Proximal minimization in CAT(\(\kappa\)) spaces

Rafa Espínola\(^1\), Adriana Nicolae\(^{1,2}\)

\(^1\) Department of Mathematical Analysis, University of Seville
Apdo. 1160, 41080 Sevilla, Spain
\(^2\) Department of Mathematics, Babeş-Bolyai University,
Kogălniceanu 1, 400084 Cluj-Napoca, Romania
E-mails: espinola@us.es, anicolae@math.ubbcluj.ro

Abstract

In this note, we provide convergence results for the proximal point algorithm and a splitting variant thereof in the setting of CAT(\(\kappa\)) spaces with \(\kappa > 0\) using a recent definition for the resolvent of a convex, lower semi-continuous function due to Kimura and Kohsaka (J. Fixed Point Theory Appl. 18 (2016), 93–115).

Keywords: Proximal minimization, convex optimization, CAT(\(\kappa\)) space.

MSC 2010: 90C25, 49M37, 53C23.

1 Introduction

In Hilbert spaces, the proximal point algorithm originates from Martinet \([16]\) and Rockafellar \([19]\) and is a well-known method used for minimizing convex, lower-semicontinuous functions. More recently, this algorithm and generalizations thereof have been introduced in nonlinear settings too such as Riemannian manifolds of nonpositive sectional curvature \([8, 13]\), of sectional curvature bounded above by \(\kappa > 0\) \([14]\) or even in geodesic spaces of nonpositive curvature in the sense of Alexandrov (also known as CAT(0) spaces) \([2]\).

Let \(X\) be a complete CAT(0) space and \(f : X \to (-\infty, \infty]\) be a proper, convex and lower semi-continuous function. The proximal point algorithm generates a sequence \((x_n)\) starting from a point \(x_0 \in X\) by the following rule: \(x_{n+1} = J^{f}_{\lambda_n}(x_n)\), where \((\lambda_n)\) is a sequence of positive real numbers and for \(\lambda > 0\),

\[
J^{f}_{\lambda}(x) = \arg\min_{y \in X} \left[ f(y) + \frac{1}{\lambda} d(y, x)^2 \right], \quad x \in X. \tag{1}
\]

The mapping \(J^{f}_{\lambda} : X \to X\), called the resolvent of \(f\), is well-defined in this context and was studied by Jost \([9, 10]\) and Mayer \([17]\) in connection to the theory of generalized harmonic maps. If there is no ambiguity concerning the function \(f\), we usually just write \(J^{f}_{\lambda_n}\). One of the remarkable properties of the resolvent is the fact that it is firmly nonexpansive in the sense of \([11]\). Moreover,

\*After the online publication of this paper, we realized that the proximal point algorithm and its splitting version discussed here had been previously obtained in \([18]\) in the setting of CAT(\(\kappa\)) spaces using in the definition of the resolvent the squared distance function, see \([11]\).
the set of fixed points of $J_\lambda$ is precisely the set of minimum points of $f$. Considering a suitable notion of weak convergence that goes back to Lim \cite{lim1984fixed} and is also referred to as $\Delta$-convergence, Ariza-Ruiz, Leustean and López-Acedo \cite{ariza2008asymptotic} showed that for any $\lambda > 0$ and $x \in X$, the sequence of Picard iterates $(J_\lambda^k x)$ $\Delta$-converges to a minimum point of $f$ (provided such a point exists). Bačák \cite{bacak2014convex} proved that if $f$ attains its minimum, then the sequence $(x_n)$ generated by the proximal point algorithm $\Delta$-converges to a minimum point of $f$. Motivated by results of Bertsekas \cite{bertsekas1979proximal} in the Euclidean setting, Bačák \cite{bacak2014convex} additionally studied in the context of CAT(0) spaces a splitting proximal point algorithm for finding a minimum point of a function that can be written as a finite sum of convex, lower semi-continuous functions and applied his findings to the computation of the geometric median and the Fréchet mean of a finite set of points.

Very recently, Kimura and Kohsaka \cite{kimura2016asymptotic} introduced in CAT($\kappa$) spaces with $\kappa > 0$ the resolvent of a convex, lower-semicontinuous function $f$ as an instance of so-called firmly spherically nonspreading mappings. They showed that, under appropriate boundedness conditions, the Picard iterates of the resolvent of $f$ $\Delta$-converges to a minimum point of $f$. In this paper we use this definition to study in the setting of CAT($\kappa$) spaces with $\kappa > 0$ the convergence of the corresponding versions of the proximal point and the splitting proximal point algorithms discussed in \cite{bacak2014convex, bacak2014convex} in CAT(0) spaces.

Finally, we would like to point out that after the submission of this paper, it was brought to our attention that Kimura and Kohsaka have also independently submitted a recent joint work (which was accepted in the meantime, see \cite{kimura2017asymptotic}) on the proximal point algorithm in CAT($\kappa$) spaces.

## 2 Preliminaries

Let $(X, d)$ be a metric space. A geodesic path joining $x, y \in X$ is a mapping $c : [0, l] \subseteq \mathbb{R} \to X$ such that $c(0) = x, c(l) = y$ and $d(c(t), c(t')) = |t - t'|$ for every $t, t' \in [0, l]$. The image $c([0, l])$ of $c$ is called a geodesic segment from $x$ to $y$. A point $z \in X$ belongs to such a geodesic segment if there exists $t \in [0, 1]$ such that $d(z, x) = td(x, y)$ and $d(z, y) = (1 - t)d(x, y)$, and in this case we write $z = (1 - t)x + ty$. We say that $(X, d)$ is a geodesic space if every two points in $X$ can be joined by a geodesic path. A subset $C$ of $X$ is convex if given two points of $C$, any geodesic segment joining them is contained in $C$. A rigorous introduction to geodesic spaces is provided in \cite{bakonyi1984geodesic}.

Let $(X, d)$ be a geodesic space. Having $C \subseteq X$ convex and $f : C \to (-\infty, \infty]$, the domain of $f$ is defined by $\text{dom } f = \{x \in C \mid f(x) < \infty\}$. The function $f$ is called proper if $\text{dom } f \neq \emptyset$. We say that $f$ is convex if for every $x, y \in C$ and $t \in [0, 1]$, $f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y)$. The function $f$ is uniformly convex on $\text{dom } f$ if there exists a nondecreasing function $\delta : [0, \infty) \to [0, \infty]$ vanishing only at 0 such that for every $x, y \in \text{dom } f$ and $t \in [0, 1],$

$$f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y) - t(1 - t)\delta(d(x, y)).$$

One can prove that this is in fact equivalent to the following condition (see also \cite{rockafellar1970convex} where uniformly convex functions are studied in Banach spaces): for every $\varepsilon > 0$ there exists $\delta > 0$ such that if $x, y \in \text{dom } f$ with $d(x, y) \geq \varepsilon$, then

$$f\left(\frac{1}{2}x + \frac{1}{2}y\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y) - \delta.$$

Fix $\kappa \in \mathbb{R}$ and denote by $M^2_\kappa$ the complete, simply connected model surface of constant sectional curvature $\kappa$. A comparison triangle for a geodesic triangle $\Delta(x_1, x_2, x_3)$ in $X$ is a triangle $\Delta = \Delta(\overline{x}_1, \overline{x}_2, \overline{x}_3)$ in $M^2_\kappa$ such that $d(x_i, x_j) = d_{M^2_\kappa}(\overline{x}_i, \overline{x}_j)$ for $i, j \in \{1, 2, 3\}$. A geodesic triangle $\Delta$ is said to satisfy the CAT($\kappa$) inequality if for every comparison triangle $\Delta$ of $\Delta$ and every $x, y \in \Delta$
we have that \(d(x, y) \leq d_{M^2}(x, y)\), where \(\overline{\tau}, \overline{\gamma} \in \Delta\) are the comparison points of \(x\) and \(y\), i.e., if 
\[
x = (1 - t)x_i + tx_j \text{ then } \overline{\tau} = (1 - t)\overline{\tau}_i + t\overline{\tau}_j.
\]
A metric space is called a CAT(\(\kappa\)) space if every two points (at distance less than \(\pi/\sqrt{\kappa}\) for \(\kappa > 0\)) can be joined by a geodesic path and every geodesic triangle (having perimeter less than \(2\pi/\sqrt{\kappa}\) for \(\kappa > 0\)) satisfies the CAT(\(\kappa\)) inequality.

Suppose in the sequel that \((X, d)\) is a complete CAT(\(\kappa\)) space with \(\kappa > 0\) such that for every 
\(v, w \in X, d(v, w) < \pi/(2\sqrt{\kappa})\). For \(x, y, z \in X\) with \(y \neq z\) and \(t \in [0, 1]\), the following inequality
\[
\cos \left(\sqrt{\kappa}d((1 - t)y + tz, x)\right) \geq \frac{\sin \left(\sqrt{\kappa}(1 - t)d(y, z)\right)}{\sin \left(\sqrt{\kappa}d(y, z)\right)} \cos \left(\sqrt{\kappa}d(y, x)\right) + \frac{\sin \left(\sqrt{\kappa}d(y, z)\right)}{\sin \left(\sqrt{\kappa}d(y, z)\right)} \cos \left(\sqrt{\kappa}d(z, x)\right)
\]
is an immediate consequence of the spherical law of cosines.

Let \((x_n)\) be a sequence in \(X\). For \(x \in X\), set \(r(x, (x_n)) = \limsup_{n \to \infty} d(x, x_n)\). The asymptotic radius of \((x_n)\) is given by \(r((x_n)) = \inf \{r(x, x_n) \mid x \in X\}\) and the asymptotic center of \((x_n)\) is the set \(A((x_n)) = \left\{x \in X : \limsup_{n \to \infty} d(x, x_n) = r((x_n))\right\}\). The sequence \((x_n)\) is said to \(\Delta\)-converge to \(x \in X\) if \(x\) is the unique point in the asymptotic center of every subsequence of \((x_n)\). In this case \(x\) is called the \(\Delta\)-limit of \((x_n)\). Assume next that \(r((x_n)) < \pi/(2\sqrt{\kappa})\). Then \(A((x_n))\) is a singleton (see [7] Proposition 4.1) and \((x_n)\) has a \(\Delta\)-convergent subsequence (see [7] Corollary 4.4). Moreover, if \((x_n)\) \(\Delta\)-converges to some \(x \in X\) and \(f : X \to (-\infty, \infty]\) is a proper, convex lower semi-continuous function, then 
\(f(x) \leq \liminf_{n \to \infty} f(x_n)\) (see [11] Lemma 3.1).

Any proper, convex and lower semi-continuous function \(f : X \to (-\infty, \infty]\) is bounded below (see [11] Theorem 3.6). If there exists a sequence \((x_n)\) such that \(\lim_{n \to \infty} f(x_n) = \inf_{y \in X} f(y)\) and \(r((x_n)) < \pi/(2\sqrt{\kappa})\), then \(f\) attains its minimum, i.e., there exists \(z \in X\) such that \(f(z) = \inf_{x \in X} f(x)\) and we call \(z\) a minimum point of \(f\). Indeed, denote \(r = \inf_{y \in X} f(y)\) and consider the sets \(C_p = \{y \in X \mid f(y) \leq r + 1/p\}\). For any \(p \geq 1\), the sequence \((x_n)\) will eventually be contained in \(C_p\). In addition, one can easily see that \((C_p)\) is a decreasing sequence of nonempty, closed and convex sets, so, by [7] Corollary 3.6, \(\bigcap_{p \geq 1} C_p \neq \emptyset\). Any point in this intersection is a minimum point of \(f\). In particular, if \(\text{diam}(X) < \pi/(2\sqrt{\kappa})\), then \(f\) always has a minimum point. Note also that, if nonempty, the set of minimum points of \(f\) has a unique closest point to any given point in \(X\) because it is closed and convex. A detailed discussion on convex analysis in CAT(0) spaces can be found in [4].

In [11], Kimura and Kohsaka define and study properties of the resolvent for a proper, convex and lower semi-continuous function \(f : X \to (-\infty, \infty]\). Consider first, for a fixed \(x \in X\), the following convex functions (see [11] Lemma 4.1)
\[
\Psi^1_x : X \to [1/\kappa, \infty], \quad \Psi^1_x(y) = \frac{1}{\kappa \cos \left(\sqrt{\kappa}d(y, x)\right)},
\]
\[
\Psi^2_x : X \to [-1/\kappa, 0], \quad \Psi^2_x(y) = -\frac{\cos \left(\sqrt{\kappa}d(y, x)\right)}{\kappa},
\]
and
\[
\Psi_x : X \to [0, \infty], \quad \Psi_x(y) = \Psi^1_x(y) + \Psi^2_x(y).
\]
Then, for \( \lambda > 0 \), the resolvent of \( f \) is defined by

\[
J^\lambda_f(x) = \argmin_{y \in X} \left[ f(y) + \frac{1}{\lambda} \Psi_x(y) \right], \quad x \in X. \tag{6}
\]

This mapping is well-defined (see [11, Theorem 4.2]) and, when \( k \searrow 0 \), one actually recovers the definition (1) of the resolvent in CAT(0) spaces. Moreover, if \( C \subseteq X \) is nonempty, closed and convex, then the indicator function \( \delta_C : X \rightarrow [0, \infty] \),

\[
\delta_C(x) := \begin{cases} 0, & \text{if } x \in C, \\ \infty, & \text{otherwise}, \end{cases}
\]

is proper, convex and lower semi-continuous and for any \( \lambda > 0 \), \( J^\lambda_{\delta_C} \) is the metric projection onto \( C \), as is the case for the resolvent of \( \delta_C \) in any CAT(0) space (see also [11, Remark 4.4]).

In [11, Theorem 4.6] it is shown that the resolvent \( J^\lambda \) is firmly spherically nonspreading, that is, for any \( x, z \in X \), the following inequality holds

\[
\cos (\sqrt{\kappa} d(x, J^\lambda x)) + \cos (\sqrt{\kappa} d(z, J^\lambda z)) \cos^2 (\sqrt{\kappa} d(J^\lambda x, J^\lambda z)) \\
\geq 2 \cos (\sqrt{\kappa} d(J^\lambda x, z)) \cos (\sqrt{\kappa} d(x, J^\lambda z)).
\]

Furthermore, if \( \text{Fix}(J^\lambda) \neq \emptyset \), then, by [11, Theorem 4.6.(i)], it follows that for every \( x \in X \) and \( z \in \text{Fix}(J^\lambda) \),

\[
\cos (\sqrt{\kappa} d(J^\lambda x, z)) \cos (\sqrt{\kappa} d(x, J^\lambda x)) \geq \cos (\sqrt{\kappa} d(x, z)),
\]

a condition which is satisfied by the projection mapping onto closed and convex subsets (see [7, Proposition 3.5]).

The following result will be used in the next section.

**Lemma 2.1.** Let \( (a_j) \) and \( (b_j) \) be sequences of nonnegative real numbers such that \( (a_j) \) is bounded, \( \sum_{j=0}^{\infty} b_j < \infty \) and there exists \( j_0 \in \mathbb{N} \) such that for all \( j \geq j_0 \), \( a_{j+1} \geq a_j - b_j \). Then \( (a_j) \) is convergent.

**Proof.** For all \( m, n \in \mathbb{N} \) with \( n \geq m \geq j_0 \), \( a_{n+1} \geq a_m - \sum_{j=m}^{n} b_j \). Thus, \( \liminf_{n \rightarrow \infty} a_n \geq a_m - \sum_{j=m}^{\infty} b_j \), from where \( \liminf_{n \rightarrow \infty} a_n \geq \limsup_{m \rightarrow \infty} a_m \), which shows that \( (a_j) \) is convergent. \( \square \)

### 3 Main results

Let \( (X, d) \) be a complete CAT(\( \kappa \)) space with \( \kappa > 0 \) such that for every \( v, w \in X \), \( d(v, w) < \pi/(2\sqrt{\kappa}) \) and suppose \( f : X \rightarrow (-\infty, \infty] \) is a proper, convex and lower semi-continuous function. The following inequality also appears in the proof of [11, Theorem 4.6] in a more particular form.

**Lemma 3.1.** If \( \lambda > 0 \) and \( J^\lambda \) is defined by (6), then for \( x, z \in X \) we have that

\[
\lambda (f(J^\lambda x) - f(z)) \leq 2 \left( 1 + \frac{1}{\cos^2 (\sqrt{\kappa} d(x, J^\lambda x))} \right) \times \left( \cos (\sqrt{\kappa} d(z, J^\lambda x)) \cos (\sqrt{\kappa} d(x, J^\lambda x)) - \cos (\sqrt{\kappa} d(z, x)) \right).
\]
Proof. If \( z = J_\lambda x \), the inequality holds with equality. Otherwise, let \( a = \sqrt{d}(x, z) \), \( b = \sqrt{d}(x, J_\lambda x) \), \( c = \sqrt{d}(z, J_\lambda x) \) and \( e = \sqrt{d}((1 - t)z + tJ_\lambda x, x) \), where \( t \in (0, 1) \). Since

\[
f(J_\lambda x) + \frac{1}{\lambda} \left( \frac{1}{\cos b} - \cos b \right) \leq f((1 - t)z + tJ_\lambda x) + \frac{1}{\lambda} \left( \frac{1}{\cos e} - \cos e \right)
\leq (1 - t)f(z) + tf(J_\lambda x) + \frac{1}{\lambda} \left( \frac{1}{\cos e} - \cos e \right)
\]

and, by (2),

\[
\cos e \geq \frac{\sin((1 - t)c)}{\sin c} \cos a + \frac{\sin(tc)}{\sin c} \cos b
\]

we obtain that

\[
\lambda(1 - t)(f(J_\lambda x) - f(z)) \leq \frac{\cos b (\sin c - \sin(tc)) - \sin((1 - t)c) \cos a}{\cos b (\sin((1 - t)c) \cos a + \sin(tc) \cos b)}
+ \left( 1 - \frac{\sin(tc)}{\sin c} \right) \cos b - \frac{\sin((1 - t)c)}{\sin c} \cos a.
\]

Dividing by \((1 - t)\) and letting \( t \searrow 1 \),

\[
\lambda(f(J_\lambda x) - f(z)) \leq \frac{c}{\sin c} \left( 1 + \frac{1}{\cos^2 b} \right) (\cos c \cos b - \cos a).
\]

Using the fact that for any \( \alpha \in [0, \pi/2) \), \( \sin \alpha \geq \alpha/2 \), we obtain the desired inequality. \( \square \)

Consider the following variant of the proximal point algorithm where one uses the resolvent defined by (6): given \((\lambda_n)\) a sequence of positive real numbers and \(x_0 \in X\), define the sequence \((x_n)\) in \(X\) by

\[
x_{n+1} = J_{\lambda_n}(x_n) = \arg\min_{y \in X} \left[ f(y) + \frac{1}{\lambda_n} \Psi_{x_n}(y) \right]. \tag{7}
\]

The next result shows that the sequence \((x_n)\) defined above \(\Delta\)-converges to a minimum point of \(f\) (provided such a point exists) and constitutes a counterpart of [2, Theorem 1.4] from the context of CAT(0) spaces.

**Theorem 3.2.** Let \((X, d)\) be a complete CAT(\(\kappa\)) space with \(\kappa > 0\) such that for every \(v, w \in X\), \(d(v, w) < \pi/(2\sqrt{\kappa})\). Suppose \(f : X \to (-\infty, \infty]\) is a proper, convex and lower semi-continuous function which attains its minimum. Then, given any \(x_0 \in X\) and any sequence of positive real numbers \((\lambda_n)\) with \(\sum_{n \geq 0} \lambda_n = \infty\), the sequence \((x_n)\) defined by (7) \(\Delta\)-converges to a minimum point of \(f\).

**Proof.** Let \(z\) be a minimum point of \(f\). By Lemma 3.1, we have that for every \(n \in \mathbb{N}\),

\[
\cos \left( \sqrt{d}(z, x_n) \right) \leq \cos \left( \sqrt{d}(z, x_{n+1}) \right) \leq \cos \left( \sqrt{d}(x_n, x_{n+1}) \right) \leq \cos \left( \sqrt{d}(z, x_{n+1}) \right).
\]

This yields \(d(z, x_{n+1}) \leq d(z, x_n)\), so \((x_n)\) is Fejér monotone with respect to the set of minimum points of \(f\). Moreover, \(\lim_{n \to \infty} d(z, x_n) \leq d(z, x_0) < \pi/(2\sqrt{\kappa})\) and \(\lim_{n \to \infty} d(x_n, x_{n+1}) = 0\), hence there exists \(n_0 \in \mathbb{N}\) such that for each \(n \geq n_0\), \(1/\cos^2 \left( \sqrt{d}(x_n, x_{n+1}) \right) < 2\). Note also that

\[
f(x_{n+1}) + \frac{1}{\lambda_n} \Psi_{x_n}(x_{n+1}) \leq f(x_n),
\]
which shows that \((f(x_n))\) is nonincreasing. In addition, again by Lemma 3.1 we get that for all \(n \geq n_0\),
\[
\lambda_n (f(x_{n+1}) - f(z)) \leq 6 \left( \cos \left( \sqrt{\kappa}d(z, x_{n+1}) \right) - \cos \left( \sqrt{\kappa}d(z, x_n) \right) \right).
\]
Thus, for \(m \geq n_0\),
\[
(f(x_{m+1}) - f(z)) \sum_{n=n_0}^{m} \lambda_n \leq \sum_{n=n_0}^{m} \lambda_n (f(x_{n+1}) - f(z)) \leq 6 \left( \cos \left( \sqrt{\kappa}d(z, x_{m+1}) \right) - \cos \left( \sqrt{\kappa}d(z, x_{n_0}) \right) \right) \leq 6,
\]
from where
\[
f(x_{m+1}) \leq f(z) + \frac{6}{\sum_{n=n_0}^{m} \lambda_n}
\]
and so \(\lim_{m \to \infty} f(x_m) = f(z)\).

Let \((x_{n_i})\) be a subsequence of \((x_n)\) which \(\Delta\)-converges to some \(x \in X\). Then \(f(x) \leq \liminf_{i \to \infty} f(x_{n_i}) = f(z)\), so \(x\) is a minimum point of \(f\). Since \((x_{n_i})\) is Fejér monotone with respect to the set of minimum points of \(f\) and the \(\Delta\)-limit of every \(\Delta\)-convergent subsequence of \((x_n)\) is a minimum point of \(f\), one can easily see that \((x_n)\) \(\Delta\)-converges to \(x\) (see, for instance, [4, Proposition 3.2.6]).

**Remark 3.3.** If we assume in the previous result that \(\text{diam}(X) < \pi/(2\sqrt{\kappa})\), then \(f\) always attains its minimum. Furthermore, if \(X\) is compact, then \((x_n)\) converges to a minimum point of \(f\).

A related method for approximating a minimum point of a convex lower semi-continuous function \(f\) was given in geodesic spaces by Jost [10, Chapter 3] by considering a regularization of \(f\) with a nonnegative, lower semi-continuous function satisfying a quantitative strict convexity condition, which is fulfilled by any uniformly convex function. We show next that the function \(\Psi_x\) defined by (5) is indeed uniformly convex.

**Remark 3.4.** The function \(\Psi_x^1\) defined by (3) is uniformly convex.

**Proof.** For \(\varepsilon > 0\), take \(\delta = \varepsilon^2/32\). Let \(y, z \in X\) with \(d(y, z) \geq \varepsilon\), \(t \in [0, 1]\) and denote \(a = \sqrt{\kappa}d(x, y)\), \(b = \sqrt{\kappa}d(x, z)\) and \(c = \sqrt{\kappa}d(y, z) \geq \sqrt{\kappa}\varepsilon\). Then
\[
\Psi_x^1 \left( \frac{1}{2}y + \frac{1}{2}z \right) \leq \frac{\sin c}{\kappa \sin(c/2)(\cos a + \cos b)} \leq 2 \cos(c/2) \frac{\cos a + \cos b}{\kappa(\cos a + \cos b)^2}
\]
\[
\leq 2 \cos(c/2) \frac{\cos a + \cos b}{4\kappa \cos a \cos b} = \frac{1}{2} \cos(c/2) \left( \Psi_x^1(y) + \Psi_x^1(z) \right).
\]
Because \(1 - \cos(c/2) = 2\sin^2(c/4) \geq 2(c/8)^2 = c^2/32\), we have that
\[
\Psi_x^1 \left( \frac{1}{2}y + \frac{1}{2}z \right) \leq \frac{1}{2} \left( 1 - \frac{c^2}{32} \right) \left( \Psi_x^1(y) + \Psi_x^1(z) \right)
\]
\[
= \frac{1}{2} \Psi_x^1(y) + \frac{1}{2} \Psi_x^1(z) - \frac{c^2}{64} \left( \Psi_x^1(y) + \Psi_x^1(z) \right).
\]
At the same time, \(\Psi_x^1(y) + \Psi_x^1(z) \geq 2/\kappa\). Therefore,
\[
\Psi_x^1 \left( \frac{1}{2}y + \frac{1}{2}z \right) \leq \frac{1}{2} \Psi_x^1(y) + \frac{1}{2} \Psi_x^1(z) - \frac{c^2}{32\kappa} \leq \frac{1}{2} \Psi_x^1(y) + \frac{1}{2} \Psi_x^1(z) - \delta.
\]
\[\blacksquare\]
Thus, $\Psi_x$ is uniformly convex as the sum of a convex and a uniformly convex function and we obtain the following immediate consequence of \cite{10} Theorem 3.1.1.

**Theorem 3.5.** Let $(X,d)$ be a complete CAT($\kappa$) space with $\kappa > 0$ such that for every $v, w \in X$, $d(v, w) < \pi/(2\sqrt{\kappa})$ and suppose $f : X \to (-\infty, \infty]$ is a proper, convex and lower semi-continuous function. If $x \in X$, $J_\lambda$ is defined by (6) and

$$\limsup_{n \to \infty} d(x, J_{\lambda_n} x) < \pi/(2\sqrt{\kappa})$$

(8)

for some sequence of positive real numbers $(\lambda_n)$ with $\lim_{n \to \infty} \lambda_n = \infty$, then $(J_{\lambda} x)_{\lambda > 0}$ converges to a minimum point of $f$ as $\lambda \to \infty$.

**Remark 3.6.** If we assume above that diam$(X) < \pi/(2\sqrt{\kappa})$, then (8) is satisfied. Moreover, one can show that $(J_{\lambda} x)_{\lambda > 0}$ actually converges to the minimum point of $f$ which is closest to $x$.

\cite{10} Chapter 4] studies energy functionals defined in an appropriate space of $L^2$-functions, which is a CAT(0) space if the functions take values in a CAT(0) space. Minimum points of such energy functionals are called generalized harmonic maps and their existence is proved via \cite{10} Theorem 3.1.1. In a similar way, Theorem 3.5 could prove to be useful for the study of energy functionals in an appropriate CAT($\kappa$) space of functions.

We focus next on the following splitting proximal point algorithm employed in the study of minimum points for a function $f : X \to (-\infty, \infty]$ which can be written as

$$f = \sum_{i=1}^{N} f_i,$$

(9)

where for each $i \in \{1, \ldots, N\}$, $f_i : X \to (-\infty, \infty]$ is proper, convex and lower semi-continuous. To this end we will use instead of the resolvent of $f$, the resolvents of the functions $f_i$,

$$J_{\lambda}^i(x) = \arg\min_{y \in X} \left[ f_i(y) + \frac{1}{\lambda} \Psi_x(y) \right]$$

and given $(\lambda_j)$ a sequence of positive real numbers and $x_0 \in X$, the sequence $(x_n)$ is defined by

$$x_{jN+1} = J_{\lambda_j}^1(x_{jN}), \quad x_{jN+2} = J_{\lambda_j}^2(x_{jN+1}), \quad \ldots, \quad x_{jN+N} = J_{\lambda_j}^N(x_{jN+N-1}).$$

(10)

This method was recently studied in CAT(0) spaces in \cite{3} Theorem 3.4] and we adapt the proof strategy to our setting.

**Theorem 3.7.** Let $(X,d)$ be a compact CAT($\kappa$) space with $\kappa > 0$ such that for every $v, w \in X$, $d(v, w) < \pi/(2\sqrt{\kappa})$. Suppose $f : X \to (-\infty, \infty]$ is a function of the form (9) which attains its minimum. For any $x_0 \in X$ and any sequence of positive real numbers $(\lambda_j)$ with $\sum_{j \geq 0} \lambda_j = \infty$ and

$$\sum_{j \geq 0} \lambda_j^2 < \infty,$$

let $(x_n)$ be defined by (10). If there exists $L > 0$ such that for every $j \in \mathbb{N}$ and $i \in \{1, \ldots, N\}$,

$$f_i(x_{jN}) - f_i(x_{jN+i}) \leq Ld(x_{jN}, x_{jN+i})$$

(11)

and

$$f_i(x_{jN+i-1}) - f_i(x_{jN+i}) \leq Ld(x_{jN+i-1}, x_{jN+i}),$$

(12)

then $(x_n)$ converges to a minimum point of $f$.
Proof. Let \( z \) be a minimum point of \( f \). For \( j \in \mathbb{N} \) and \( i \in \{1, \ldots, N\} \), apply Lemma 3.1 to the function \( f_i \), the resolvent \( J_{\lambda_j}^i \) and the points \( x_{jN+i-1} \) and \( z \) to get that

\[
\lambda_j (f_i(x_{jN+i}) - f_i(z)) \leq 2 \left( 1 + \frac{1}{\cos^2 \left( \sqrt{\kappa d(x_{jN+i-1}, x_{jN+i})} \right)} \right) \times \left( \cos \left( \sqrt{\kappa d(z, x_{jN+i})} \right) - \cos \left( \sqrt{\kappa d(z, x_{jN+i-1})} \right) \right).
\]

Let \( m \in \{1, \ldots, N\} \). Because

\[
f_m(x_{jN+m}) + \frac{1}{\lambda_j} \Psi_{x_{jN+m-1}}(x_{jN+m}) \leq f_m(x_{jN+m-1})
\]

and

\[
\Psi_{x_{jN+m-1}}(x_{jN+m}) = \frac{\sin^2 \left( \sqrt{\kappa d(x_{jN+m-1}, x_{jN+m})} \right)}{\cos \left( \sqrt{\kappa d(x_{jN+m-1}, x_{jN+m})} \right)} \geq \frac{\kappa d(x_{jN+m-1}, x_{jN+m})^2}{4},
\]

by \([12]\), we have that

\[
\frac{\kappa d(x_{jN+m-1}, x_{jN+m})^2}{4} \leq \lambda_j (f_m(x_{jN+m-1}) - f_m(x_{jN+m})) \leq \lambda_j Ld(x_{jN+m-1}, x_{jN+m}),
\]

from where

\[
d(x_{jN+m-1}, x_{jN+m}) \leq 4\lambda_j L/\kappa. \tag{13}
\]

Take \( j_0 \in \mathbb{N} \) such that \( \lambda_j \leq 1 \) for \( j \geq j_0 \) and denote \( \alpha = 1 + 1/\cos^2(4L/\sqrt{\kappa}) \). Then if \( j \geq j_0 \),

\[
1 + 1/\cos^2 \left( \sqrt{\kappa d(x_{jN+i-1}, x_{jN+i})} \right) \leq \alpha \quad \text{and so}
\]

\[
\lambda_j (f_i(x_{jN+i}) - f_i(z)) \leq 2\alpha \left( \cos \left( \sqrt{\kappa d(z, x_{jN+i})} \right) - \cos \left( \sqrt{\kappa d(z, x_{jN+i-1})} \right) \right).
\]

Note that

\[
\sum_{i=1}^{N} (f_i(x_{jN+i}) - f_i(z)) = f(x_{jN}) - f(z) + \sum_{i=1}^{N} (f_i(x_{jN+i}) - f_i(x_{jN})).
\]

Hence, for all \( j \geq j_0 \),

\[
\lambda_j (f(x_{jN}) - f(z)) \leq 2\alpha \left( \cos \left( \sqrt{\kappa d(z, x_{jN+N})} \right) - \cos \left( \sqrt{\kappa d(z, x_{jN})} \right) \right) + \lambda_j \sum_{i=1}^{N} (f_i(x_{jN}) - f_i(x_{jN+i})).
\]

Using \([13]\) we obtain that

\[
d(x_{jN}, x_{jN+i}) \leq d(x_{jN}, x_{jN+1}) + \ldots + d(x_{jN+i-1}, x_{jN+i}) \leq 4i\lambda_j L/\kappa,
\]

which, by \([11]\), yields that for \( j \geq j_0 \),

\[
\lambda_j (f(x_{jN}) - f(z)) \leq 2\alpha \left( \cos \left( \sqrt{\kappa d(z, x_{(j+1)N})} \right) - \cos \left( \sqrt{\kappa d(z, x_{jN})} \right) \right) + 2N(N+1)\lambda_j^2 L^2/\kappa. \tag{14}
\]
Since $z$ is a minimum point of $f$, we have that for $j \geq j_0$,
\[
\cos \left( \sqrt{\kappa d(z, x_{j+1}N)} \right) \geq \cos \left( \sqrt{\kappa d(z, x_{jN})} \right) - N(N + 1)\lambda_j^2 L^2 / (\kappa \alpha),
\]
which, by Lemma 2.1, shows that the sequence $(\cos (\sqrt{\kappa d(z, x_{jN})}))_j$ is convergent and so the sequence $(d(z, x_{jN}))_j$ converges too. Moreover, using (14), we get that
\[
\sum_{j=0}^{\infty} \lambda_j (f(x_{jN}) - f(z)) < \infty.
\]
This implies that there exists a subsequence $(x_{j_l}N)_l$ of $(x_{jN})$ such that $\lim_{l \to \infty} f(x_{j_l}N) = f(z)$. We may assume that $(x_{j_l}N)_l$ converges to some $p \in X$ (otherwise take a convergent subsequence of it). Since $f$ is lower semi-continuous, $f(p) \leq \lim_{l \to \infty} f(x_{j_l}N) = f(z)$, so $p$ is a minimum point of $f$, which means that $(d(p, x_{j_l}N))_j$ is convergent and must converge to 0 since $(x_{j_l}N)_l$ converges to $p$. Now, one only needs to use (13) to obtain that $(x_{jN+m})_j$ converges to $p$ for all $m \in \{1, \ldots, N\}$ which finally yields that $(x_n)$ converges to $p$.

**Remark 3.8.** Other choices for the function $\Psi_x$ in the definition of the resolvent (6) are also possible, even if the image of $\Psi_x$ is not $[0, \infty)$. For instance, it is easy to see that the resolvent is well-defined and that similar convergence results hold for the sequence generated by the proximal point algorithm when considering $\Psi_x = \Psi^1_x$ or $\Psi_x = \Psi^2_x$.

**Remark 3.9.** Although the proximal point algorithm as given in [2] can be applied in any CAT$(\kappa)$ space with $\kappa \leq 0$, as before one could also consider for $\kappa < 0$ another algorithm of this type by taking, for example, in (6) $\Psi_x : X \to [0, \infty)$, $\Psi_x(y) = -\frac{1}{\kappa} \left( \cosh \left( \sqrt{-\kappa d(y, x)} \right) - \frac{1}{\cosh \left( \sqrt{-\kappa d(y, x)} \right)} \right)$.

Note that when $\kappa \nearrow 0$, we obtain the squared distance function as for CAT(0) spaces. It turns out that the resolvent is indeed well-defined and that analogous convergence results can be proved in this case too.

**Acknowledgements:** The authors have been partially supported by DGES (MTM2015-65242-C2-1-P). A. Nicolae would also like to acknowledge the Juan de la Cierva-incorporación Fellowship Program of the Spanish Ministry of Economy and Competitiveness.

**References**

[1] D. Ariza-Ruiz, L. Leustean and G. López-Acedo, *Firmly nonexpansive mappings in classes of geodesic spaces*, Trans. Amer. Math. Soc. 366 (2014), 4299–4322.

[2] M. Bačák, *The proximal point algorithm in metric spaces*, Israel J. Math. 194 (2013), 689–701.

[3] M. Bačák, *Computing means and medians in Hadamard spaces*, SIAM J. Optim. 24 (2014), 1542–1566.

[4] M. Bačák, *Convex analysis and optimization in Hadamard spaces*, De Gruyter, Berlin, 2014.

[5] D. P. Bertsekas, *Incremental proximal methods for large scale convex optimization*, Math. Program. 129 (2011), 163–195.

[6] M. R. Bridson and A. Häfliger, *Metric spaces of non-positive curvature*, Springer-Verlag, Berlin, 1999.
[7] R. Espínola and A. Fernández-León, CAT(κ)-spaces, weak convergence and fixed points, J. Math. Anal. Appl. 353 (2009), 410–427.

[8] O. P. Ferreira and P. R. Oliveira, Proximal point algorithm on Riemannian manifolds, Optimization 51 (2002), 257–270.

[9] J. Jost, Convex functionals and generalized harmonic maps into spaces of non positive curvature, Comment. Math. Helvetici 70 (1995), 659–673.

[10] J. Jost, Nonpositive Curvature: Geometric and Analytic Aspects, in: Lect. in Math., ETH Zürich, Birkhäuser, 1997.

[11] Y. Kimura and F. Kohsaka, Spherical nonspreadingness of resolvents of convex functions in geodesic spaces, J. Fixed Point Theory Appl. 18 (2016), 93–115.

[12] Y. Kimura and F. Kohsaka, The proximal point algorithm in geodesic spaces with curvature bounded above, Linear Nonlinear Anal. (accepted).

[13] C. Li, G. López and V. Martín-Márquez, Monotone vector fields and the proximal point algorithm on Hadamard manifolds, J. London Math. Soc. 79 (2009), 663–683.

[14] C. Li and J. C. Yao, Variational inequalities for set-valued vector fields on Riemannian manifolds: convexity of the solution set and the proximal point algorithm, SIAM J. Control Optim. 50 (2012), 2486–2514.

[15] T. C. Lim, Remarks on some fixed point theorems, Proc. Amer. Math. Soc. 60 (1976), 179–182.

[16] B. Martinet, Régularisation d’inéquations variationnelles par approximations sucessives, Rev. Française Informat. Recherche Opérationnelle 4 (1970), 154–158.

[17] U. F. Mayer, Gradient flows on nonpositively curved metric spaces and harmonic maps, Comm. Anal. Geom. 6 (1998), 199–253.

[18] S. Ohta, M. Pálfia, Discrete-time gradient flows and law of large numbers in Alexandrov spaces, Calc. Var. Partial Differential Equations 54 (2015), 1591–1610.

[19] T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976), 877–898.

[20] C. Zălinescu, On uniformly convex functions, J. Math. Anal. Appl. 95 (1983), 344–374.