method to minimize convex functionals on a CAT(0) space.

1. Introduction

The alternating projection method in Hilbert space, which originated in the early 1930s from work by von Neumann, has flourished enormously during the last two decades. It has given rise to both a beautiful theory and a number of useful algorithms, see for instance [2, 3, 5, 6, 7, 8] and references therein. There is also a fruitful connection to other well-known algorithms (like the proximal point algorithm), see [4, 8]. In this paper, we show that the underlying linear structure of the space is dispensable and that the whole machinery works also in metric spaces, namely in CAT(0) spaces, which include Hilbert spaces, classical hyperbolic spaces, simply connected Riemannian manifolds of non-positive sectional curvature, $\mathbb{R}$-trees and Euclidean buildings. (Another important CAT(0) space will appear in Example 5.1.) Let us state the main result of this paper here. We refer the reader to Section 2 for the notation and terminology.

Main result (See Theorem 4.1 below)**.** *Let X be a complete CAT(0) space and $A, B \subset X$ convex closed subsets such that $A \cap B \neq \emptyset$. Let $x_0 \in X$ be a starting point and $(x_n) \subset X$ be the sequence generated by Algorithm* (2)*. Then:*

- (i) *(x_n) weakly converges to a point $x \in A \cap B$.*
- (ii) *If A and B are boundedly regular, then $x_n \to x$.*
- (iii) *If A and B are boundedly linearly regular, then $x_n \to x$ linearly.*
- (iv) *If A and B are linearly regular, then $x_n \to x$ linearly with a rate independent of the starting point.*

The alternating projection algorithm in Hilbert spaces is not only an interesting mathematical object, but also plays a key role in optimization (for instance, in convex feasibility problems) and has found many applications outside mathematics,

Date: June 18, 2011.

1991 *Mathematics Subject Classification.* Primary: 46C50, 53C23. Secondary: 47H09, 49J27.

Key words and phrases. Non-positive curvature, CAT(0) space, alternating projections, feasibility problem, nearest point mapping, non-expansive mapping, weak convergence, convex optimization, algorithm.

such as in medical imaging [6]. The results of the present paper allow the use of the alternating projection method in a much more general setting where there may be no natural linear structure. Indeed, there is a plethora of such situations (tree spaces in phylogenomics, some models of cognition, configuration spaces in robotics, etc.), when we recognize a CAT(0) space as an underlying space of a given problem. We refer the interested reader to [1, 9, 14, 15, 16, 17], and the references therein. Since convex sets in CAT(0) spaces are of great importance we expect that the alternating projection method in this setting will find further applications.

Relatedly, let us mention that there is a rich fixed point theory in CAT(0) spaces, mainly due to Art Kirk [24, 25, 26]. For a different approach to alternating projection method on manifolds, see [28].

1.1. **Alternating projections in Hilbert space.** Here we briefly describe the alternating projection method in Hilbert spaces. As a reference we recommend [6]. Let H be a Hilbert space and $A, B \subset H$ closed convex sets. Symbols P_A, P_B denote the metric projections (i.e. the nearest-point mappings) onto A and B respectively. Given a starting point $x_0 \in H$, define the sequence

$$x_{2n-1} = P_A\left(x_{2n-2}\right), \quad x_{2n} = P_B\left(x_{2n-1}\right), \quad n \in \mathbb{N}. \tag{1}$$

Algorithm (1) was developed by von Neumann who also proved norm convergence in the case when A and B are two closed subspaces.

Theorem 1.1 (von Neumann). *Let H be a Hilbert space and $A, B \subset H$ its closed subspaces. For any starting point $x_0 \in H$, the sequence defined by* (1) *converges in norm to a point from $A \cap B$.*

Proof. See [8, Theorem 3.1]. □

Weak convergence in the general case was established by Bregman in 1965.

Theorem 1.2 (Bregman). *Let H be a Hilbert space and $A, B \subset H$ closed convex sets with $A \cap B \neq \emptyset$. Assume $x_0 \in H$ is a starting point and $(x_n) \subset H$ the sequence generated by Algorithm* (1)*. Then (x_n) weakly converges to a point from $A \cap B$.*

Proof. See [10], or [8, Theorem 3.3]. □

A decades-old problem as to whether or not the convergence of (1) has to be in norm was answered quite recently in the negative [19].

Example 1.3 (Hundal). There exist a hyperplane $A \subset \ell_2$, a convex cone $B \subset \ell_2$ and a point $x_0 \in \ell_2$ such that the sequence generated by Algorithm (1) from the starting point x_0 converges weakly to a point in $A \cap B$ but not in norm [19].

1.2. **Paper outline.** We generalize results on the convergence of the alternating projection method in Hilbert spaces (see [3, 5, 6]) to CAT(0) spaces using the approach of [5]. Section 2 establishes our terminology, fixes notation and presents some preliminary facts. Auxiliary results, mainly on the weak convergence and Fejér monotone sequences, are contained in Section 3. The main results (various types of convergence of the alternating projection method) are gathered in Theorem 4.1 in Section 4. Section 5 contains an application of the alternating projection method to convex optimization in CAT(0) spaces.

2. Preliminaries

We begin by outlining the framework of CAT(0) spaces. For further details on the subject, the reader is referred to [11]. Let X be a CAT(0) space. When no confusion is likely, we do not distinguish between a geodesic and its geodesic segment. Having two points $x, y \in X$, we denote the geodesic segment from x to y by $[x, y]$. A set $C \subset X$ is *convex* if $x, y \in C$ implies $[x, y] \subset C$. Let A be a subset of X. Then $\overline{\mathrm{co}}\, A$ stands for its closed convex hull defined as

$$\overline{\mathrm{co}}\, A = \bigcap \{C \subset X : A \subset C,\ C \text{ convex, closed}\}.$$

We say that a geodesic $\gamma \subset X$ goes *through* a point $p \in X$ if p lies on the geodesic segment of γ. Note, this definition allows p to be an endpoint of γ.

2.1. **Angle between geodesics.** Given $x, y, z \in X$, the symbol $\alpha(y, x, z)$ denotes the (Alexandrov) angle between the geodesics $[x, y]$ and $[x, z]$. The corresponding angle in the comparison triangle is denoted $\alpha'(y, x, z)$.

2.2. **Projections.** For any metric space X and $C \subset X$, define the *distance function* by

$$d(x, C) = \inf_{c \in C} d(x, c), \quad x \in X.$$

Interchangeably we use the symbol d_C for $d(\cdot, C)$. Note that the function d_C is convex and continuous provided X is CAT(0) and C is convex and complete [11, Cor. 2.5, p.178]. The following Proposition 2.1 is of principal importance for developing the alternating projection method in CAT(0) space.

Proposition 2.1. *Let X be a CAT(0) space and $C \subset X$ be complete and convex. Then:*

(i) *For every $x \in X$, there exists a unique point $P_C(x) \in C$ such that*
$$d(x, P_C(x)) = d(x, C).$$
(ii) *If $y \in [x, P_C(x)]$, then $P_C(x) = P_C(y)$.*
(iii) *If $x \in X \setminus C$ and $y \in C$ such that $P_C(x) \neq y$, then $\alpha(x, P_C(x), y) \geq \frac{\pi}{2}$.*
(iv) *The mapping P_C is a non-expansive retraction from X onto C. Further, the mapping $H : X \times [0, 1] \to X$ associating to (x, t) the point a distance $t d(x, P_C(x))$ on the geodesic $[x, P_C(x)]$ is a continuous homotopy from the identity map of X to P_C.*

Proof. See [11, Proposition 2.4, p.176]. □

The mapping P_C is called the *(metric) projection* onto C.

2.3. **Weak convergence.** Example 1.3 shows that we cannot do without weak convergence even in Hilbert spaces. Fortunately, there is an analogous tool at our disposal for use in all CAT(0) spaces. A notion of weak convergence in CAT(0) spaces was first introduced by Jürgen Jost in [23, Definition 2.7]. Sosov later defined his ψ- and ϕ-convergences, both generalizing Hilbert space weak convergence into geodesic metric spaces [31]. Recently, Kirk and Panyanak extended Lim's Δ-convergence [29] into CAT(0) spaces [27] and finally, Espínola and Fernández-León [13] modified Sosov's ϕ-convergence to obtain an equivalent formulation of Δ-convergence in CAT(0) spaces. This is, however, exactly the original weak convergence due to Jost [23].

Let X be a complete CAT(0) space. Suppose $(x_n) \subset X$ is a bounded sequence and define its *asymptotic radius* about a given point $x \in X$ as

$$r(x_n, x) = \limsup_{n\to\infty} d(x_n, x),$$

and *asymptotic radius* as

$$r(x_n) = \inf_{x\in X} r(x_n, x).$$

Further, we say that a point $x \in X$ is the *asymptotic center* of (x_n) if

$$r(x_n, x) = r(x_n).$$

Since X is a complete CAT(0) space we know that the asymptotic center of (x_n) exists and is unique [12, Proposition 7].

We shall say that $(x_n) \subset X$ *weakly converges* to a point $x \in X$ if x is the asymptotic center of each subsequence of (x_n). We use the notation $x_n \xrightarrow{w} x$.

If there is a subsequence (x_{n_k}) of (x_n) such that $x_{n_k} \xrightarrow{w} z$ for some $z \in X$, we say that z is a *weak cluster point* of the sequence (x_n). Each bounded sequence has a weak cluster point, see [23, Theorem 2.1], or [27, p. 3690].

Proposition 2.2. *A bounded sequence $(x_n) \subset X$ weakly converges to a point $x \in X$ if and only if, for any geodesic γ through x, we have*

$$d(x, P_\gamma(x_n)) \to 0, \quad \text{as } n \to \infty.$$

Proof. See [13, Proposition 5.2]. □

Clearly, if $x_n \to x$, then $x_n \xrightarrow{w} x$.

We shall say that a function $f : X \to \mathbb{R}$ is *weakly lsc* at a given point $x \in X$ if

$$\liminf_{n\to\infty} f(x_n) \ge f(x)$$

for each sequence $x_n \xrightarrow{w} x$.

2.4. **Alternating projections in CAT(0) space.** Let X be a complete CAT(0) space and $A, B \subset X$ be convex closed sets. As in Hilbert spaces, the *alternating projection method* produces the sequence

$$x_{2n-1} = P_A(x_{2n-2}), \quad x_{2n} = P_B(x_{2n-1}), \quad n \in \mathbb{N}, \tag{2}$$

where $x_0 \in X$ is a given starting point. This sequence is sometimes referred to as the alternating sequence.

3. Auxiliaries

For each of the following Lemmas let X be a complete CAT(0) space and $C \subset X$ a closed convex set. The following lemma is an analogue of one from Banach Space folklore.

Lemma 3.1. *If $(x_n) \subset C$ and $x_n \xrightarrow{w} x \in X$, then $x \in C$.*

Proof. Assume that $x \notin C$ and denote $\gamma = [x, P_C(x)]$. We claim that $P_\gamma(x_n) = P_C(x)$ for all $n \in \mathbb{N}$. Indeed, if for some $m \in \mathbb{N}$ we had $P_\gamma(x_m) \neq P_C(x)$, then by Proposition 2.1, we would have both

$$\alpha(x_m, P_C(x), P_\gamma(x_m)) \ge \frac{\pi}{2}, \quad \alpha(x_m, P_\gamma(x_m), P_C(x)) \ge \frac{\pi}{2},$$

which is impossible.

Finally,

$$d(P_\gamma(x_n), x) = d\left(P_C(x), x\right) \nrightarrow 0, \quad n \to \infty,$$

which, by Proposition 2.2, contradicts $x_n \xrightarrow{w} x$. □

Lemma 3.2. *The distance function d_C is weakly lsc.*

Proof. By contradiction. Let $(x_n) \subset X, x \in X$ and $x_n \xrightarrow{w} x$. Suppose that

$$\liminf_{n\to\infty} d_C(x_n) < d_C(x).$$

That is, there exist a subsequence $(x_{n_k}), k_0 \in \mathbb{N}$ and $\delta > 0$ such that $d_C(x_{n_k}) < d_C(x) - \delta$ for all $k > k_0$. By continuity and convexity of the distance function, we get

$$d_C(y) \le d_C(x) - \delta$$

for all $y \in \overline{\text{co}}\{x_{n_k} : k > k_0\}$. But this, through Lemma 3.1, yields a contradiction to $x_n \xrightarrow{w} x$. □

3.1. **Regularity of sets in CAT(0) space.** We say that $A, B \subset X$ are *boundedly regular* if for any bounded set $S \subset X$ and any $\varepsilon > 0$ there exists $\delta > 0$ such that if $x \in S$ and $\max\{d(x, A), d(x, B)\} \le \delta$ then $d(x, A \cap B) < \varepsilon$.

We say that $A, B \subset X$ are *boundedly linearly regular* if for any bounded set $S \subset X$ there exists $\kappa > 0$ such that for $x \in S$ we have

$$d(x, A \cap B) \le \kappa \max\{d(x, A), d(x, B)\}.$$

We say that $A, B \subset X$ are *linearly regular* if there exists $\kappa > 0$ such that for any $x \in X$ we have

$$d(x, A \cap B) \le \kappa \max\{d(x, A), d(x, B)\}.$$

3.2. **Linear convergence.** A sequence (x_n) converges *linearly* to a point $x \in X$ if there exist $K \ge 0$ and $\beta \in [0, 1)$ such that

$$d(x, x_n) \le K\beta^n, \qquad n \in \mathbb{N}.$$

The parameter β is called a *rate* of linear convergence.

3.3. **Fejér monotone sequence.** A sequence $(x_n) \subset X$ is *Fejér monotone* with respect to C if, for any $c \in C$,

$$d(x_{n+1}, c) \le d(x_n, c), \qquad n \in \mathbb{N}.$$

Proposition 3.3. *Let $(x_n) \subset X$ be a Fejér monotone sequence with respect to C. Then:*

(i) *(x_n) is bounded,*
(ii) *$d_C(x_{n+1}) \le d_C(x_n)$ for each $n \in \mathbb{N}$.*
(iii) *(x_n) weakly converges to some $x \in C$ if and only if all weak cluster points of (x_n) belong to C.*
(iv) *(x_n) converges to some $x \in C$ if and only if $d(x_n, C) \to 0$.*
(v) *(x_n) converges linearly to some $x \in C$, provided there exists $\theta \in [0, 1)$ such that $d(x_{n+1}, C) \le \theta d(x_n, C)$ for each $n \in \mathbb{N}$.*

Proof. (i) and (ii) are easy. Let us prove the nontrivial implication of (iii). Assume that all weak cluster points of (x_n) lie in C. It suffices to show that (x_n) has a unique cluster point. By contradiction, let $c_1, c_2 \in C$, with $c_1 \neq c_2$, be weak cluster points of (x_n). That is, there are subsequences (x_{n_k}) and (x_{m_k}) such that $x_{n_k} \xrightarrow{w} c_1$ and $x_{m_k} \xrightarrow{w} c_2$. Without loss of generality, assume $r(x_{n_k}) \leq r(x_{m_k})$. For any $\varepsilon > 0$ there exists $k_0 \in \mathbb{N}$ such that $d(x_{n_k}, c_1) < r(x_{n_k}) + \varepsilon$, for all $k \geq k_0$. By Fejér monotonicity we also have $d(x_{m_k}, c_1) < r(x_{n_k}) + \varepsilon$, for all $m_k \geq n_{k_0}$. Hence, there exists $k_1 \in \mathbb{N}$ such that $d(x_{m_k}, c_1) < r(x_{m_k}) + \varepsilon$, for all $k \geq k_1$. But this contradicts the fact that c_2 is the unique asymptotic center of (x_{m_k}).

Now we prove (iv). Suppose $d(x_n, C) \to 0$. Since for all $k \in \mathbb{N}$ we have

$$d(x_{n+k}, x_n) \leq d\left(x_{n+k}, P_C(x_n)\right) + d\left(x_n, P_C(x_n)\right) \tag{3a}$$

and hence, by Fejér monotonicity,

$$d(x_{n+k}, x_n) \leq d\left(x_n, P_C(x_n)\right) + d\left(x_n, P_C(x_n)\right) \leq 2d(x_n, C), \tag{3b}$$

hence (x_n) is Cauchy and so converges to a point from C. The converse implication in (iv) is trivial. It remains to prove (v). From (3) we get

$$d(x_{n+k}, x_n) \leq 2d(x_n, C) \leq 2\theta^n d(x_0, C)$$

for all $n, k \in \mathbb{N}$. The sequence (x_n) obviously converges to some $x \in C$ and thus letting $k \to \infty$ we have

$$d(x, x_n) \leq 2d(x_n, C) \leq 2\theta^n d(x_0, C) \quad n \in \mathbb{N}.$$

In other words $x_n \to x$ linearly with rate θ, completing the proof. □

Lemma 3.4. *Let $A, B \subset X$ be convex closed sets with $A \cap B \neq \emptyset$. Then the sequence generated by Algorithm* (2) *is Fejér monotone with respect to $A \cap B$.*

Proof. Pick $c \in A \cap B$. Fix $n \in \mathbb{N}$ and without loss of generality assume that $x_n \in A$. Recall $x_{n+1} = P_B(x_n)$. If $x_{n+1} = c$ we are done. Otherwise we have $\alpha(c, x_{n+1}, x_n) \geq \frac{\pi}{2}$, by Proposition 2.1. Then we also have $\alpha'(c, x_{n+1}, x_n) \geq \frac{\pi}{2}$ and so can conclude that $d(x_{n+1}, c) \leq d(x_n, c)$. □

4. Convergence results

The following theorem contains our main results.

Theorem 4.1. *Let X be a complete CAT(0) space and $A, B \subset X$ convex closed subsets such that $A \cap B \neq \emptyset$. Let $x_0 \in X$ be a starting point and $(x_n) \subset X$ be the sequence generated by Algorithm* (2)*. Then:*

(i) *(x_n) weakly converges to a point $x \in A \cap B$.*
(ii) *If A and B are boundedly regular, then $x_n \to x$.*
(iii) *If A and B are boundedly linearly regular, then $x_n \to x$ linearly.*
(iv) *If A and B are linearly regular, then $x_n \to x$ linearly with a rate independent of the starting point.*

Proof. We start by proving the following inequality, for any $n \in \mathbb{N}$,

$$\max\left\{d^2(x_n, A), d^2(x_n, B)\right\} \leq d^2(x_n, A \cap B) - d^2(x_{n+1}, A \cap B). \tag{4}$$

Indeed, fix $n \in \mathbb{N}$ and without loss of generality assume $x_n \in A$ and $x_{n+1} \notin A \cap B$. Since, by Proposition 2.1, we have

$$\alpha'\left(x_n, x_{n+1}, P_{A\cap B}(x_n)\right) \geq \alpha\left(x_n, x_{n+1}, P_{A\cap B}(x_n)\right) \geq \frac{\pi}{2},$$

it follows that

$$\begin{aligned} d^2\left(x_n, P_{A\cap B}(x_n)\right) &\geq d^2(x_n, x_{n+1}) + d^2\left(P_{A\cap B}(x_n), x_{n+1}\right), \\ d^2(x_n, A\cap B) &\geq d^2(x_n, B) + d^2(A\cap B, x_{n+1}), \end{aligned}$$

which yields (4). Now, by Fejér monotonicity (Lemma 3.4), Proposition 3.3(ii) and (4) we get

$$\max\left\{d(x_n, A), d(x_n, B)\right\} \to 0 \quad \text{as } n \to \infty. \tag{5}$$

Let us prove (i). Using Fejér monotonicity (Lemma 3.4), we obtain that (x_n) is bounded and hence it has a weak cluster point $x \in X$. Take a subsequence (x_{n_k}) which weakly converges to x. Using Lemma 3.2 and (5), we have $d(x, A) = d(x, B) = 0$. Hence $x \in A \cap B$ and we conclude, by Proposition 3.3(iii), that $x_n \xrightarrow{w} x$.

As for (ii), bounded regularity of A and B along with (5) gives $d(x_n, A\cap B) \to 0$ as $n \to \infty$. Applying Lemma 3.4 and Proposition 3.3(iv) yields (ii).

To prove (iii), recall that (x_n) is bounded. Hence, by bounded linear regularity, there exists $\kappa > 0$ such that, for every $n \in \mathbb{N}$,

$$d(x_n, A\cap B) \leq \kappa \max\left\{d(x_n, A), d(x_n, B)\right\}.$$

Using (4), we arrive at

$$\begin{aligned} d^2(x_n, A\cap B) &\leq \kappa^2\left(d^2(x_n, A\cap B) - d^2(x_{n+1}, A\cap B)\right), \\ d(x_{n+1}, A\cap B) &\leq \sqrt{1 - \frac{1}{\kappa^2}}\, d(x_n, A\cap B) \end{aligned}$$

Applying Proposition 3.3(v) finishes the proof of (iii).

Finally, the proof of (iv) is similar to that one of (iii). □

5. Applications: Minimizing convex functions in CAT(0)

Let (X, d) be a complete CAT(0) space. Since there is a bijective correspondence between the class of closed convex subsets of X and the class of lower semicontinuous (lsc, for short) convex functions on X, we get many natural examples of closed convex sets in X. Namely, let $f : X \to (-\infty, \infty]$ be a lsc convex function, then the α-sublevel set, where $\alpha \geq \inf_X f$, defined as

$$A_f^\alpha = \{x \in X : f(x) \leq \alpha\}$$

is a closed convex subset of X.

In this final section, we would like to present an application of the alternating projection method to convex optimization in CAT(0). Let us first recall that examples of continuous convex functions on (X, d) include the following.

(i) The function
$$x \mapsto d\left(x, x_0\right),$$
where x_0 is a fixed point of X.

(ii) The square of the function in (i), which is even *strictly* convex.

More generally,

(iii) the distance function to a closed convex subset $C \subset X$, defined in Paragraph 2.2.

(iv) Displacement functions for isometries [11, Definition II.6.1]. Let $T : X \to X$ be an isometry. The *displacement function* of T is the function $d_T : X \to [0, \infty)$ defined by $d_T(x) = d(x, Tx)$. It is convex and Lipschitz.

(v) Busemann functions [11, Definition II.8.7]. Let $c : [0, \infty) \to X$ be a geodesic ray. The function $b_c : X \to \mathbb{R}$ defined by

$$b_c(x) = \lim_{t \to \infty} \left[d\left(x, c(t)\right) - t \right], \qquad x \in X$$

is called the *Busemann function* associated to the ray c. Busemann functions are convex and 1-Lipschitz. Concrete examples of Busemann functions are given in [11, p. 273]. Another explicit example of a Busemann function in the CAT(0) space of positive definite $n \times n$ matrices with real entries is found in [11, Proposition 10.69]. The sublevel sets of Busemann functions are called *horoballs* and carry a lot of information about the geometry of the space in question, see [11] and references therein.

The above mentioned convex functions are well defined on any complete CAT(0) space. A further, very different, example is provided by the energy functional on a special CAT(0) space that is very important in many areas of analysis and geometry [20, 21, 22]. An understanding of this example requires some rudimentary knowledge of differential geometry and algebraic topology; we refer the reader to [20, 18]. The reader who does not wish to go into the details may skip over the following example without compromising their understanding of the remainder of the section.

Example 5.1. We shall follow [20, Chapter 7], a more general construction is given in [22, Chapter 4]. Let M and N be compact Riemannian manifolds, with N having nonpositive sectional curvature. For $f \in L^2(M, N)$ and $h > 0$ define

$$E_h(f) = \int_M \int_{B(x,h)} \eta_h(x, y) \frac{d^2\left(f(x), f(y)\right)}{h^2} \mathrm{dVol}(y) \mathrm{dVol}(x), \tag{7}$$

where dVol is the Riemannian volume form on M, and $\eta_h(x, y)$ is a calibrating kernel. The energy of f is then defined as

$$E(f) = \lim_{h \to 0} E_h(f). \tag{8}$$

The functionals E_h are convex and continuous, whereas the energy functional E is convex and lsc. Minimizers of the energy functional are called harmonic maps. However, in many situations (like in [20, Theorem 7.5.2]) it turns out that instead of considering the energy functional on $L^2(M, N)$, it is more convenient to extend it to the CAT(0) space of equivariant maps between the universal covers. We do that now.

Let $g : M \to N$ be a continuous map. Given a point $p \in M$, the homomorphism between the fundamental groups $\pi_1(M, p)$ and $\pi_1\left(N, g(p)\right)$ induced by the map g is denoted

$$g_\sharp : \pi_1(M, p) \to \pi_1\left(N, g(p)\right).$$

We will, for simplicity, denote $g_\sharp$ by ρ. Let $\tilde{M}$ and $\tilde{N}$ be universal covers of M and N, respectively, and let

$$\tilde{g} : \tilde{M} \to \tilde{N}$$

be the lift of g. More precisely, it is the lift of $g \circ \pi$, where $\pi : \tilde{M} \to M$ is the covering map. The lift exists since the lifting condition is trivially satisfied: the universal cover $\tilde{M}$ is simply connected and hence the fundamental group $\pi_1(\tilde{M})$ is trivial. Also the map $\tilde{g}$ is ρ-equivariant, that is,

$$\tilde{g}(\lambda x) = \rho(\lambda)\tilde{g}(x),$$

for all $x \in \tilde{M}$ and $\lambda \in \pi_1(M, p)$, where the fundamental groups operate by deck transformations. For ρ-equivariant maps $h_1, h_2 : \tilde{M} \to \tilde{N}$, we define an L^2-distance by

$$d(h_1, h_2) = \left(\int d^2 (h_1(x), h_2(x)) \, \mathrm{dVol}(M) \right)^{1/2},$$

where we integrate with respect to the volume form on M and over some fundamental domain in $\tilde{M}$. Then we put

$$L^2_\rho(M, N) = \left\{ h : \tilde{M} \to \tilde{N},\ h \text{ is } \rho\text{-equivariant},\ d(h, \tilde{g}) < \infty \right\}.$$

Since $\tilde{N}$ is a complete simply connected Riemannian manifold of nonpositive sectional curvature, the space $L^2_\rho(M, N)$ is a complete CAT(0) space.

We can now consider the energy functionals E and E_h, defined in (7) and (8), as functionals on the space $L^2_\rho(M, N)$. Then E_h is convex and continuous on $L^2_\rho(M, N)$, and the energy functional E is convex and lsc.

Moreover, the space $L^2_\rho(M, N)$ is very different from all the examples of CAT(0) spaces mentioned in the Introduction. In particular, it is different from Hilbert spaces since it is not flat, and it is different from Riemannian manifolds since it is not locally compact. □

Let X be a complete CAT(0) space. Consider now the following optimization problem. We are given a function $F : X \to (-\infty, \infty]$ of the form $F = \max(f, g)$, where $f, g : X \to (-\infty, \infty]$ are lsc and convex, and we wish to find a minimizer of F, that is some $x \in A^\alpha_F = \{x \in X : F(x) \le \alpha\}$, where $\alpha = \inf_X F$, of course we are assuming that $\inf_X F$ is finite and the set A^α_F is nonempty. Or, alternatively we may seek an approximative minimizer for F; that is, given some $\alpha > \inf_X F$ we want to find some $x \in A^\alpha_F$. Then, in case the projections onto A^α_f and A^α_g are easy to compute, we can find the desired $x \in A^\alpha_F$ as the limit of the alternating sequence since

$$A^\alpha_F = A^\alpha_f \cap A^\alpha_g.$$

In general when the functions f and g are only lsc and convex, we have weak convergence of the alternating sequence by Theorem 4.1(i), and this is the best we can hope for. If, however, we impose additional assumptions on the functions f and g, we get strong convergence, as we shall see in Proposition 5.2 below.

We will first recall that a function $h : X \to (-\infty, \infty]$ is *uniformly convex* if there exists $\lambda > 0$ such that for any $x, y \in X$ and $u \in [x, y]$ we have

$$h(u) \le (1 - t)h(x) + th(y) - \lambda t(1 - t)d^2(x, y),$$

where $t = \frac{d(x,u)}{d(x,y)}$. We remark that uniform convexity is also essential in Mayer's approach to energy minimization [30].

The following proposition provides the promised sufficient conditions on the functions f and g to ensure the sets A^α_f and A^α_g are 'more regular', and hence allows us to obtain, via Theorem 4.1, strong convergence for the alternating sequence to an (approximative) minimizer of the functional $F = \max(f, g)$.

Proposition 5.2. *Let X be a complete CAT(0) space, and $F : X \to (-\infty, \infty]$ be a functional of the form $F = \max(f, g)$, where $f, g : X \to (-\infty, \infty]$ are lsc convex. Let $\alpha \geq \inf_X F > -\infty$, and A_F^α be nonempty. If the function f is both uniformly convex and uniformly continuous on bounded sets of X, then the sets A_f^α and A_g^α are boundedly regular.*

Proof. Assume $S \subset X$ is a given bounded set and $\varepsilon > 0$. We will look for $\delta > 0$ such that if one picks $x \in S$ with

$$\max\left[d\left(x, A_f^\alpha\right), d\left(x, A_g^\alpha\right)\right] < \delta, \tag{9}$$

then

$$d\left(x, A_F^\alpha\right) < \varepsilon. \tag{10}$$

Let $b = P_{A_g^\alpha}(x)$, the projection of x onto the set A_g^α. If $b \in A_f^\alpha$, we can take $\delta = \varepsilon$ in (9) to fulfill (10). If $b \notin A_f^\alpha$, then denoting the projections $P_{A_F^\alpha}(b)$ and $P_{A_f^\alpha}(b)$ by c and a respectively and taking m to be the midpoint of the geodesic $[b, c]$ we have, by the uniform convexity of f, that there exists $\lambda > 0$ such that

$$f(m) \leq \frac{1}{2}\left[f(b) + f(c)\right] - \lambda d^2(b, c),$$

and hence,

$$d^2(b, c) \leq \frac{1}{\lambda}\left[\frac{f(b) + f(c)}{2} - f(m)\right]. \tag{11}$$

By uniform continuity of f, there exists $\delta' > 0$ such that

$$|f(b) - f(x')| < \frac{\varepsilon^2 \lambda}{2},$$

whenever $d(b, x') < \delta'$. Therefore, if $d(b, a) < \delta'$, from (11) we further have

$$d^2(b, c) < \frac{1}{\lambda}\left[\frac{f(a) + f(c)}{2} + \frac{\varepsilon^2 \lambda}{4} - f(m)\right] \leq \frac{\varepsilon^2}{4},$$

which yields

$$d(b, c) = d\left(b, A_F^\alpha\right) < \frac{\varepsilon}{2},$$

and hence, if we choose $\delta < \frac{1}{2}\max(\varepsilon, \delta')$ in (9) we obtain

$$d\left(x, A_f^\alpha \cap A_g^\alpha\right) \leq d(x, b) + d(b, c) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

That is, the sets A_f^α and A_g^α are boundedly regular. □

Notice that in the above Proposition 5.2 we only make additional assumptions on the function f, whereas the function g is arbitrary lsc convex.

References

[1] F. Aurenhammer: *Voronoi diagrams - A survey of a fundamental geometric data structure.* ACM Computing Surveys, Vol. 23, No. 3 (1991), 345–405.

[2] H.H. Bauschke: *A norm convergence result on random products of relaxed projections in Hilbert space.* Trans. Amer. Math. Soc. 347 (1995), no. 4, 1365–1373.

[3] H.H. Bauschke: *Projection algorithms: results and open problems.* Inherently parallel algorithms in feasibility and optimization and their applications (Haifa, 2000), 11–22, Stud. Comput. Math., 8, North-Holland, Amsterdam, 2001.

[4] H.H. Bauschke, J.V. Burke, F.R. Deutsch, H.S. Hundal, and J.D. Vanderwerff: *A new proximal point iteration that converges weakly but not in norm.* Proc. Amer. Math. Soc. 133 (2005), no. 6, 1829–1835.

[5] H.H. Bauschke, J.M. Borwein: *On the convergence of von Neumann's alternating projection algorithm for two set.* Set-Valued Analysis 1 (1993), 185–212.

[6] H.H. Bauschke, J.M. Borwein: *On projection algorithms for solving convex feasibility problems.* SIAM Rev. 38 (1996), 367–426.

[7] H.H. Bauschke, F.R. Deutsch, and H.S. Hundal: *Characterizing arbitrarily slow convergence in the method of alternating projections.* Int. Trans. Oper. Res. 16 (2009), 413–425.

[8] H.H. Bauschke, E. Matoušková, and S. Reich: *Projection and proximal point methods: convergence results and counterexamples.* Nonlinear Anal. 56 (2004), no. 5, 715–738.

[9] L.J. Billera, S.P. Holmes, and K. Vogtmann: *Geometry of the space of phylogenetic trees.* Adv. in Appl. Math. 27 (2001), no. 4, 733–767.

[10] L.M. Bregman: *Finding the common point of convex sets by the method of successive projection.* (Russian) Dokl. Akad. Nauk SSSR 162 (1965), 487–490. (English) Soviet Math. Dokl. 6 (1965), 688–692.

[11] M.R. Bridson, A. Haefliger: *Metric spaces of non-positive curvature.* Springer-Verlag, Berlin, 1999.

[12] S. Dhompongsa, W.A. Kirk, and B. Sims: *Fixed points of uniformly Lipschitzian mappings.* Nonlinear Anal. 65 (2006), no. 4, 762–772.

[13] R. Espínola, A. Fernández-León: *CAT(k)-spaces, weak convergence and fixed points.* J. Math. Anal. Appl. 353 (2009), no. 1, 410–427.

[14] P. Gärdenfors, M.-A. Williams: *Reasoning about categories in conceptual spaces.* Proceedings of the Fourteenth International Joint Conference on Arfitial Inteligence, Morgan Kaufmann, 385–392, 2001.

[15] R. Ghrist: *Configuration spaces, braids, and robotics.* Braids, 263–304, Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., 19, World Sci. Publ., Hackensack, NJ, 2010.

[16] R. Ghrist, S.M. Lavalle: *Nonpositive curvature and Pareto optimal coordination of robots.* SIAM J. Control Optim. 45 (2006), no. 5, 1697–1713.

[17] R. Ghrist, V. Peterson: *The geometry and topology of reconfiguration.* Adv. in Appl. Math. 38 (2007), no. 3, 302–323.

[18] A. Hatcher: *Algebraic topology.* Cambridge University Press, Cambridge, 2002.

[19] H.S. Hundal: *An alternating projection that does not converge in norm.* Nonlinear Anal. 57 (2004), no. 1, 35–61.

[20] J. Jost: *Riemannian geometry and geometric analysis.* Fifth edition. Universitext. Springer-Verlag, Berlin, 2008.

[21] J. Jost: *Nonlinear Dirichlet forms.* New directions in Dirichlet forms, 1–47, AMS/IP Stud. Adv. Math., 8, Amer. Math. Soc., Providence, RI, 1998.

[22] J. Jost: *Nonpositive curvature: geometric and analytic aspects.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1997.

[23] J. Jost: *Equilibrium maps between metric spaces.* Calc. Var. Partial Differential Equations 2 (1994), no. 2, 173–204.

[24] W.A. Kirk: *Geodesic geometry and fixed point theory.* Seminar of Mathematical Analysis (Malaga/Seville, 2002/2003), 195–225, Colecc. Abierta, 64, Univ. Sevilla Secr. Publ., Seville, 2003.

[25] W.A. Kirk: *Fixed point theorems in CAT(0) spaces and R-trees.* Fixed Point Theory Appl. 2004, no. 4, 309–316.

[26] W.A. Kirk: *Geodesic geometry and fixed point theory. II.* International Conference on Fixed Point Theory and Applications, 113–142, Yokohama Publ., Yokohama, 2004.

[27] W.A. Kirk, B. Panyanak: *A concept of convergence in geodesic spaces.* Nonlinear Anal. 68 (2008), no. 12, 3689–3696.
[28] A.S. Lewis, J. Malick: *Alternating projections on manifolds.* Math. Oper. Res. 33 (2008), no. 1, 216–234.
[29] T.C. Lim: *Remarks on some fixed point theorems.* Proc. Amer. Math. Soc. 60 (1976), 179–182.
[30] U. Mayer: *Gradient flows on nonpositively curved metric spaces and harmonic maps.* Comm. Anal. Geom. 6 (1998), no. 2, 199–253.
[31] E.N. Sosov: *On analogues of weak convergence in a special metric space.* (Russian) Izv. Vyssh. Uchebn. Zaved. Mat. 2004, no. 5, 84–89; translation in Russian Math. (Iz. VUZ) 48 (2004), no. 5, 79–83.

Miroslav Bačák, Max Planck Institute, Inselstrasse 22, 04103 Leipzig, Germany
E-mail address: bacak@mis.mpg.de

Ian Searston, School of Mathematical and Physical Sciences, University of Newcastle, Callaghan, NSW 2308, Australia
E-mail address: Ian.Searston@newcastle.edu.au

Brailey Sims, School of Mathematical and Physical Sciences, University of Newcastle, Callaghan, NSW 2308, Australia
E-mail address: Brailey.Sims@newcastle.edu.au