Canad. Math. Bull. Vol. 59 (4), 2016 pp. 673–681
http://dx.doi.org/10.4153/CMB-2016-033-3

Lipschitz Retractions in Hadamard Spaces via Gradient Flow Semigroups

Miroslav Bačák and Leonid V. Kovalev

Abstract. Let $X(n)$, for $n \in \mathbb{N}$, be the set of all subsets of a metric space (X, d) of cardinality at most n. The set $X(n)$ equipped with the Hausdorff metric is called a finite subset space. In this paper we are concerned with the existence of Lipschitz retractions $r: X(n) \to X(n-1)$ for $n \geq 2$. It is known that such retractions do not exist if X is the one-dimensional sphere. On the other hand, Kovalev has recently established their existence if X is a Hilbert space, and he also posed a question as to whether or not such Lipschitz retractions exist when X is a Hadamard space. In this paper we answer the question in the positive.

1 Introduction

Let (X, d) be a metric space. For each $n \in \mathbb{N}$, we denote the set of all subsets of X with cardinality at most n by $X(n)$. The set $X(n)$ equipped with the Hausdorff metric d_{H} is called a *finite subset space.* Unlike Cartesian powers X^n or the space of unordered n-tuples X^n/S_n, finite subset spaces admit *canonical* isometric embeddings

$$\iota: X(n-1) \to X(n).$$

Following [6, 7, 9], we are interested in Lipschitz retractions $r: X(n) \to X(n-1)$. Kovalev proved their existence for X being a Euclidean space [6] and X being a Hilbert space [7]. On the other hand, a result of Mostovoy [9] yields that in general there is no continuous mapping $r: X(n) \to X(n-1)$ with $r \circ \iota = \mathrm{id}$ if X is the one-dimensional sphere $\mathbb{S}^1$; here id stands for the identity operator on $X(n-1)$. It is therefore natural to ask whether Lipschitz retractions $r: X(n) \to X(n-1)$ exist if X is a nonpositively curved metric space. Indeed, this question appears explicitly in [7, Question 3.3]. It was originally motivated by Remark 4.5 in [6], which observes that the existence of Lipschitz retractions between finite subset spaces $X(n)$ enables induction on n in certain extension arguments. Indeed, under appropriate assumptions, a Lipschitz map into $X(n)$ can be decomposed as a pair of maps into $X(n-1)$; after applying an inductive hypothesis one ends up with a map into $X(2n-2)$, at which point a retraction onto $X(n)$ is applied to obtain an extended Lipschitz map into $X(n)$. Our main result (Theorem 3.2) provides the positive solution to [7, Question 3.3].

The solution for Hilbert spaces from [7] is based on the existence of gradient flow trajectories in a finite dimensional subspace, which is assured by the classical ODE

Received by the editors March 1, 2016; revised May 9, 2016.
Published electronically June 28, 2016.
Author L. V. K. was supported by the National Science Foundation grant DMS-1362453.
AMS subject classification: **53C23**, 47H20, 54E40, 58D07.
Keywords: finite subset space, gradient flow, Hadamard space, Lie–Trotter–Kato formula, Lipschitz retraction.

theory. In the present paper, we also define the desired retractions via gradient flows of certain convex functionals on (the n-th power of) a Hadamard space and make use of the Lie–Trotter–Kato formula proved recently in [3, 11].

Finally, note that given a Hadamard space $(\mathcal{H}, d)$, we obtain Lipschitz retractions $r: \mathcal{H}(n) \to \mathcal{H}(n-1)$ with Lipschitz constant $\max(4n^{\frac{3}{2}}+1, 2n^2+n^{\frac{1}{2}})$, whereas the result from [7] for X being a Hilbert space has Lipschitz constant $\max(2n-1, n^{\frac{3}{2}})$.

2 Preliminaries

We first recall some basic facts about Hadamard spaces as well as more recent results that shall be used in our proof. For further details, we refer the reader to [2].

Let $(\mathcal{H}, d)$ be a Hadamard space, that is, a complete metric space with geodesics satisfying

$$d(z, x_t)^2 \le (1-t)d(z, x_0)^2 + td(z, x_1)^2 - t(1-t)d(x_0, x_1)^2, \tag{2.1}$$

for each $z, x_0, x_1 \in \mathcal{H}$ and $t \in [0,1]$, where $x_t := (1-t)x_0 + tx_1$ is a unique point on the geodesic $[x_0, x_1]$ such that $d(x_0, x_t) = td(x_0, x_1)$. An equivalent (and more geometric) formulation of inequality (2.1) is the following relation between a triangle with vertices $p, q, r \in \mathcal{H}$ and its *comparison triangle* with vertices $\overline{p}, \overline{q}, \overline{r} \in \mathbb{R}^2$, where $d(p,q) = \|\overline{p}-\overline{q}\|$, $d(r,q) = \|\overline{r}-\overline{q}\|$, and $d(p,r) = \|\overline{p}-\overline{r}\|$. If $x := (1-t)p + tq$ and $y := (1-s)p + sr$ for some $s, t \in [0,1]$, and we denote their comparison points by $\overline{x} := (1-t)\overline{p} + t\overline{q}$ and $\overline{y} := (1-s)\overline{p} + s\overline{r}$, respectively, inequality (2.1) implies that

$$d(x, y) \le \|\overline{x} - \overline{y}\|. \tag{2.2}$$

Here the symbol $\|\cdot\|$ stands for the Euclidean norm on $\mathbb{R}^2$.

Given two geodesics $[x_0, x_1]$ and $[y_0, y_1]$, we have

$$d(x_t, y_t) \le (1-t)d(x_0, y_0) + td(x_1, y_1), \tag{2.3}$$

for each $t \in [0,1]$; see [2, (1.2.4)].

Given a function $f: \mathcal{H} \to (-\infty, \infty]$, denote its *domain* by

$$\operatorname{dom} f := \{x \in \mathcal{H} : f(x) < \infty\}$$

and the set of its minimizers by $\operatorname{Min} f := \{x \in \mathcal{H} : f(x) = \inf f\}$. We say that a function $f: \mathcal{H} \to (-\infty, \infty]$ is *convex* if for each geodesic $\gamma: [0,1] \to \mathcal{H}$, the function $f \circ \gamma$ is convex. Given a convex lower semicontinuous (lsc, for short) function $f: \mathcal{H} \to (-\infty, \infty]$, point $x \in \mathcal{H}$, and $\lambda > 0$, there exists a unique minimizer of the function

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

which we denote by $J_\lambda x$. The mapping $J_\lambda: x \mapsto J_\lambda x$ is called the *resolvent* of f with parameter λ. It satisfies the following important inequality

$$f(J_\lambda x) + \frac{1}{2\lambda} d(x, J_\lambda x)^2 + \frac{1}{2\lambda} d(J_\lambda x, y)^2 \le f(y) + \frac{1}{2\lambda} d(x,y)^2, \tag{2.4}$$

for every $x, y \in \mathcal{H}$. Given $x \in \overline{\operatorname{dom}} f$, the *gradient flow semigroup* associated to f is defined by

$$S_t x := \lim_{k\to\infty} (J_{\frac{t}{k}})^k x, \quad x \in \overline{\operatorname{dom}} f, \tag{2.5}$$

https://doi.org/10.4153/CMB-2016-033-3 Published online by Cambridge University Press

for every $t \in [0,\infty)$. As in Hilbert spaces, the semigroup is comprised of nonexpansive operators, that is

$$d(S_t x, S_t y) \le d(x,y), \tag{2.6}$$

for each $t \in [0,\infty)$ and $x, y \in \overline{\operatorname{dom} f}$. The above-mentioned theory of gradient flows in Hadamard spaces was first studied by Jost [5] and Mayer [8]. A more recent result established the asymptotic behavior of a gradient flow [1]. It relies upon the notion of weak convergence in Hadamard spaces, which was introduced by Jost [4]. Let us recall that a bounded sequence $(x_k) \subset \mathcal{H}$ converges *weakly* to a point $x \in \mathcal{H}$ provided $\lim_{k\to\infty} d\left(P_\gamma(x_k), x\right) = 0$ for each geodesic $\gamma\colon [0,1] \to \mathcal{H}$ with $x \in \gamma$. Here P_γ stands for the metric projection onto (the image of) γ. Now we are ready to state the theorem on asymptotic behavior of a gradient flow.

Theorem 2.1 *Let $f\colon \mathcal{H} \to (-\infty,\infty]$ be a convex lsc function which attains its infimum on $\mathcal{H}$. Then, given $x \in \overline{\operatorname{dom} f}$, the associated gradient flow semigroup $S_t x$ weakly converges to a point $x^* \in \operatorname{Min} f$.*

Proof See [1] or [2, Theorem 5.1.16]. ∎

The function values then converge to the infimum of f.

Theorem 2.2 *Let $f\colon \mathcal{H} \to (-\infty,\infty]$ be a convex lsc function which attains its infimum on $\mathcal{H}$. Then, given $x \in \overline{\operatorname{dom} f}$, we have $f(S_t x) \to \inf f$ as $t \to \infty$.*

Proof This can be seen from the proof of Theorem 2.1 in [1] or, for instance, as in [2, Proposition 5.1.12]. ∎

The proof of our main theorem uses gradient flows of convex functions to define a desired Lipschitz retraction. These convex functions have a special form, namely they are given as a finite sum of some elementary convex functions, and their gradient flow can be approximated by the Lie–Trotter–Kato formula. We will now state the necessary facts precisely. Let $N \in \mathbb{N}$ and consider a function $f\colon \mathcal{H} \to (-\infty,\infty]$ of the form

$$f := \sum_{j=1}^{N} f_j, \tag{2.7}$$

where $f_j\colon \mathcal{H} \to (-\infty,\infty]$ are convex lsc functions for every $j = 1,\dots,N$. Let us denote the resolvent of the function f_j by $J_\lambda^{[j]}$ and the gradient flow semigroup of f by S_t.

Theorem 2.3 (Lie–Trotter–Kato formula) *Let $f\colon \mathcal{H} \to (-\infty,\infty]$ be of the form (2.7). Then we have*

$$S_t x = \lim_{k\to\infty} \left(J_{\frac{t}{k}}^{[N]} \circ \cdots \circ J_{\frac{t}{k}}^{[1]} \right)^k x, \tag{2.8}$$

for every $t \in [0,\infty)$ and $x \in \overline{\operatorname{dom} f}$.

Proof The original proof appeared in [11]. For a simplified proof, see [3]. ∎

https://doi.org/10.4153/CMB-2016-033-3 Published online by Cambridge University Press

In fact, the gradient flow we are going to use in the proof of Theorem 3.2 is not on $\mathcal{H}$, but on its n-th power. The n-th power of $\mathcal{H}$, denoted by $\mathcal{H}^n$, is equipped with the metric

$$d(x,y) := \Big(\sum_{j=1}^{n} d\left(x_j, y_j\right)^2 \Big)^{\frac{1}{2}}, \quad x, y \in \mathcal{H}^n,$$

and is then also a Hadamard space. Note that we use the same symbol d for the original metric on $\mathcal{H}$, as well as for the metric on $\mathcal{H}^n$. It is always clear from the arguments which one is meant. Given a point $x = (x_1, \ldots, x_n) \in \mathcal{H}^n$, we shall denote $\{x\} := \{x_1, \ldots, x_n\}$, a subset of $\mathcal{H}$. However, given $x, y \in \mathcal{H}^n$, we shall denote the Hausdorff distance between $\{x\}$ and $\{y\}$ by $d_{\mathrm{H}}(x,y)$ instead of $d_{\mathrm{H}}(\{x\},\{y\})$.

3 The Existence of Lipschitz Retractions

The desired Lipschitz retractions $r: \mathcal{H}(n) \to \mathcal{H}(n-1)$ will be defined via a gradient flow of a convex functional on $\mathcal{H}^n$. Specifically, we define this functional as

$$F(x) := \sum_{1 \le i < j \le n} d(x_i, x_j), \quad x = (x_1, \ldots, x_n) \in \mathcal{H}^n. \tag{3.1}$$

and show that it is indeed convex and Lipschitz.

Lemma 3.1 *The function $F: \mathcal{H}^n \to \mathbb{R}$ is convex and $n^{\frac{3}{2}}$-Lipschitz.*

Proof Convexity follows from (2.3). For the Lipschitz property, we estimate

$$\begin{aligned} |F(x) - F(y)| &\le \sum_{1 \le i < j \le n} |d(x_i, x_j) - d(y_i, y_j)| \\ &\le \sum_{1 \le i < j \le n} d\left(x_i, y_i\right) + d(x_j, y_j) \le n^{\frac{3}{2}} d(x,y), \end{aligned}$$

where we twice used the triangle inequality and then the Cauchy–Scharz inequality. ∎

We are now ready to prove the main theorem.

Theorem 3.2 *Let $(\mathcal{H}, d)$ be a Hadamard space. Then for each integer $n \ge 2$ there exists a Lipschitz retraction $r: \mathcal{H}(n) \to \mathcal{H}(n-1)$ with Lipschitz constant*

$$\max(4n^{\frac{3}{2}} + 1, 2n^2 + n^{\frac{1}{2}}).$$

Proof We divide the proof into several steps.

Step 1. Let J_λ and S_t be the resolvent and gradient flow semigroup, respectively, associated with the function F from (3.1). Let us denote

$$D := \left\{ x = (x_1, \ldots, x_n) \in \mathcal{H}^n : x_i = x_j \text{ for some } 1 \le i < j \le n \right\}.$$

Given $x \in \mathcal{H}^n$, we define

$$\delta(x) := \min_{1 \le i < j \le n} d(x_i, x_j), \quad \text{and} \quad T(x) := \inf\left\{ t > 0 : S_t x \in D \right\}.$$

https://doi.org/10.4153/CMB-2016-033-3 Published online by Cambridge University Press

We will first show that

$$T(x) \le \frac{1}{2}\delta(x). \tag{3.2}$$

In order to be able to apply formula (2.8), denote by $J_\lambda^{[i,j]}$ the resolvent associated with the function $x \mapsto d(x_i, x_j)$, $x = (x_1, \dots, x_n) \in \mathcal{H}^n$, where $1 \le i < j \le n$. One can easily verify that for the l-th coordinate ($l = 1, \dots, n$) we have

$$(J_\lambda^{[i,j]}x)_l = \begin{cases} (1-\alpha)x_i + \alpha x_j & \text{if } l = i, \\ (1-\alpha)x_j + \alpha x_i & \text{if } l = j, \\ x_l & \text{if } l \notin \{i,j\}, \end{cases}$$

where $\alpha = \min(\frac{1}{2}, \frac{\lambda}{d(x_i,x_j)})$. Indeed, this value of α minimizes the function

$$d((1-\alpha)x_i+\alpha x_j, (1-\alpha)x_j+\alpha x_i)+\frac{1}{2\lambda}\Big(d((1-\alpha)x_i+\alpha x_j, x_i)^2+d((1-\alpha)x_j+\alpha x_i, x_j)^2\Big),$$

because for $\alpha \le 1/2$ this function simplifies to

$$\frac{d(x_i,x_j)}{\lambda}(\lambda - 2\lambda\alpha + \alpha^2 d(x_i,x_j)).$$

Then (2.8) reads

$$S_t x = \lim_{k\to\infty} (R_{\frac{t}{k}})^k x, \quad x \in \overline{\operatorname{dom}} f, \tag{3.3}$$

where

$$\begin{aligned} R_{\frac{t}{k}} &:= J_{\frac{t}{k}}^{[n-1,n]} \circ \cdots \circ J_{\frac{t}{k}}^{[1,6]} \\ &\circ J_{\frac{t}{k}}^{[4,5]} \circ J_{\frac{t}{k}}^{[3,5]} \circ J_{\frac{t}{k}}^{[2,5]} \circ J_{\frac{t}{k}}^{[1,5]} \circ J_{\frac{t}{k}}^{[3,4]} \circ J_{\frac{t}{k}}^{[2,4]} \circ J_{\frac{t}{k}}^{[1,4]} \circ J_{\frac{t}{k}}^{[2,3]} \circ J_{\frac{t}{k}}^{[1,3]} \circ J_{\frac{t}{k}}^{[1,2]}. \end{aligned}$$

Next we turn our attention to the first two coordinates. This choice will be justified by (3.4). We want to show that applying the mapping $J_\lambda^{[2,i]} J_\lambda^{[1,i]}$ can extend the first two coordinates of a given point only if they were less than λ apart, and this extension is at most λ. More precisely, we claim that for $\lambda > 0$ and $i = 3, \dots, n$, the following holds for every $y = (y_1, \dots, y_n) \in \mathcal{H}^n$ and $z := J_\lambda^{[2,i]} J_\lambda^{[1,i]} y$:

(i) if $d(y_1, y_2) \ge \lambda$, then $d(z_1, z_2) \le d(y_1, y_2)$.

(ii) if $d(y_1, y_2) < \lambda$, then $d(z_1, z_2) \le d(y_1, y_2) + \lambda$.

We will now show both (i) and (ii) by using comparison triangles. To this end denote $u := J_\lambda^{[1,i]} y$ and consider the triangle with vertices y_1, y_2, u_i in $\mathcal{H}$ along with its comparison triangle with vertices $\overline{y_1}, \overline{y_2}, \overline{u_i} \in \mathbb{R}^2$. Then denote the comparison points of z_1 and z_2 by $\overline{z_1}$ and $\overline{z_2}$, respectively. Next observe that (i) and (ii) hold true if we replace all the points involved by their comparison points (and consider the Euclidean distance in $\mathbb{R}^2$, of course). This can be seen by elementary geometry arguments about triangles in $\mathbb{R}^2$, presented below this proof as Lemma 3.3 (where $A = y_1$, $B = y_2$, $C = u_i$). Finally, applying (2.2) gives (i) and (ii).

Choose $x \in \mathcal{H}^n$ and $k \in \mathbb{N}$. Denote $\lambda := \frac{\delta(x)}{2k}$. Without loss of generality one may assume

$$d(x_1, x_2) = \delta(x). \tag{3.4}$$

https://doi.org/10.4153/CMB-2016-033-3 Published online by Cambridge University Press

Define now $x^{l,[i,j]} := J_\lambda^{[i,j]} \circ \cdots \circ J_\lambda^{[2,3]} \circ J_\lambda^{[1,3]} \circ J_\lambda^{[1,2]} \circ (R_\lambda)^{l-1} x$, for each $l = 1, \dots, k$ and $1 \le i < j \le n$, and observe that (i) and (ii) imply

$$d\Big(x_1^{k,[1,n]}, x_2^{k,[2,n]} \Big) \le \lambda = \frac{\delta(x)}{2k}, \tag{3.5}$$

where the subscript indices denote the coordinates in $\mathcal{H}^n$. Indeed, each application of $J_\lambda^{[1,2]}$ shortens the distance between the first two coordinates by additive constant 2λ, while the application of $J_\lambda^{[2,i]} J_\lambda^{[1,i]}$, with $i = 3, \dots, n$, does not expand it, or expands it by the additive constant λ at most, as we know from (i) and (ii). More precisely, we have

$$\begin{aligned}
d\left(x_1, x_2\right) &= \delta(x),\\
d\Big(x_1^{1,[1,2]}, x_2^{1,[1,2]} \Big) &= \max\Big(0, \delta(x) - \frac{\delta(x)}{k} \Big),\\
d\Big(x_1^{1,[n-1,n]}, x_2^{1,[n-1,n]} \Big) &\le \max\Big(\frac{\delta(x)}{k}, \delta(x) - \frac{\delta(x)}{k} \Big),\\
d\Big(x_1^{2,[1,2]}, x_2^{2,[1,2]} \Big) &\le \max\Big(0, \delta(x) - 2\frac{\delta(x)}{k} \Big),\\
d\Big(x_1^{2,[n-1,n]}, x_2^{2,[n-1,n]} \Big) &\le \max\Big(\frac{\delta(x)}{k}, \delta(x) - 2\frac{\delta(x)}{k} \Big),\\
&\vdots\\
d\Big(x_1^{k-1,[n-1,n]}, x_2^{k-1,[n-1,n]} \Big) &\le \max\Big(\frac{\delta(x)}{k}, \delta(x) - (k-1)\frac{\delta(x)}{k} \Big) = \frac{\delta(x)}{k},\\
d\Big(x_1^{k,[1,2]}, x_2^{k,[1,2]} \Big) &= 0.
\end{aligned}$$

and hence (3.5) holds true.

Passing to the limit $k \to \infty$ in (3.5) and recalling (3.3), then give

$$d\Big((S_{\frac12 \delta(x)} x)_1, (S_{\frac12 \delta(x)} x)_2 \Big) = 0,$$

or, in other words, we have just proved (3.2).

Step 2. Let $x, y \in \mathcal{H}^n$. By (2.4) we have

$$\frac{1}{2\lambda} d\left(J_\lambda x, y\right)^2 \le F(y) - F\left(J_\lambda x\right) + \frac{1}{2\lambda} d(x,y)^2.$$

Consider now $t \in [0, T(x)]$. Fix $k \in \mathbb{N}$ and employ the above inequality k times to obtain

$$\begin{aligned}
d(J_{\frac{t}{k}} x, y)^2 &\le \frac{2t}{k}\Big[F(y) - F(J_{\frac{t}{k}} x) \Big] + d(x,y)^2,\\
d\Big((J_{\frac{t}{k}})^2 x, y \Big)^2 &\le \frac{2t}{k}\Big[F(y) - F\Big((J_{\frac{t}{k}})^2 x \Big) \Big] + d(J_{\frac{t}{k}} x, y)^2,\\
&\vdots\\
d\Big((J_{\frac{t}{k}})^k x, y \Big)^2 &\le \frac{2t}{k}\Big[F(y) - F\Big((J_{\frac{t}{k}})^k x \Big) \Big] + d\Big((J_{\frac{t}{k}})^{k-1} x, y \Big)^2.
\end{aligned}$$

Summing up these inequalities, dividing by t^2, and putting $x := y$ gives

$$\frac{d\big((J_{\frac{t}{k}})^k x, x\big)^2}{t^2} \le 2\frac{F(x) - F\big((J_{\frac{t}{k}})^k x\big)}{t} \le 2n^{\frac{3}{2}} \frac{d\big(x, (J_{\frac{t}{k}})^k x\big)}{t},$$

and after taking $\limsup_{k\to\infty}$, we obtain $\frac{d(S_t x, x)}{t} \le 2n^{\frac{3}{2}}$. Hence, by virtue of (3.2),

$$d_{\mathrm{H}}\left(S_t x, x\right) \le d\left(S_t x, x\right) \le 2tn^{\frac{3}{2}} \le \delta(x) n^{\frac{3}{2}}. \tag{3.6}$$

For future reference we also record that the nonexpansiveness of the gradient flow semigroup (2.6) implies

$$d_{\mathrm{H}}(S_t x, S_t y) \le d(S_t x, S_t y) \le d(x, y) \le n^{\frac{1}{2}} \max_{1\le j\le n} d(x_j, y_j). \tag{3.7}$$

Step 3. Given $x \in \mathcal{H}(n)$, we number its elements $\{x_1, \dots, x_n\}$ and consider $x' := (x_1, \dots, x_n) \in \mathcal{H}^n$. We may assume that $d\left(x_1, x_2\right) = \delta\left(x'\right)$. Then we define $r(x) := \left\{S_{T(x')} x'\right\}$. However, we will write x instead of x' in the sequel. Let us now show that $r: \mathcal{H}(n) \to \mathcal{H}(n-1)$ is a Lipschitz retraction. First of all, observe that r is the identity on (the canonical embedding of) $\mathcal{H}(n-1)$. To prove the Lipschitz property, choose $x, y \in \mathcal{H}(n)$ and examine the following two alternatives. If $\delta(x) + \delta(y) \le 4d_{\mathrm{H}}(x, y)$, then

$$\begin{aligned} d_{\mathrm{H}}\left(r(x), r(y)\right) &\le d_{\mathrm{H}}\left(r(x), x\right) + d_{\mathrm{H}}\left(x, y\right) + d_{\mathrm{H}}\left(y, r(y)\right) \\ &\le n^{\frac{3}{2}}\delta(x) + d_{\mathrm{H}}(x, y) + n^{\frac{3}{2}}\delta(y) \le (4n^{\frac{3}{2}} + 1) d_{\mathrm{H}}(x, y) \end{aligned}$$

where we used (3.6) to obtain the second inequality.

If, on the other hand, $\delta(x) + \delta(y) > 4d_{\mathrm{H}}(x, y)$, then we may assume $\delta(x) > 2d_{\mathrm{H}}(x, y)$ without loss of generality. The fact that $\delta(x) > 2d_{\mathrm{H}}(x, y)$ then implies that we can renumber the points $\{y_1, \dots, y_n\}$ in such a way that

$$d\left(x_j, y_j\right) \le d_{\mathrm{H}}(x, y) \tag{3.8}$$

for each $j = 1, \dots, n$. In the remainder of the proof, we will use (3.8) only, without referring to $\delta(x) > 2d_{\mathrm{H}}(x, y)$. We can, hence, without loss of generality assume $T(x) \le T(y)$ on account of the fact that the roles of x and y in (3.8) are interchangeable. Recall that $r(x) = S_{T(x)} x$ and put $z := S_{T(x)} y$. Inequality (3.7) implies that $d_{\mathrm{H}}\left(r(x), z\right) \le n^{\frac{1}{2}} d_{\mathrm{H}}(x, y)$. Consequently, $\delta(z) = \delta(z) - \delta\left(r(x)\right) \le 2d_{\mathrm{H}}\left(z, r(x)\right) \le 2n^{\frac{1}{2}} d_{\mathrm{H}}(x, y)$. By (3.6) we have $d_{\mathrm{H}}\left(z, r(z)\right) \le n^{\frac{3}{2}}\delta(z) \le 2n^2 d_{\mathrm{H}}(x, y)$. Finally, one arrives at

$$\begin{aligned} d_{\mathrm{H}}\left(r(x), r(y)\right) &\le d_{\mathrm{H}}\left(r(x), z\right) + d_{\mathrm{H}}\left(z, r(z)\right) + d_{\mathrm{H}}\left(r(z), r(y)\right) \\ &\le n^{\frac{1}{2}} d_{\mathrm{H}}(x, y) + 2n^2 d_{\mathrm{H}}(x, y) + 0, \end{aligned}$$

where the zero on the right-hand side is due to the semigroup property of the gradient flow. The proof is complete. ∎

The following lemma provides justification for inequalities (i) and (ii) used in the preceding proof.

Lemma 3.3 *Let ABC be a triangle in $\mathbb{R}^2$. Fix $\lambda > 0$. Let A_1, C_1 be two points on the segment AC such that $|AA_1| = |CC_1| = \min(|AC|/2, \lambda)$. Also let B_1 be the point on BC_1 such that $|BB_1| = \min(|BC_1|/2, \lambda)$. Then*

$$|A_1B_1| \le \begin{cases} |AB| & \text{if } |AB| \ge \lambda, \\ |AB| + \lambda & \text{if } |AB| < \lambda. \end{cases}$$

Proof The inequality $|A_1B_1| \le |AB| + \lambda$ holds because A_1, B_1 are obtained by translating the points A, B toward C_1 by amounts between 0 and λ. It remains to prove that $|A_1B_1| \le |AB|$ when $|AB| \ge \lambda$. There are three cases to consider.

Case 1: $|AC| \ge 2\lambda$ and $|BC_1| \ge 2\lambda$. Then both A_1 and B_1 are obtained from A and B by translating them toward C_1 by the same distance λ. Hence $|A_1B_1| < |AB|$.

Case 2: $|AC| < 2\lambda$ and $|BC_1| \ge 2\lambda$. Then $A_1 = C_1$, hence $|A_1B_1| = |A_1B| - \lambda$. And since $|A_1B| \le |AB| + \lambda$, we have $|A_1B_1| \le |AB|$ again.

Case 3: $|BC_1| < 2\lambda$. Then $|A_1B_1| = \frac{1}{2}|A_1B| \le \frac{1}{2}(|AB| + \lambda) \le |AB|$. ∎

Remark 3.4 (Asymptotic behavior of the flow) Given $x \in \mathcal{H}^n$, denote $\Delta(x) := \max_{1\le i<j\le n} d(x_i, x_j)$. Using the same arguments as in the previous proof, we can show that for $\tau := \frac{1}{2}\Delta(x)$, one obtains $S_\tau x \in \operatorname{Min} F$. Alternatively, we can obtain the asymptotic behavior of the flow as follows. By Theorem 2.1, given $x \in \mathcal{H}^n$, the flow $S_t x$ weakly converges to a point $x^* = (x_1^*, \dots, x_n^*) \in \operatorname{Min} F$; and obviously $x_1^* = \cdots = x_n^*$. Next we show that this convergence is in fact strong. To this end, we first observe that $x_1^* \in \bigcap_{t\in[0,\infty)} \overline{\mathrm{co}}\,\{S_t x\}$. Indeed, by virtue of (2.5) and the semigroup property it is sufficient to show that $\{J_\lambda x\} \subset \overline{\mathrm{co}}\{x\}$. This inclusion however follows directly by a projection argument. Now use Theorem 2.2 to conclude $F(S_t x) \to 0$ and therefore $\operatorname{diam} \overline{\mathrm{co}}\,\{S_t x\} \to 0$. Hence we have $S_t x \to x^*$.

Remark 3.5 (Open questions) We end the paper by posing a few questions, many of which have appeared already in [7]. The Lipschitz constant of $r\colon \mathcal{H}(n) \to \mathcal{H}(n-1)$ guaranteed by Theorem 3.2 is $\max(4n^{\frac{3}{2}} + 1, 2n^2 + n^{\frac{1}{2}})$. Can one improve upon this constant? Can one show that, for every $n \in \mathbb{N}$ with $n \ge 2$, there exist Lipschitz retractions $r\colon \mathcal{H}(n) \to \mathcal{H}(n-1)$ with Lipschitz constants independent of n? Can one extend Theorem 3.2 into spaces of nonpositive curvature in the sense of Busemann? In particular, does an analog of Theorem 3.2 hold in strictly convex or uniformly convex Banach spaces? Can one extend Theorem 3.2 into p-uniformly convex spaces? Recall that a geodesic metric space (X, d) is called p-uniformly convex (for $p \ge 2$) if there exists $K > 0$ such that

$$d(z, x_t)^p \le (1-t)d(z, x_0)^p + td(z, x_1)^p - Kt(1-t)d(x_0, x_1)^p$$

for each $z, x_0, x_1 \in \mathcal{H}$ and $t \in [0,1]$, where $x_t := (1-t)x_0 + tx_1$. This definition was introduced in [10, Definition 3.2] as a generalization of p-uniform convexity in Banach spaces.

Acknowledgement We would like to thank the referee for carefully reading the preprint and suggesting several improvements on the exposition.

References

[1] M. Bačák, *The proximal point algorithm in metric spaces*, Israel J. Math. **194**(2013), no. 2, 689–701. http://dx.doi.org/10.1007/s11856-012-0091-3

[2] ———, *Convex analysis and optimization in Hadamard spaces*. De Gruyter Series in Nonlinear Analysis and Applications, 22. De Gruyter, Berlin, 2014. http://dx.doi.org/10.1515/9783110361629

[3] ———, *A new proof of the Lie–Trotter–Kato formula in Hadamard spaces*. Commun. Contemp. Math. **16**(2014), p. 1350044 (15 pages). http://dx.doi.org/10.1142/S0219199713500442

[4] J. Jost, *Equilibrium maps between metric spaces*. Calc. Var. Partial Differential Equations **2**(1994), 173–204. http://dx.doi.org/10.1007/BF01191341

[5] ———, *Nonlinear Dirichlet forms*. In: New directions in Dirichlet forms, AMS/IP Stud. Adv. Math. 8. Amer. Math. Soc., Providence, RI, 1998, pp. 1–47.

[6] L. V. Kovalev, *Symmetric products of the line: embeddings and retractions*. Proc. Amer. Math. Soc. **143**(2015), 801–809. http://dx.doi.org/10.1090/S0002-9939-2014-12280-5

[7] ———, *Lipschitz retraction of finite subsets of Hilbert spaces*. Bull. Aust. Math. Soc. **93**(2016), 146–151. http://dx.doi.org/10.1017/S0004972715000672

[8] U. F. Mayer, *Gradient flows on nonpositively curved metric spaces and harmonic maps*. Comm. Anal. Geom. **6**(1998), 199–253. http://dx.doi.org/10.4310/CAG.1998.v6.n2.a1

[9] J. Mostovoy, *Lattices in* $\mathbb{C}$ *and finite subsets of a circle*. Amer. Math. Monthly **111**(2004), 357–360. http://dx.doi.org/10.2307/4145248

[10] A. Naor and L. Silberman, *Poincaré inequalities, embeddings, and wild groups*. Compos. Math. **147**(2011), 1546–1572. http://dx.doi.org/10.1112/S0010437X11005343

[11] I. Stojkovic, *Approximation for convex functionals on non-positively curved spaces and the Trotter-Kato product formula*. Adv. Calc. Var. **5**(2012), 77–126.

Max Planck Institute for Mathematics in the Sciences, Inselstr. 22, 04 103 Leipzig, Germany
e-mail: bacak@mis.mpg.de

Department of Mathematics, Syracuse University, Syracuse, NY 13244, USA
e-mail: lvkovale@syr.edu