A new proof of the Lie-Trotter-Kato formula in Hadamard spaces

by

Miroslav Bačák

Preprint no.: 18 2013
A NEW PROOF OF THE LIE-TROTTER-KATO FORMULA
IN HADAMARD SPACES

MIROSLAV BÁČÁK

Abstract. The Lie-Trotter-Kato product formula was recently extended into Hadamard spaces by [Stojkovic, Adv. Calc. Var., 2012]. The aim of our short note is to give a simpler proof relying upon weak convergence. Unlike in the original proof, we completely avoid using ultrapowers of Hadamard spaces and any additional compactness assumptions.

1. Introduction

Let \( f : \mathcal{H} \to (-\infty, \infty] \) be a convex lower semicontinuous function (lsc) defined on an Hadamard space \((\mathcal{H}, d)\). For instance \( \mathcal{H} \) can be a Hilbert space and \( d \) its natural metric induced by the inner product. For the notation and terminology not explained here, the reader is referred to Section 2. Given \( \lambda > 0 \), define the resolvent of \( f \) as

\[
J_\lambda(x) := \arg \min_{y \in \mathcal{H}} \left[ f(y) + \frac{1}{2\lambda} d(x, y)^2 \right], \quad x \in \mathcal{H}.
\]

The mapping \( J_\lambda : \mathcal{H} \to \text{dom} f \) is well-defined for all \( \lambda > 0 \); see [12, Lemma 2] and [21, Theorem 1.8]. We also put \( J_0(x) := x \), for all \( x \in \mathcal{H} \). The gradient flow semigroup of \( f \) is given as

\[
S_t(x) := \lim_{n \to \infty} \left( J_{\frac{t}{n}} \right)^{(n)}(x), \quad x \in \text{dom} f,
\]

for any \( t \in [0, \infty) \). The limit in (2) is uniform with respect to \( t \) on bounded subintervals of \([0, \infty)\), and \((S_t)_{t \geq 0}\) is a strongly continuous semigroup of nonexpansive mappings on \( \mathcal{H} \); see [15, Theorem 1.3.13], and [21, Theorem 1.13].

Remark 1.1. If \( \mathcal{H} \) is a Hilbert space, \( u_0 \in \text{dom} f \), and we put \( u(t) = S_t(u_0) \), for \( t \in [0, \infty) \), we obtain a “classical” solution to the parabolic problem

\[
\dot{u}(t) \in -\partial f(u(t)), \quad t \in (0, \infty),
\]

with the initial condition \( u(0) = u_0 \).
In the present paper, we consider a function $f : \mathcal{H} \to (-\infty, \infty]$ of the form

$$f := \sum_{j=1}^{k} f_j,$$

where $f_j : \mathcal{H} \to (-\infty, \infty]$ are convex lsc functions, $j = 1, \ldots, k$, and $k \in \mathbb{N}$. This covers a surprisingly large spectrum of problems and has become a classical framework in applications; see for instance [7, Proposition 27.8] for the so-called parallel splitting algorithm. In the Hadamard space setting, functions of the form (3) naturally emerged in connection with the following example.

**Example 1.2.** Given a finite number of points $a_1, \ldots, a_N \in \mathcal{H}$, and positive weights $w_1, \ldots, w_N$ with $\sum_{n=1}^{N} w_n = 1$, we define the function

$$f(x) := \sum_{n=1}^{N} w_n d(x, a_n)^p, \quad x \in \mathcal{H},$$

where $p \in [1, \infty)$. Then $f$ is convex continuous and we are especially interested in two important cases:

(i) If $p = 1$, then $f$ becomes the objective function in the Fermat-Weber problem for optimal facility location. If, moreover, all the weights $w_n = \frac{1}{N}$, a minimizer is called a median of the points $a_1, \ldots, a_N$.

(ii) If $p = 2$, then a minimizer of $f$ is the barycenter of the probability measure

$$\mu = \sum_{n=1}^{N} w_n \delta_{a_n},$$

where $\delta_{a_n}$ stands for the Dirac measure at the point $a_n$. For further details on barycenters, the reader is referred to [14, Chapter 3] and [27]. If, moreover, all the weights $w_n = \frac{1}{N}$, the (unique) minimizer of $f$ is called the Fréchet mean of the points $a_1, \ldots, a_N$.

Both medians and means of points in an Hadamard space are currently a subject of intensive research for their applications in computational biology; see [3, 5, 22] and the references therein. Another area where Fréchet means play an important role are so-called consensus algorithms [10].

Having demonstrated the importance of functions of the form (3), we will now turn to a nonlinear version of the Lie-Trotter-Kato product formula, which states that the gradient flow semigroup of $f$ can be approximated by the resolvents or semigroups of the individual functions $f_j$, with $j = 1, \ldots, k$. We first need some notation.

**Notation 1.3.** Let $f : \mathcal{H} \to (-\infty, \infty]$ be a function of the form (3), and we of course assume that it is not identically equal $\infty$. Its resolvent $J_\lambda$ and semigroup $S_t$ are given by (1) and (2), respectively. The resolvent of the function $f_j$ is denoted $J_{\lambda_j}$, for each $j = 1, \ldots, k$, and likewise the semigroup of $f_j$ is denoted $S_{t_j}$, for each $j = 1, \ldots, k$.

The symbol $P_j$ will denote the metric projection onto $\text{dom } f_j$, where $j = 1, \ldots, k$. If $F : \mathcal{H} \to \mathcal{H}$ is a mapping, we denote its $k$th power, with $k \in \mathbb{N}$, by

$$F^{(k)} x := (F \circ \cdots \circ F) x, \quad x \in \mathcal{H},$$
where $F$ appears $k$-times on the right hand side. Having this notation at hand, we are able to state a nonlinear version of the Lie-Trotter-Kato product formula due to Stojkovic [24, Theorems 4.4, 4.5, 4.8].

**Theorem 1.4** (Stojkovic). Let $(\mathcal{H}, d)$ be an Hadamard space and $f : \mathcal{H} \to (-\infty, \infty]$ be of the form (3). We use Notation 1.3. Then, for any $t \geq 0$, and $x \in \text{dom } f$, we have

\[
S_t(x) = \lim_{n \to \infty} \left( J_{\frac{t}{n}} \circ \cdots \circ J_{\frac{1}{n}} \right)^{(n)}(x),
\]

and,

\[
S_t(x) = \lim_{n \to \infty} \left( S_{\frac{k}{n}} \circ P_k \cdots \circ S_{\frac{1}{n}} \circ P_1 \right)^{(n)}(x).
\]

The convergence in both (4a) and (4b) is uniform with respect to $t$ on any compact subinterval of $[0, \infty)$.

The original proof in [24] uses ultrapowers of Hadamard spaces or additional assumptions (local compactness of the underlying space). The aim of the present note is to simplify this proof by using the notion of weak convergence. Such simplification is believed to be desirable with regard to the importance of the Lie-Trotter-Kato formula.

We shall finish this Introduction by recalling a brief development of the theory of gradient flows in Hadamard spaces, which has recently attracted considerable interest. It started independently by the work of Jost [15] and Mayer [21], when the existence of the gradient flow semigroup was established. The study of the relationship with the Mosco and $\Gamma$- convergences, initiated already in [15], was treated in greater detail in [4, 18]. In [6], the author describes large time behavior of the gradient flow as well as its discrete version called the proximal point algorithm. As already mentioned above, the the Lie-Trotter-Kato product formula was proved in [24]. There have been also many related results in some special instances of Hadamard spaces, namely, in manifolds of nonpositive sectional curvature ([19, 20, 23]), and the Hilbert ball ([16] and the references therein). On the other hand this theory can be partially extended into more general metric spaces and plays an important role in optimal transport theory, PDEs, and probability [1]. For another viewpoint; see [2].

To demonstrate that the gradient flow theory in Hadamard spaces applies in various situations, we now present a number of natural examples of convex lsc functions on an Hadamard space $(\mathcal{H}, d)$.

**Example 1.5** (Indicator functions). Let $K \subset \mathcal{H}$ be a convex set. Define the indicator function of $K$ by

\[
\iota_K(x) := \begin{cases} 
0, & \text{if } x \in K, \\
\infty, & \text{if } x \notin K.
\end{cases}
\]

Then $\iota_K$ is a convex function, and it is lsc if and only if $K$ is closed.

**Example 1.6** (Distance functions). The function

\[
x \mapsto d(x, x_0), \quad x \in \mathcal{H},
\]

where $x_0$ is a fixed point of $\mathcal{H}$, is convex and continuous. The function $d(\cdot, z)^p$ for $p > 1$ is strictly convex. More generally, the distance function to a closed convex
subset $C \subset \mathcal{H}$, defined as
\[
d_C(x) := \inf_{c \in C} d(x,c), \quad x \in \mathcal{H},
\]
is convex and 1-Lipschitz [8, p.178].

**Example 1.7** (Displacement functions). Let $T : \mathcal{H} \to \mathcal{H}$ be an isometry. The displacement function of $T$ is the function $\delta_T : \mathcal{H} \to [0, \infty)$ defined by
\[
\delta_T(x) := d(x,Tx), \quad x \in \mathcal{H}.
\]
It is convex and Lipschitz [8, p.229].

**Example 1.8** (Busemann functions). Let $c : [0, \infty) \to \mathcal{H}$ be a geodesic ray. The function $b_c : \mathcal{H} \to \mathbb{R}$ defined by
\[
b_c(x) := \lim_{t \to \infty} [d(x,c(t)) - t], \quad x \in \mathcal{H},
\]
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 [8, p. 273]. Another explicit example of a Busemann function in the Hadamard space of positive definite $n \times n$ matrices with real entries can be found in [8, 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 [8] and the references therein.

**Example 1.9** (Energy functional). The energy functional is another important instance of a convex function on an Hadamard space [11, 12, 13, 17]. Indeed, the energy functional is convex and lsc on a suitable Hadamard space of $L^2$-mappings. Minimizers of the energy functional are called harmonic maps, and are of an immense importance in both geometry and analysis. For a probabilistic approach to harmonic maps, see [25, 26, 28].

2. Preliminaries

We first recall basic notation and facts concerning Hadamard spaces. For further details on the subject, the reader is referred to [8].

2.1. Hadamard spaces. If a geodesic metric space $(X,d)$ satisfies the following inequality
\[
d(x,\gamma(t))^2 \leq (1-t)d(x,\gamma(0))^2 + td(x,\gamma(1))^2 - t(1-t)d(\gamma(0),\gamma(1))^2,
\]
for any $x \in X$, any geodesic $\gamma : [0,1] \to X$, and any $t \in [0,1]$, we say it has nonpositive curvature (in the sense of Alexandrov), or that it is a CAT(0) space. A complete CAT(0) space is called an Hadamard space.

The class of Hadamard spaces includes Hilbert spaces, $\mathbb{R}$-trees, Euclidean Bruhat-Tits buildings, classical hyperbolic spaces, complete simply connected Riemannian manifolds of nonpositive sectional curvature, the Hilbert ball, CAT(0) complexes, and many other important spaces included in none of the above classes [8].

Let $(\mathcal{H},d)$ be an Hadamard space. Having two points $x,y \in \mathcal{H}$, we denote the geodesic segment from $x$ to $y$ by $[x,y]$. We usually do not distinguish between a geodesic and its geodesic segment, as no confusion can arise. For a point $z \in [x,y]$, we write $z = (1-t)x + ty$, where $t = d(x,z)/d(x,y)$.

For a function $f : \mathcal{H} \to (-\infty, \infty]$ we denote $\text{dom } f = \{x \in \mathcal{H} : f(x) < \infty \}$. If $\text{dom } f \neq \emptyset$, we say $f$ is proper. To avoid trivial situations we often assume this
property without explicit mentioning. As usually, the symbol \( \text{dom} f \) stands for the closure of \( \text{dom} f \). A point \( x \in \mathcal{H} \) is called a minimizer of \( f \) if \( f(x) = \inf_{\mathcal{H}} f \).

2.2. Convex sets and functions on Hadamard spaces. Recall that a set \( C \subseteq \mathcal{H} \) is convex if \( x, y \in C \) implies \( [x, y] \subseteq C \). A function \( f : \mathcal{H} \to (-\infty, \infty] \) is convex provided \( f \circ \gamma : [0, 1] \to (-\infty, \infty] \) is convex for any geodesic \( \gamma : [0, 1] \to \mathcal{H} \). Note that the distance function \( d_C \) is convex and continuous; see Example 1.6.

**Proposition 2.1.** Let \((\mathcal{H}, d)\) be an Hadamard space and \( C \subseteq \mathcal{H} \) be closed and convex. Then:

(i) For every \( x \in \mathcal{H} \), there exists a unique point \( P_C(x) \in C \) such that
\[
d(x, P_C(x)) = d_C(x).
\]

(ii) If \( x \in \mathcal{H} \) and \( y \in C \), then
\[
d(x, y)^2 \geq d(x, P_C(x))^2 + d(y, P_C(x))^2
\]

(iii) The mapping \( P_C : \mathcal{H} \to C \) is nonexpansive and is called the metric projection onto \( C \).

**Proof.** See [8, Proposition 2.4, p.176]. □

The following result comes from [9, Lemma 2.2]. We include its short proof for readers’ convenience.

**Proposition 2.2.** Let \((\mathcal{H}, d)\) be an Hadamard space. If \((C_\alpha)_{\alpha \in I}\) is a nonincreasing family of bounded closed convex sets in \( \mathcal{H} \). Then \( \bigcap_{\alpha \in I} C_\alpha \neq \emptyset \).

**Proof.** Choose \( x \in \mathcal{H} \) and denote its projection onto \( C_\alpha \) by \( x_\alpha = P_{C_\alpha}(x) \). Then \((d(x, x_\alpha))_\alpha\) is a nonincreasing net of nonnegative numbers, and hence has a limit \( l \). If \( l = 0 \), then \( x \in \bigcap_\alpha C_\alpha \). If \( l > 0 \), then we claim that \((x_\alpha)\) is Cauchy. Indeed, denote \( x_{\alpha \beta} = \frac{1}{2}x_\alpha + \frac{1}{2}x_\beta \) and apply (6) with \( t = 1/2 \) to obtain
\[
d(x, x_{\alpha \beta})^2 \leq \frac{1}{2}d(x, x_\alpha)^2 + \frac{1}{2}d(x, x_\beta)^2 - \frac{1}{4}d(x_\alpha, x_\beta)^2,
\]
which implies that \((x_\alpha)\) is Cauchy. The limit point clearly lies in \( \bigcap_\alpha C_\alpha \). □

As a consequence of Proposition 2.2 we obtain that convex lsc functions are locally bounded.

**Lemma 2.3.** Let \((\mathcal{H}, d)\) be an Hadamard, and \( f : \mathcal{H} \to (-\infty, \infty] \) be a lsc convex function. Then \( f \) is bounded from below on bounded sets.

**Proof.** Let \( C \subseteq \mathcal{H} \) be bounded, and without loss of generality assume that \( C \) is closed convex. If \( \inf_C f = -\infty \), then the sets \( S_N = \{x \in C : f(x) \leq -N\} \) for \( N \in \mathbb{N} \) are all nonempty. Since all \( S_N \) are closed convex and bounded, Proposition 2.2 yields a point \( z \in \bigcap_{N \in \mathbb{N}} S_N \). Clearly \( f(z) = -\infty \), which is not possible. □

Let \((\mathcal{H}, d)\) be an Hadamard space, \( f : \mathcal{H} \to (-\infty, \infty] \) be convex lsc and \( x \in \mathcal{H} \). Then
\[
\frac{1}{2\lambda}d(J_\lambda x, v)^2 - \frac{1}{2\lambda}d(x, v)^2 + \frac{1}{2\lambda}d(J_\lambda x)^2 + f(J_\lambda x) \leq f(v),
\]
for each \( v \in \text{dom} f \). See [1, Theorem 4.1.2]. Furthermore, if \( x \in \text{dom} f \) and we set \( x(t) = S_t x \), for \( t \in [0, \infty) \), then
\[
\frac{1}{2t}d(x(t), v)^2 - \frac{1}{2t}d(x(0), v)^2 + f(x(t)) \leq f(v),
\]
for any $t > 0$ and $v \in \text{dom } f$. See [1, (4.0.13)].

2.3. Weak convergence in Hadamard spaces. Here we recall the definition and basic properties of the weak convergence in Hadamard spaces. For a systematic account, the reader is referred to [4, Section 3].

We shall say that a bounded sequence $(x_n) \subset \mathcal{H}$ weakly converges to a point $x \in \mathcal{H}$ if $P_{x_n} \to x$ as $n \to \infty$ for any geodesic $\gamma : [0,1] \to \mathcal{H}$ with $\gamma(0) = x$.

If there is a subsequence $(x_{n_p})$ of $(x_n)$ such that $x_{n_p} \rightharpoonup z$ for some $z \in \mathcal{H}$, we say that $z$ is a weak cluster point of the sequence $(x_n)$. The following important result first appeared in [11, Theorem 2.1].

Proposition 2.4. Each bounded sequence has a weakly convergent subsequence, or in other words, each bounded sequence has a weak cluster point.

Proof. See [11, Theorem 2.1] or [4, Proposition 3.2].

Lemma 2.5. Let $C \subset \mathcal{H}$ a closed convex set. If $(x_n) \subset C$ and $x_n \rightharpoonup x \in \mathcal{H}$, then $x \in C$.

Proof. See [4, Lemma 3.7].

Definition 2.6. We shall say that a function $f : \mathcal{H} \to (-\infty, \infty]$ is weakly lsc at a given point $x \in \text{dom } f$ if

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

for each sequence $x_n \rightharpoonup x$. We say that $f$ is weakly lsc if it is lsc at any $x \in \text{dom } f$.

Lemma 2.7. If $f : \mathcal{H} \to (-\infty, \infty]$ a lsc convex function, then it is weakly lsc.

Proof. See [4, Lemma 3.9].

2.4. Nonexpansive mappings in Hadamard spaces.

Definition 2.8. Let $(\mathcal{H},d)$ be an Hadamard space and $(F_\rho)_{\rho>0}$ be a family of nonexpansive maps $F_\rho : \mathcal{H} \to \mathcal{H}$. Fix $\lambda, \rho > 0$ and $x \in \mathcal{H}$. Then the map

(10)

$$y \mapsto \frac{1}{1+\lambda \rho}x + \frac{\lambda}{1+\lambda \rho}F_\rho y, \quad y \in \mathcal{H},$$

is a contraction with Lipschitz constant $\frac{\lambda}{1+\lambda \rho}$, and hence has a unique fixed point, which will be denoted $J_{\lambda,\rho}(x)$, see [24, Lemma 3.1].

Theorem 2.9. Let $(\mathcal{H},d)$ be an Hadamard space and $f : \mathcal{H} \to (-\infty, \infty]$ be of the form (3). We use Notation 1.3.

(i) Let $J_{\lambda,\rho}$ be the mapping from Definition 2.8 associated with the nonexpansive map $F_\rho = (J_\rho^k \circ \cdots \circ J_\rho^1)$. If we have

$$J_\lambda(x) = \lim_{\rho \to 0^+} J_{\lambda,\rho}(x),$$

for any $x \in \text{dom } f$ and $\lambda > 0$, then,

$$S_t(x) = \lim_{n \to \infty} \left( J_{\frac{t}{n}}^k \circ \cdots \circ J_{\frac{1}{n}}^1 \right)^{(n)}(x),$$

for any $x \in \text{dom } f$ and $t > 0$, and the convergence is uniform with respect to $t$ on any compact subinterval of $[0,\infty)$.
(ii) Let now $J_{\lambda,\rho}$ be the mapping from Definition 2.8 associated with the non-expansive map $F_\rho = (S^k_\rho \circ P_k \cdot \cdot \cdot \circ S^1_\rho \circ P_1)$. If we have

$$J_\lambda(x) = \lim_{\rho \to 0^+} J_{\lambda,\rho}(x),$$

for any $x \in \text{dom } f$ and $\lambda > 0$, then,

$$S_t(x) = \lim_{n \to \infty} \left( S^k_{\frac{t}{n}} \circ P_k \cdot \cdot \cdot \circ S^1_{\frac{t}{n}} \circ P_1 \right)^{(n)}(x),$$

for any $x \in \text{dom } f$ and $t > 0$, and the convergence is uniform with respect to $t$ on any compact subinterval of $[0, \infty)$.

Proof. See [24, Theorem 3.13].

3. Proof of the main result

In this section, we give the promised alternative proof of Theorem 1.4. For reader’s convenience, we recall the statement here.

Theorem 3.1 (Stojkovic). Let $(H, d)$ be an Hadamard space and $f : H \to (-\infty, \infty]$ be of the form (3). We use Notation 1.3. Then, for any $t \geq 0$, and $x \in \text{dom } f$, we have

$$S_t(x) = \lim_{n \to \infty} \left( J^k_{\frac{t}{n}} \circ \cdot \cdot \cdot \circ J^1_{\frac{t}{n}} \right)^{(n)}(x),$$

and,

$$S_t(x) = \lim_{n \to \infty} \left( S^k_{\frac{t}{n}} \circ P_k \cdot \cdot \cdot \circ S^1_{\frac{t}{n}} \circ P_1 \right)^{(n)}(x).$$

The convergence in both (13a) and (13b) is uniform with respect to $t$ on any compact subinterval of $[0, \infty)$.

Proof. We mix various inequalities derived in [24] and employ the weak convergence as appropriate.

We first show (13a). By Theorem 2.9, it suffices to show $J_{\lambda,t}x \to J_\lambda x$ as $t \to 0$, where $J_{\lambda,\rho}$ now corresponds to the choice $F_\rho = (J^k_{\rho} \circ \cdot \cdot \cdot \circ J^1_{\rho})$. Put $x_0(t) = J_{\lambda,t}x$ and $x_j(t) = J^j_{\frac{t}{n}}x_{j-1}(t)$ for $j = 1, \ldots, k$. By the definition of $J_{\lambda,t}$ we have

$$x_0(t) = \frac{1}{1 + \lambda/t} x + \frac{\lambda/t}{1 + \lambda/t} x_1(t),$$

and consequently,

$$d(x, x_k(t)) = \frac{t + \lambda}{\lambda} d(x, x_0(t)),$$

together with,

$$d(x_0(t), x_k(t)) = \frac{t}{\lambda} d(x, x_0(t)).$$

Applying (8) for each $f_j$, with $j = 1, \ldots, k$, and summing the resulting inequalities up, we arrive at

$$2t f(v) \geq 2t \sum_{j=1}^k f_j(x_j(t)) + d(x_k(t), v)^2 - d(x_0(t), v)^2 + \sum_{j=1}^k d(x_{j-1}(t), x_j(t))^2,$$
for every \( v \in \text{dom } f \). The inequality (6) yields
\[
\frac{\lambda}{1 + \lambda/t} d(v, x_k(t))^2 \geq d(v, x_0(t))^2 - \frac{1}{1 + \lambda/t} d(v, x)^2 + \frac{\lambda}{(1 + \lambda/t)^2} d(x, x_k(t))^2.
\]
(18)

Combining this inequality with (15) and (17) gives after some elementary calculations that
\[
2\lambda f(v) \geq 2\lambda \sum_{j=1}^{k} f_j(x_j(t)) + d(x_0(t), x)^2 + d(x_0(t), v)^2 - d(x, v)^2.
\]
(19) for every \( v \in \text{dom } f \).

Fix now a sequence \( t_n \to 0 \). For every \( j = 0, 1, \ldots, k \), the sequence \( (x_j(t_n))_n \) is bounded due to [24, Proposition 4.2]. Then Lemma 2.3 yields that the sequence \( (f_j(x_j(t_n)))_n \) is bounded from below and the inequality (19) implies that it is also bounded from above.

Consider (17) with \( t_n \) and take the limit \( n \to \infty \) to obtain via (16) that
\[
\lim_{n \to \infty} \sum_{j=1}^{k} d(x_{j-1}(t_n), x_j(t_n))^2 = 0.
\]
(20)

Let \( z \in \mathcal{H} \) be a weak cluster point of \( x_0(t_n) \) and \( x_0(t_{n_p}) \) be a sequence weakly converging to \( z \). Recall that the existence of a weak cluster point was guaranteed by Proposition 2.4. By (20), also \( x_j(t_{n_p}) \) weakly converges to \( z \), for all \( j = 1, \ldots, k \). Consequently, \( z \in \text{dom } f \) due to Lemma 2.5.

Consider next the inequality (19) with \( t_{n_p} \) and take the limit \( p \to \infty \). Since the functions \( f_j \), with \( j = 1, \ldots, k \), are weakly \( \mathcal{L}_c \) by Lemma 2.7, we obtain \( z = J_\lambda x \).

In particular, \( z \in \text{dom } f \). Since \( z \) was an arbitrary weak cluster point of \( x_0(t_n) \), we get \( x_0(t_n) \rightharpoonup J_\lambda x \). Applying once again (19) with \( v = J_\lambda x \), gives \( x_0(t_n) \to J_\lambda x \) as \( n \to \infty \). This finishes the proof of (13a).

Next we show (13b). By Theorem 2.9, it suffices to show \( J_{\lambda, p} x \to J_\lambda x \) as \( t \to 0 \), where \( J_{\lambda, p} \) now corresponds to the choice \( F_p = (S^{k}_{\rho} \circ P_k \cdots \circ S^{1}_{\rho} \circ P_1) \).

Put \( x_0(t) = J_{\lambda, p} x \) and \( x_j(t) = (S^{j}_{\rho} \circ P_j) x_{j-1}(t) \) for \( j = 1, \ldots, N \). Since we again have (14), the same arguments as above yield that (15) and also (18) hold true.

Applying (9) with \( f_j \), for \( j = 1, \ldots, N \), and summing the resulting inequalities up gives
\[
f(v) \geq \frac{1}{2t} d(x_k(t), v)^2 - \frac{1}{2t} d(x_0(t), v)^2 + \sum_{j=1}^{k} f_j(x_j(t)),
\]
for any \( v \in \text{dom } f \). Combining this inequality with (15) and (18) gives after some elementary calculations that (19) holds.

Fix now a sequence \( t_n \to 0 \). For every \( j = 0, 1, \ldots, k \), the sequence \( (x_j(t_n))_n \) is bounded due to [24, Proposition 4.2]. Then Lemma 2.3 yields that the sequence \( (f_j(x_j(t_n)))_n \) is bounded from below and the inequality (19) implies that it is also bounded from above.

The inequality (9) implies that
\[
d(x_j(t_n), v)^2 - d(P_j x_{j-1}(t_n), v)^2 + 2t_n f_j(x_j(t_n)) \leq 2t_n f_j(v),
\]
for each $j = 1, \ldots, k$ and $v \in \text{dom} f_j$. Take the limit $n \to \infty$ to obtain
\begin{equation}
\limsup_{n \to \infty} \left[ d(x_j(t_n), v)^2 - d(P_j x_{j-1}(t_n), v)^2 \right] \leq 0,
\end{equation}
for each $j = 1, \ldots, k$ and $v \in \text{dom} f_j$, and even $v \in \overline{\text{dom}} f_j$. In combination with the fact that
\[ d(x_j(t_n), v)^2 - d(x_0(t_n), v)^2 \to 0, \quad \text{as } n \to \infty, \]
for each $j = 1, \ldots, k$ and $v \in \text{dom} f$, contained in \cite[(4.20)]{24}, we obtain
\begin{equation}
\limsup_{n \to \infty} \left[ d(x_{j-1}(t_n), v)^2 - d(P_j x_{j-1}(t_n), v)^2 \right] \leq 0,
\end{equation}
for each $j = 1, \ldots, k$ and $v \in \overline{\text{dom}} f$. The inequality (7) now reads
\[ d(P_j x_{j-1}(t_n), v)^2 + d(x_{j-1}(t_n), P_j x_{j-1}(t_n))^2 \leq d(x_{j-1}(t_n), v)^2, \]
for any $v \in \overline{\text{dom}} f_j$. Therefore
\begin{equation}
\limsup_{n \to \infty} d(x_{j-1}(t_n), P_j x_{j-1}(t_n))^2 \leq 0,
\end{equation}
for any $j = 1, \ldots, k$.

Let $z \in \mathcal{H}$ be a weak cluster point $x_0(t_n)$ and $x_0(t_{n_p})$ be a sequence weakly converging to $z$. Again, the existence of a weak cluster point was guaranteed by Proposition 2.4. By (23) we have $z \in \overline{\text{dom}} f_j$. Apply now the inequality (22) with $v = P_j x_{j-1}(t_n)$ and (23) to obtain $x_1(t_{n_p}) \overset{w}{\to} z$. By the same argument then also $z \in \overline{\text{dom}} f$ and $x_j(t_{n_p})$ weakly converges to $z$, for $j = 2, \ldots, k$. Consider next the inequality (19) with $v = J_{\lambda}x$ and $t = t_{n_p}$ and take the limit $p \to \infty$. Since the functions $f_j$, with $j = 1, \ldots, k$, are weakly lsc by Lemma 2.7, we obtain $z = J_{\lambda}x$. In particular, $z \in \text{dom } f$. Since $z$ was an arbitrary weak cluster point of $x_0(t_n)$, we get $x_0(t_n) \overset{w}{\to} J_{\lambda}x$. Applying once again (19) with $v = J_{\lambda}x$ and $t = t_n$ gives $x_0(t_n) \to J_{\lambda}x$ as $n \to \infty$. This finishes the proof of (13b). \hfill \Box

\section*{References}
\begin{enumerate}
\item M. Băcăk, \textit{Computing medians and means via the proximal point algorithm}, Preprint.
\end{enumerate}