Canad. Math. Bull. Vol. **65** (4), 2022, pp. 994–1003
http://dx.doi.org/10.4153/S000843952200008X

A note on Mosco convergence in CAT(0) spaces

A. Bërdëllima

Abstract. In this note, we show that in a complete CAT(0) space pointwise convergence of proximal mappings under a certain normalization condition implies Mosco convergence.

1 Introduction

A classical result of [2] establishes a set of equivalences between Mosco convergence, convergence of proximal mappings and of Moreau–Yosida approximates for a sequence of proper closed convex functions defined on a smooth reflexive Banach space. This result has a natural extension to any metric space equipped with a convex structure that allows for an appropriate notion of weak convergence. We consider these equivalences in the setting of a complete CAT(0) space. A CAT(0) space is a uniquely geodesic metric space that is endowed with a canonical convex structure and allows for a suitable notion of weak convergence. The main feature that distinguishes CAT(0) spaces from other geodesic metric spaces is the fact that every geodesic triangle in a CAT(0) space is at least as thin as its comparison triangle in the Euclidean plane. These spaces were formally introduced by Gromov [11] in recognition of the seminal work of Alexandrov [1] and are often referred to as spaces of nonpositive curvature, e.g., [6, 8]. The main motivation for studying Mosco convergence in a CAT(0) space is the implications it has in the theory of the gradient flow initiated by Mayer [15] and Jost [13]. More recently, Mosco convergence is applied in numerical schemes for the so called metamorphosis models on Hadamard manifolds [10]. In relation to the theory of the gradient flow, Bačak [3] proved that Mosco convergence implies pointwise convergence of Moreau approximates and of proximal mappings, and in turn convergence of proximal mappings yields pointwise convergence of the gradient flow semigroup. Later in [5], Bačak et al. showed that pointwise convergence of Moreau approximates yields Mosco convergence. However, it is not known whether pointwise convergence of the proximal mappings implies Mosco convergence. That convergence of proximal mappings alone is not enough was known to Bačak [4]. Indeed, consider a sequence of constant functions $0, 1, 0, 1, \dots$ defined on $\mathbb{R}$. These are closed and convex functions but they do not converge in the sense of Mosco to any function f, however, their proximal mappings $x \mapsto J_\lambda x$ equal the identity for all $\lambda > 0$.

Received by the editors September 25, 2021; revised January 11, 2022; accepted January 12, 2022.
Published online on Cambridge Core January 17, 2021.
AMS subject classification: **46N10**, 54E50.
Keywords: CAT(0) spaces, Mosco convergence, Moreau–Yosida approximate, proximal mapping.

It is the purpose of this note to establish a suitable condition under which convergenve of proximal mappings guarantees Mosco convergence for a sequence of proper closed convex functions.

2 Preliminaries

We follow standard terminology from metric geometry theory about CAT(0) spaces, e.g., [8, 9]. By (X,d) we denote a complete CAT(0) space, unless otherwise stated. Given $x \in X$ let

$$\Gamma_x(X) := \{\gamma : [0,1] \to X \mid \gamma(0) = x\}$$

denote the set of all geodesic segments emanating from x. A convex combination of two elements $x, y \in X$ with parameter $t \in [0,1]$ is an element $x_t \in [x,y]$ satisfying $d(x,x_t) = td(x,y)$. We write $x_t := (1-t)x \oplus ty$. Such a convex combination is uniquely determined. We often denote by $[x,y]$ the unique geodesic segment connecting x with y. A set $C \subseteq X$ is a convex set if for every $x, y \in C$ and $t \in [0,1]$ it holds that $(1-t)x \oplus ty \in C$. A set C is closed if its complement $X \backslash C$ is open in the usual metric topology. In particular, geodesic segments are closed and convex. Let $d(x,C) := \inf\{d(x,y) \mid y \in C\}$ and $P_C x := \{z \in C \mid d(x,z) = d(x,C)\}$. We use the following definition for weak convergence in a CAT(0) space which is due to [12].

Definition 2.1 A sequence (x_n) weakly converges to x and we write $x_n \overset{w}{\to} x$ if and only if $\lim_{n\to+\infty} P_\gamma x_n = x$ for every $\gamma \in \Gamma_x(X)$.

This notion of convergence is well defined since projection onto a closed convex set always exists and is unique. On bounded sets, weak convergence coincides with the so-called Δ-convergence introduced by Lim [14] and shares with it the desirable properties of Opial and Kadec–Klee. Moreover, every bounded sequence has a weakly convergent subsequence [7, Chapter 3]. Moreover, weak limits are unique. Indeed if $x_n \overset{w}{\to} x$ and $x_n \overset{w}{\to} y$ then for $\gamma := [x,y]$ and $\tilde{\gamma} := [y,x]$ we obtain $x = \lim_{n\to+\infty} P_\gamma x_n = \lim_{n\to+\infty} P_{\tilde{\gamma}} x_n = y$.

Definition 2.2 A sequence of functions $f^n : X \to (-\infty, +\infty]$ is said to be Mosco convergent to $f : X \to (-\infty, +\infty]$ and we write $M - \lim_{n\to+\infty} f^n = f$ if for each $x \in X$:

(1) $f(x) \le \liminf_{n\to+\infty} f^n(x_n)$ whenever $x_n \overset{w}{\to} x$ and
(2) there exists some sequence $(y_n) \subset X$ such that $y_n \to x$ and $f(x) \ge \limsup_{n\to+\infty} f^n(y_n)$.

One of the difficulties in Attouch's theorem is that it involves the dual space in its formulation. This is problematic in the setting of a CAT(0) space since a proper dual space corresponding to the weak convergence in Definition 2.1 is not yet fully understood. However the following concept becomes helpful and plays the role of a "subgradient."[1]

[1] Refer to [16] for this and related concepts in convex analysis.

https://doi.org/10.4153/S000843952200008X Published online by Cambridge University Press

Definition 2.3 Let $f : X \to (-\infty, +\infty]$ be some (extended) real-valued function and let $x \in \operatorname{dom} f := \{x \in X \mid f(x) < +\infty\}$. The absolute slope of f at x is defined as

$$|\partial f|(x) := \limsup_{y\to x} \frac{\max\{f(x) - f(y), 0\}}{d(x,y)}. \tag{2.1}$$

If $f(x) = +\infty$ we set $|\partial f|(x) := +\infty$.

To this end, we assume that our functions are real valued and satisfy:

(1) (proper) there exists no $x \in X$ such that $f(x) = -\infty$ and $f \not\equiv +\infty$.
(2) (convex) $f((1-t)x \oplus ty) \le (1-t)f(x) + t(y)$ for all $x, y \in X$ and $t \in [0,1]$.
(3) (closed) $f(x) \le \liminf_{y\to x} f(y)$ for every $x \in X$.

Lemma 2.1 [3, Lemma 5.1.2] *Let $f : X \to (-\infty, +\infty]$ be a closed convex function. Then*

$$|\partial f|(x) = \sup_{y\in H\setminus\{x\}} \frac{\max\{f(x) - f(y), 0\}}{d(x,y)}, \quad x \in \operatorname{dom} f. \tag{2.2}$$

Given $x \in X$ and a geodesic segment γ containing x, we define the function $\operatorname{sign}(\cdot;\gamma) : \gamma \to \{-1,1\}$ as $\operatorname{sign}(y;\gamma) = 1$ if $y \in (x, \gamma(1)]$ and $\operatorname{sign}(y;\gamma) = -1$ if $y \in [\gamma(0), x]$. Another notion for the slope of a function f is that of the geodesic derivative.

Definition 2.4 Let $f : X \to (-\infty, +\infty]$. The geodesic lower derivative of f at $x \in X$ along a geodesic γ containing x is defined as[2]

$$f'_-(x;\gamma) := \liminf_{y \xrightarrow{\gamma} x} \frac{f(y) - f(x)}{\operatorname{sign}(y;\gamma)\, d(y,x)}. \tag{2.3}$$

Analogously, the geodesic upper derivative, denoted by $f'_+(x;\gamma)$, is defined with $\liminf$ replaced by $\limsup$. If both limits exist and coincide then we say that f is geodesically differentiable at x along γ and denote it by $f'(x;\gamma)$.

For a given closed convex function $f : X \to (-\infty, +\infty]$ and parameter $\lambda > 0$ the *Moreau approximate* f_λ of f is defined as

$$f_\lambda(x) := \inf_{y\in X} \left\{ f(y) + \frac{1}{2\lambda} d(y,x)^2 \right\}, \text{ for each } x \in X \tag{2.4}$$

and the proximal mapping of f

$$J_\lambda x := \arg\min_{y\in X} \left\{ f(y) + \frac{1}{2\lambda} d(y,x)^2 \right\}, \text{ for each } x \in X. \tag{2.5}$$

In general, $J_\lambda x$ can be a multivalued operator if X is some arbitrary metric space, however, in a CAT(0) space $J_\lambda x$ exists and is always unique (see [3]). We use also the following two results for complete CAT(0) spaces.

[2] A related notion *directional derivative* is found in [17].

Theorem 2.2 [3, Theorem 5.2.4] *Let $f^n : X \to (-\infty, +\infty]$ a sequence of closed convex functions. If $M - \lim_{n\to+\infty} f^n(x) = f(x)$, then it holds $\lim_{n\to+\infty} f^n_\lambda(x) = f_\lambda(x)$ and $\lim_{n\to+\infty} J^n_\lambda x = J_\lambda x$ for each $x \in X$.*

Theorem 2.3 [5, Theorem 3.2] *Let $f, f^n : X \to (-\infty, +\infty]$ be a sequence of closed convex functions. If $\lim_{n\to+\infty} f^n_\lambda(x) = f_\lambda(x)$ then $M - \lim_{n\to+\infty} f^n(x) = f(x)$ for all $x \in X$.*

3 Main result

3.1 A normalization condition

Definition 3.1 A sequence of functions $(f^n)_{n\in\mathbb{N}}, f$ satisfies the normalization condition if there exists some sequence $(x_n)_{n\in\mathbb{N}} \subset X$ and $x \in X$ such that $x_n \to x$, $f^n(x_n) \to f(x)$ and $|\partial f^n|(x_n) \to |\partial f|(x)$ as $n \to +\infty$.

Lemma 3.1 *A sequence of closed convex functions $(f^n)_{n\in\mathbb{N}}, f : X \to (-\infty, +\infty]$ satisfies the normalization condition whenever $M - \lim_{n\to+\infty} f^n = f$.*

Proof Let $x_0 \in X$ and $M - \lim_{n\to+\infty} f^n = f$. Then by Theorem 2.2, we have $\lim_{n\to+\infty} J^n_\lambda x_0 = J_\lambda x_0$ for any $\lambda > 0$. Take $x_n := J^n_\lambda x_0$ and $x := J_\lambda x_0$. By definition of J_λ we have

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

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

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

implying $\limsup_{n\to+\infty} f^n(x_n) \le \limsup_{n\to+\infty} f^n(\xi_n)_{n\in\mathbb{N}}$. On the other hand, by definition of Mosco convergence there exists a sequence $(\xi_n)_{n\in\mathbb{N}}$ such that $\limsup_{n\to+\infty} f^n(\xi_n) \le f(x)$. Hence, $\limsup_{n\to+\infty} f^n(x_n) \le f(x)$. Moreover, $\lim_{n\to+\infty} x_n = x$ implies, in particular, that $x_n \overset{w}{\to} x$. Again by definition of Mosco convergence, we obtain $f(x) \le \liminf_{n\to+\infty} f^n(x_n)$. Therefore, $f(x) = \lim_{n\to+\infty} f^n(x_n)$ as desired. Next, we need to show the property about the slopes. Note that by Lemma 2.1, we have

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

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

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

https://doi.org/10.4153/S000843952200008X Published online by Cambridge University Press

which in turn yields

$$\frac{\max\{f(x)-\limsup_{n\to+\infty} f^n(\xi_n),0\}}{d(x,y)} \le \liminf_{n\to+\infty} |\partial f^n|(x_n).$$

Using $\limsup_{n\to+\infty} f^n(\xi_n) \le f(y)$, we get

$$\frac{\max\{f(x)-f(y),0\}}{d(x,y)} \le \liminf_{n\to+\infty} |\partial f^n|(x_n).$$

Because the last inequality holds for any $y \in X\backslash\{x_n\}$ then $|\partial f|(x) \le \liminf_{n\to+\infty} |\partial f^n|(x_n)$. Now by Definition 2.3, we obtain

$$|\partial f^n|(x_n) \le \frac{\max\{f^n(x_n)-f^n(y_n),0\}}{d(x_n,y_n)} + \varepsilon_n, \quad \forall n \in \mathbb{N}$$

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 it has a weakly convergent subsequence (y_{n_k}). But $\mathrm{cl}\,\mathbb{B}(x,\delta)$ is a closed convex set and since weak convergence coincides on bounded sets with the Δ-convergence then by [3, Lemma 3.2.1] $y_{n_k} \overset{w}{\to} y \in \mathrm{cl}\,\mathbb{B}(x,\delta)$. One can choose (ε_n) such that $\lim_{k\to+\infty} \varepsilon_{n_k} = 0$. Moreover, $d(x,\cdot)$ is weakly lsc [3, Corollary 3.2.4] implying

$$\limsup_{k\to+\infty} |\partial f^{n_k}|(x_{n_k}) \le \frac{\max\{f(x)-\liminf_k f^{n_k}(y_{n_k}),0\}}{d(x,y)}.$$

By definition of Mosco convergence, it follows that $\liminf_{k\to+\infty} f^{n_k}(y_{n_k}) \ge f(y)$. Hence,

$$\limsup_{n\to+\infty} |\partial f^n|(x_n) \le \limsup_{k\to+\infty} |\partial f^{n_k}|(x_{n_k}) \le \frac{\max\{f(x)-f(y),0\}}{d(x,y)}.$$

The last inequality implies $\limsup_{n\to+\infty} |\partial f^n|(x_n) \le |\partial f|(x)$. ■

Definition 3.2 A family of functions $f^n : X \to (-\infty,+\infty]$ is said to be equi locally Lipschitz if for any bounded set $K \subseteq X$, there is a constant $C_K > 0$ such that

$$|f^n(x)-f^n(y)| \le C_K d(x,y), \quad \forall x,y \in K, \forall n \in \mathbb{N}. \tag{3.1}$$

Lemma 3.2 *Let $f^n : X \to (-\infty,+\infty]$ be a sequence of closed convex functions such that $\lim_{n\to+\infty} f^n_\lambda(x_0) = \alpha_0 \in \mathbb{R}$ for some $x_0 \in X$ and some $\lambda > 0$. Then $(f^n_\lambda)_{n\in\mathbb{N}}$ are equi locally Lipschitz functions.*

Proof By virtue of [2, Theorem 2.64 (ii)], it suffices to show that there is $r > 0$ and $x_0 \in X$ such that $f^n(x) + r(d(x,x_0)^2 + 1) \ge 0$ for all $x \in X$ and all $n \in \mathbb{N}$. Let $x_0 \in X$ be such that $\lim_{n\to+\infty} f^n_\lambda(x_0) = \alpha_0 \in \mathbb{R}$. Notice that by definition of Moreau envelope, we have

$$f^n(x) \ge f^n_\lambda(x_0) - \frac{1}{2\lambda} d(x_0,x)^2 \ge \alpha_0 - \delta - \frac{1}{2\lambda} d(x_0,x)^2$$

https://doi.org/10.4153/S000843952200008X Published online by Cambridge University Press

for some $\delta > 0$ and sufficiently large n. If one takes $\delta = \alpha_0 + 1/2\lambda$, then one gets

$$f^n(x) \geq -\frac{1}{2\lambda}(d(x_0, x)^2 + 1), \quad \forall x \in X.$$

For any $r \geq 1/2\lambda$, we obtain $f^n(x) + r(d(x_0, x)^2 + 1) \geq 0$ for all $x \in X$ and all $n \in \mathbb{N}$.

3.2 Attouch's theorem for complete $\mathrm{CAT}(0)$ spaces

Theorem 3.3 *Let $f^n, f : X \to (-\infty, +\infty]$ be proper closed convex functions such that*

(1) $\forall \lambda > 0, \forall x \in X$ *it holds* $\lim_{n\to+\infty} J_\lambda^n x = J_\lambda x$.
(2) $(f^n)_{n\in\mathbb{N}}$ *satisfies the normalization condition.*
(3) $\lim_{n\to+\infty}(f_\lambda^n)'(x;\gamma) = f_\lambda'(x;\gamma)$ *for all* $\gamma \in \Gamma_{x_0}(X)$ *and* $x \in \gamma$.

Then $\lim_{n\to+\infty} f_\lambda^n(x) = f_\lambda(x)$ *for any* $\lambda > 0, x \in X$.

Proof Let $(f^n)_{n\in\mathbb{N}}, f$ satisfy the normalization condition. Then there exists $(x_n), x_0 \subset X$ such that $\lim_{n\to+\infty} x_n = x_0, \lim_{n\to+\infty} f^n(x_n) = f(x_0)$ and $\lim_{n\to+\infty} |\partial f^n|(x_n) = |\partial f|(x_0)$. Assume that $\lambda > 0$. First, we claim that $\lim_{n\to+\infty} f_\lambda^n(x_0) = f_\lambda(x_0)$. Introduce the variables $u_n := J_\lambda^n x_n$ for each $n \in \mathbb{N}$ and $u_0 := J_\lambda x_0$. Note that by assumption (3.3) for each fixed $m \in \mathbb{N}$, we have $\lim_{n\to+\infty} J_\lambda^n x_m = J_\lambda x_m$. Since the mapping $x \mapsto J_\lambda x$ is nonexpansive and therefore continuous, then $\lim_m J_\lambda x_m = J_\lambda x_0$. By triangle inequality $d(J_\lambda^n x_n, J_\lambda x_0) \leq d(J_\lambda^n x_n, J_\lambda^n x_m) + d(J_\lambda^n x_m, J_\lambda x_0)$ and nonexpansiveness of J_λ^n, we have

$$d(J_\lambda^n x_n, J_\lambda x_0) \leq d(x_n, x_m) + d(J_\lambda^n x_m, J_\lambda x_0).$$

Passing in the limit as $m, n \to +\infty$, we get $\lim_{n\to+\infty} u_n = \lim_{n\to+\infty} J_\lambda^n x_n = J_\lambda x_0 = u_0$. On the other hand,

$$|f^n(u_n) - f(u_0)| \leq |f^n(u_n) - f^n(x_n)| + |f^n(x_n) - f(x_0)| + |f(x_0) - f(u_0)|.$$

By normalization condition (3.3) and using $\lim_{\lambda\downarrow 0} u_n = \lim_{\lambda\downarrow 0} J_\lambda^n x_n = x_n$ and $\lim_{\lambda\downarrow 0} u_0 = \lim_{\lambda\downarrow 0} J_\lambda x_0 = x_0$ together with the lsc of f^n and f imply in the limit as $\lambda \downarrow 0$ and $n \to +\infty$ that $\lim_{n\to+\infty} f^n(u_n) = f(u_0)$. Again by definition of Moreau envelope as $n \to +\infty$

$$f_\lambda^n(x_n) = f^n(u_n) + \frac{1}{2\lambda} d(x_n, u_n)^2 \to f(u_0) + \frac{1}{2\lambda} d(x_0, u_0)^2 := f_\lambda(x_0).$$

Note that

$$f_\lambda^n(x_0) \leq f^n(x_n) + \frac{1}{2\lambda} d(x_0, x_n)^2 \to f(x_0) \quad \text{as} \quad n \to +\infty.$$

On the other hand, we have

$$\begin{aligned} f_\lambda^n(x_0) \geq f^n(J_\lambda^n x_0) &\geq f^n(x_n) - |\partial f^n|(x_n) d(J_\lambda^n x_0, x_n) \\ &\to f(x_0) - |\partial f|(x_0) d(J_\lambda x_0, x_0) > -\infty \quad \text{as} \quad n \to +\infty. \end{aligned}$$

In particular we get $-\infty < \liminf_{n\to+\infty} f_\lambda^n(x_0) \leq \limsup_{n\to+\infty} f_\lambda^n(x_0) < +\infty$ (we can assume wlog that $x_0 \in \operatorname{dom} f$). By Lemma 3.2, we get that $(f_\lambda^n)_{n\in\mathbb{N}}$ is equi locally

Lipschitz in X. Then for any bounded domain $K \subseteq X$, there is $C_K > 0$ such that

$$|f_\lambda^n(x) - f_\lambda^n(y)| \le C_K d(x,y), \quad \forall x, y \in K, \forall n \in \mathbb{N}.$$

From this and the estimate

$$\begin{aligned}|f_\lambda^n(x_0) - f_\lambda(x_0)| &\le |f_\lambda^n(x_0) - f_\lambda^n(x_n)| + |f_\lambda^n(x_n) - f_\lambda(x_0)| \\ &\le C_K d(x_n, x_0) + |f_\lambda^n(x_n) - f_\lambda(x_0)|\end{aligned}$$

follows that $\lim_{n\to+\infty} f_\lambda^n(x_0) = f_\lambda(x_0)$. Now define $g_{n,\lambda}(t) := f_\lambda^n(x_t)$ and $g_\lambda(t) := f_\lambda(x_t)$, where $x_t := (1-t)x_0 \oplus tx$ and $x \in X$ is arbitrary. Consider

$$g'_{n,\lambda}(t) := \lim_{s\to 0} \frac{g_{n,\lambda}(t+s) - g_{n,\lambda}(s)}{s}, \ t \in (0,1).$$

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

$$g_{n,\lambda}(1) = g_{n,\lambda}(0) + \int_0^1 g'_{n,\lambda}(t)\,dt. \tag{3.2}$$

On the other hand, $g'_{n,\lambda}(t) = (f_\lambda^n)'(x_t;\gamma)d(x_0,x)$ where $\gamma \in \Gamma_{x_0}(X)$ connects x_0 with x and $x_t \in \gamma$. Assumption (3.3) implies $\lim_{n\to+\infty} g'_{n,\lambda}(t) = g'_\lambda(t)$ for all $t \in [0,1]$. Moreover equi locally Lipschitz property of $(f_\lambda^n)_{n\in\mathbb{N}}$ implies that $\sup_n |g'_{n,\lambda}(t)| \le C_K d(x_0,x)$ for any bounded domain K around x_0 and $x \in K$. Upon replacing $g_{n,\lambda}(1)$ with $f_\lambda^n(x)$ and $g_{n,\lambda}(0)$ with $f_\lambda^n(x_0)$ in (3.2), by Lebesgue dominated convergence theorem we obtain in the limit

$$\lim_{n\to+\infty} f_\lambda^n(x) = f_\lambda(x_0) + \int_0^1 \lim_{n\to+\infty} g'_{n,\lambda}(t)\,dt = f_\lambda(x_0) + \int_0^1 g'_\lambda(t)\,dt = f_\lambda(x).$$

Example 3.4 Consider again the sequence of functions $f^n : [0,1] \to \mathbb{R}$ given by

$$f^n(x) = \begin{cases} 0 & n \text{ odd}, \\ 1 & n \text{ even}, \end{cases} \qquad \forall x \in [0,1].$$

Let $f(x) = 1/2$ for all $x \in [0,1]$. Then $(f^n)_{n\in\mathbb{N}}, f$ is a family of proper closed convex functions satisfying $\lim_{n\to+\infty} J_\lambda^n x = J_\lambda x$ for any $\lambda > 0$ and every $x \in [0,1]$. However, assumption (2) is not fulfilled since for any sequence $(x_n)_{n\in\mathbb{N}} \subset X$ and for any $x_0 \in X$ such that $\lim_{n\to+\infty} x_n = x_0$ we have $\liminf_{n\to+\infty} f^n(x_n) = 0 < 1/2 = f(x_0)$. Moreover, for any $x \in [0,1]$ it holds that $\liminf_{n\to+\infty} f^n(x) = 0 < 1/2 = f(x)$ consequently since $f_\lambda^n(x) = f^n(x)$ and $f_\lambda(x) = f(x)$ for all $\lambda > 0$ it follows that $\lim_{n\to+\infty} f_\lambda^n(x) \ne f_\lambda(x)$. Note that in this case, $(f_\lambda^n)'(x;\gamma) = (f^n)'(x;\gamma) = 0$ and $(f_\lambda)'(x;\gamma) = f'(x;\gamma) = 0$ i.e., $\lim_{n\to+\infty}(f_\lambda^n)'(x;\gamma) = f'(x;\gamma) = 0$ for all $\gamma \in \Gamma_{x_0}(X)$ and $x \in \gamma$. This example shows that condition (2) cannot be removed.

Example 3.5 Let $(\mathcal{H}, \|\cdot\|)$ be a separable Hilbert space equipped with the canonical norm $\|\cdot\|$. Let $\{e_k\}_{k\in\mathbb{N}}$ be the usual orthonormal basis in $\mathcal{H}$. Denote by $X = \bigcup_{k\in\mathbb{N}} X_k$ where

$$X_k = \{x \in \mathcal{H} : x = t\,e_k, \ 0 \le t \le 1\}.$$

https://doi.org/10.4153/S000843952200008X Published online by Cambridge University Press

When X is equipped with the length metric, that we denote by d_ℓ, then (X, d_ℓ) is a complete CAT(0) space. Let $f^n : X \to \mathbb{R}$ be defined by the formula $f^n(x) = d_\ell(\mathbf{0}, x)/n$ for all $x \in X$ and $n \in \mathbb{N}$. Here, $\mathbf{0}$ is the usual origin in $\mathcal{H}$, i.e., $\{\mathbf{0}\} = \bigcap_{k \in \mathbb{N}} X_k$. Let $f : X \to \mathbb{R}$ be the identical zero function i.e., $f(x) = 0$ for all $x \in X$. Two elements $x, y \in X$ are equivalent whenever $x, y \in X_k$ for some $k \in \mathbb{N}$, and we write $x \sim y$. Note that

$$d_\ell(x, y) = \begin{cases} \|x - y\| & x \sim y, \\ \|x\| + \|y\| & \text{else.} \end{cases}$$

In particular, we obtain $f^n(x) = \|x\|/n$ for all $x \in X$ and $n \in \mathbb{N}$. Then

$$f^n(y) + \frac{1}{2\lambda} d_\ell^2(x, y) = \begin{cases} \dfrac{\|y\|}{n} + \dfrac{1}{2\lambda}\|x - y\|^2 & x \sim y, \\[2ex] \dfrac{\|y\|}{n} + \dfrac{1}{2\lambda}(\|x\| + \|y\|)^2 & \text{else.} \end{cases}$$

By virtue of triangle inquality $\|x - y\| \le \|x\| + \|y\|$, we obtain that

$$\inf_{y \in X}\left\{f^n(y) + \frac{1}{2\lambda} d_\ell^2(x, y)\right\} = \inf_{y \sim x}\left\{f^n(y) + \frac{1}{2\lambda} d_\ell^2(x, y)\right\}, \quad \forall x \in X.$$

Consequently

$$f_\lambda^n(x) = \inf_{y \sim x}\left\{f^n(y) + \frac{1}{2\lambda}\|x - y\|^2\right\}.$$

Since $x \sim y$ then there are $s, t \in [0, 1]$ such that $x = te_k$ and $y = se_k$ for some $k \in \mathbb{N}$. Hence, in terms of the variables s, t we can write

$$g(t) = f_\lambda^n(te_k) = \inf_{s \in [0,1]}\left\{\frac{s}{n} + \frac{1}{2\lambda}(s - t)^2\right\}$$

minimum of which is achieved for $s_n = t - \lambda/n$. Hence, $J_\lambda^n x = J_\lambda^n(t\,e_k) = (t - \lambda/n)\,e_k \to te_k = x$ as $n \to +\infty$. On the other hand, $J_\lambda x = x$ since trivially $f(x) = 0$ for all $x \in X$. This means that $\lim_{n\to+\infty} J_\lambda^n x = J_\lambda x$ for all $x \in X$ and for all $\lambda > 0$. Consequently condition (1) is satisfied. Now let $x_n = e_n/n$ for $n \in \mathbb{N}$. Then $\lim_{n\to+\infty} x_n = \mathbf{0}$ and $\lim_{n\to+\infty} f^n(x_n) = \lim_{n\to+\infty} 1/n^2 = 0 = f(\mathbf{0})$. Moreover, by virtue of Lemma 2.1, we have

$$\begin{aligned} |\partial f^n|(x_n) &= \sup_{y \in X \setminus \{x_n\}} \frac{\max\{\|x_n\|/n - \|y\|/n, 0\}}{d_\ell(x_n, y)} \\ &\le \sup_{y \in X \setminus \{x_n\}} \frac{\max\{\|x_n\|/n - \|y\|/n, 0\}}{\|x_n - y\|} \le \frac{1}{n} \to 0, \quad \text{as } n \to +\infty. \end{aligned}$$

Also $|\partial f|(\mathbf{0}) = 0$, hence $\lim_{n\to+\infty} |\partial f^n|(x_n) = |\partial f|(\mathbf{0})$. Therefore, condition (2) is also satisfied. Now let $\gamma \in \Gamma_0(X)$ and $x \in \gamma$. From the above calculations, we obtain that

$$f_\lambda^n(x) = f^n(J_\lambda^n x) + \frac{1}{2\lambda}\|x - J_\lambda^n x\|^2 = \frac{|t - \lambda/n|}{n} + \frac{\lambda}{2n^2} := g_{n,\lambda}(t). \tag{3.3}$$

For sufficiently large n we get $g'_{n,\lambda}(t) = 1/n^2 \to 0$ as $n \to +\infty$. The relation $g'_{n,\lambda}(t) = (f_\lambda^n)'(x;\gamma)\, d_\ell(\mathbf{0}, \gamma(1))$ yields $(f_\lambda^n)'(x;\gamma) \to 0$ as $n \to +\infty$. On the other side, $f_\lambda(x) = 0$ for all $x \in X$ implies $(f_\lambda)'(x;\gamma) = 0$ for every $\gamma \in \Gamma_{\mathbf{0}}(X)$ and all $x \in X$. Consequently, we get $\lim_{n\to+\infty}(f_\lambda^n)'(x;\gamma) = (f_\lambda)'(x;\gamma)$. This confirms condition (3). By Theorem 3.3, $\lim_{n\to+\infty} f_\lambda^n(x) = f_\lambda(x)$ for all $x \in X$ and $\lambda > 0$. Last conclusion can be easily verified directly from (3.3).

Theorem 3.6 *Let $(f^n), f : X \to (-\infty, +\infty]$ be proper closed and convex functions such that for some $x_0 \in X$ it holds that $\lim_{n\to+\infty}(f_\lambda^n)'(x;\gamma) = f'_\lambda(x;\gamma)$ for all $\gamma \in \Gamma_{x_0}(X)$ and $x \in \gamma$. Then the following statements are equivalent:*

(1) $M - \lim_{n\to+\infty} f^n = f$.
(2) *$\forall\lambda > 0, \forall x \in X$ it holds $\lim_{n\to+\infty} J_\lambda^n x = J_\lambda x$ and $(f^n)_{n\in\mathbb{N}}$ satisfies the normalization condition with some $(x_n)_{n\in\mathbb{N}}$ converging to x_0.*
(3) *$\lim_{n\to+\infty} f_\lambda^n(x) = f_\lambda(x)$ for any $\lambda > 0, \forall x \in X$.*

Proof Follows from Theorems 2.2, 2.3, and 3.3. ∎

Acknowledgment I am grateful to my advisor Prof. D. R. Luke for his helpful comments in improving the quality of this work. I would like to thank as well the anonymous referee for their qualitative suggestions. The author was supported by Deutscher Akademischer Austauschdienst (DAAD) and partially by Germany's Excellence Strategy–The Berlin Mathematics Research Center MATH+ (EXC-2046/1, Projektnummer: 390685689).

References

[1] A. D. Alexandrov, *A theorem on triangles in a metric space and some of its applications*. Trudy Mat. Inst. Steklova **38**(1951), 5–23.
[2] H. Attouch, *Variational convergence for functions and operators*. Applicable Mathematics Series, Pitman (Advanced Publishing Program), Boston, MA, 1984.
[3] M. Bačák, *Convex analysis and optimization in Hadamard spaces*. De Gruyter Series in Nonlinear Analysis and Applications, 22, De Gruyter, Berlin, 2014.
[4] M. Bačák, *Old and new challenges in Hadamard spaces*. Preprint, 2018. arXiv:1807.01355
[5] M. Bačák, M. Montag, and G. Steidl, *Convergence of functions and their Moreau envelopes on Hadamard spaces*. J. Appr. Theory (C) **224**(2017), 1–12.
[6] W. Ballmann, *Lectures on Spaces of Nonpositive Curvature*. Birkhäuser, Basel, 1995.
[7] A. Bërdëllima, *Investigations in Hadamard spaces*. Ph.D. thesis, Georg-August-Universität Göttingen, Göttingen, Germany, 2020.
[8] M. R. Bridson and A. Haefliger, *Metric spaces of nonpositive curvature*. A Series of Comprehensive Studies in Mathematics, 319, Birkhäuser Boston Inc., Boston, 1999.
[9] Y. Burago, D. Burago, and S. Ivanov, *A Course in Metric Geometry*. Graduate Studies in Mathematics, vol. 33, Birkhäuser Boston Inc., Boston, 2001.
[10] A. Effland, S. Neumayer, and M. Rumpf, *Convergence of the time discrete metamorphosis model on Hadamard manifolds*. SIAM J. Imag. Sci. **13**(2020), no. 2, 557–588.
[11] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, *vol. 152*, Birkhäuser Boston Inc., Boston, 1999 (Based on the 1981 French original).
[12] J. Jost, *Equilibrium maps between metric spaces*. Calc. Var. Partial Diff. Equat. **2**(1994), 173–204.
[13] J. Jost, *Nonlinear Dirichlet forms, new directions in Dirichlet forms*. AMP/IP Stud. Adv. Math **8**(1998), 1–47.
[14] T. C. Lim, *Remarks on some fixed point theorems*. Proc. Amer. Math. Soc. **60**(1976), 179–182.

[15] U. F. Mayer, *Gradient flows on nonpositively curved metric spaces and harmonic maps*. Comm. Anal. Geom. **6**(1998), 199–253.
[16] R. T. Rockafellar, *Convex analysis*. Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1970.
[17] L. I. L. Tean, D. Ariza-Ruiz, and G. López-Acedo, *Firmly nonexpansive mappings in classes of geodesic spaces*. Trans. of the Amer. Math. Soc. **366**(2014), 4299–4322.

Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany
e-mail: berdellima@math.tu-berlin.de

https://doi.org/10.4153/S000843952200008X Published online by Cambridge University Press