Approximation for convex functionals on metric spaces of non-positive curvature

Igor Stojkovic
Mathematical Institute
Leiden University
Niels Bohrweg 1, 2333 CA Leiden
Netherlands

2008

Abstract

We cosider two geodesicly convex functional on a complete lenght space which is non-positivly curved. In analogy to the linear setting we buid Josida like approximation resolvents associated to these two functionals and the corresponding semigroups: yet only with the distance function and geodesics at hand. Next we show that convergence of these Josida resolvents impies the convergence of the semigroups. Finaly we show that the Trotter product formula holds under suitable assumptions. The main difficuly is the lack of linear structure, and finding the nonlinear countepart of the linear theory in terms of distance and geodesics.
Keywords: Semigroups, Convexity, Approximation, Trotter Product Formula
Subject Classification: 28A33, 28A50, 35K55, 35K90, 47H05, 47J35,49J40, 65M15

1 Introduction

During the past few decades much of the classical theory of linear spaces haas been extended to a mertic space setting of one or another kind. For example we have the beautiful theory of Variational Analysis which is a very powerfull way to treat Partial Differential Equations. Several other disciplines of matematics have been brought to life as mathematicians worked on these problems. What one ussually does is define the resolvent operators my means of minimization, and this scheme converges to a limit which defines a semigroup of solutions of the PDE in question. One can take a more abbstract point of wiew on this, that is instead of starting with a PDE that one wants to be solve, just concider a convex lower semicontinuous functional, and by the same procedure construct a semigroup which than automaticly has a certain properties. Already during the 1980's Di Giorgi and his school have been trying to buit such theory in a purely metric setting, i.e. without assumptions such as that the space is a convex subset of a Banach or a Hilbert space. They have introduced some fundamental concepts such as the metric slope and curves of maximal slope, which has opened the way to develop the theory of Gradient Flows on Metric spaces. A true breakthrough has came at the end of 1990's with the seminal papers [3] and [4] which show that the Fokker-Planck and Porous Medium equations can be interpreted as a Gradient Flow on the Wasserstein-2 space of probability measures on $\mathbb{R}^d$. Around the same time U.Mayer in his work [1] has shown by a Crandal-Liggett like approach (see [5]) that on a complete metric space which is nonpositivly curved (see Definition 1.) and such that each two points can be connected by a constant speed geodesic (i.e. a lenght space), one can deifine resolvents associated to such functional and the exponential formula defines a strongly continuous semigroup. Said differently we have a situation analogous to that of fuctionals on a Hilbert space. A few years later the authors of [2] have constructed a sound mathematical theory of Gradient flows on metric spaces.

Next to the above described in certain situations one can consider several functionals and their sum. A natural question which has very usefull applications, is can we "split" the scheme. To be precise can we approximate the flow associated to the sum functional, by taking the smaal steps in direction of the flow of these functionals separately? This problem in the linear setting has been treated in various papers. Let us mention two of these works. The first one is the work of Brezis and Pazy [0] which discusses aproximation of (sum)semigroups on Banach spaces induceed by accreative operators. The second one is the work of Kato and masuda [6] which show that with the aid of the results of [0] one has that the Trotter product formula holds for an arbitrary finite sequence of convex lower semicontinuous functionals on a Hilbert space.

The aim of this paper is to extend these results to the setting of complete lenghth spaces which are nonpositivly curved. More precisely we consider a complete lenght space X which is nonpositivly curved and two geodesicly convex lower semicontinuous functionals with values in $[0, \infty)$. In the first part of our investigation we buid the resolvents of the approximated semigroups and their smigroups. These should be seen as an approximation of the semigroup induceed by the sum of the two functionals under consideration. In the second part we show that under suitable, yet not too strong conditions that the Trotter Product formula holds in our case. In particular it always holds if the closed balls in X are compact. In particular if X is a Riemanian manifold with nonpositive sectional curvature.

Definition Non positively curved metric space is a complete metric space (NPC) s.t.

1. For any $u, v \in X$ there is a rectifiable curve from u to v with length equal to $d(u, v)$ i.e. X is a length space.

2. For any $v, u_0, u_1 \in X$ and any constant speed geodesic $u : [0, 1] \to X$ $u(0) = u_0$, $u(1) = u_1$

$$d^2(v, u_t) \leqslant (1-t)d^2(v, u_0) + td^2(v, u_1) - t(1-t)d^2(u_0, u_1) \qquad (1)$$

holds.

Theorem 1.1 *If (X,d) is a NPC space than for any $u,v \in X$ there is a unique constant speed geodesic connecting u and v, with length $d(u,v)$ ($[0,1] \to X$).*

Proof proof Mayer Thm 1.1.a)

Lemma 1.2 *For three points x,y,z in a NPC space X and any $t \in (0,1)$*

$$d(y_t, z_t) \leqslant t d(y,z) \tag{2}$$

where y_t, z_t denote the unique constant speed geodesic $[0,1] \to X$ from x to y and from x to z respectively.

Proof First of all by Def. 0.0 b)

$$d^2(y_t, z_t) \leqslant (1-t)d^2(y_t, x) + t d^2(y_t, z) - (1-t)d^2(x,z)$$

$$d^(z, y_t) \leqslant (1-t)d^2(z,x) + t d^2(z,y) - t(1-t)d^2(x,y)$$

and $d(y_t, x) = t d(x,y)$. Inserting the second and the third information into the first one we obtain

$$d^2(y_t,z_t) \leqslant t^2(1-t)d(x,y) + t(1-t)d^2(z,x) + t^2 d^2(z,y) - t^2(1-t)d^2(x,y) - t(1-t)d^2(x,z) = t^2 d^2(z,y)$$

Lemma 1.3 *Let X be a NPC space and let $x : [0,1] \to X$ be the unique geodesic connecting x_0 and x_1, and let $y \in X$ be any point. Than for $t \in [0,1]$*

$$d(x_t, y) \leqslant (1-t)x_0 + t d(x_1, y)$$

Proof Pick a $t \in [0,1]$ and points $\bar{x}_0, \bar{x}_1, \bar{y} \in \mathbb{R}^2$ such that $d(x_0,y) = |\bar{x}_0 - \bar{y}|$, $d(x_1, y) = |x_1 - y|$, $d(x_0, x_1) = |\bar{x}_0 - \bar{x}_1|$ where $|\cdot|$ denotes the Euclidian norm on $\mathbb{R}^2$. Such three points exist by the triangle inequality on X. Set $\bar{x}_t := (1-t)\bar{x}_0 + t\bar{x}_1$ and notice that by 1.1 $d^2(x_t, y) \leqslant |\bar{x}_t - \bar{y}|$ as 1.1 holds with equality if the space under consideration is a Hilbert space. Now as

$$|\bar{x}_t - \bar{y}| \leqslant (1-t)|\bar{x}_0 - \bar{y}| + t|\bar{x}_1 - \bar{y}| = (1-t)d(x_0,y) + td)x_1, y)$$

we are done. ∎

2 Basic Approximation Theorems

Let $\{\mathbb{F}_\rho | \rho > 0\}$ be a family of contactions $X \to X$. In order to prove the Trotter product formula we want to define "resolvents" $\left(I + \frac{\lambda}{\rho}(I - F_\rho)\right)^{-1}$ for $\lambda, \rho > 0$.

In case X is a Banach space if

$$\left(I + \frac{\lambda}{\rho}(I - F_\rho)\right)^{-1} x = y$$
$$\Leftrightarrow x = y + \frac{\lambda}{\rho} y - \frac{\lambda}{\rho} F_\rho y$$
$$y = \frac{1}{1+\frac{\lambda}{\rho}} x + \frac{\frac{\lambda}{\rho}}{1+\frac{\lambda}{\rho}} F_\rho y$$

Following the lines of Breses & Pazy Lemma 2.2.i) we define for $x \in X$, $\lambda, \rho > 0$ the mapping $G_{\lambda,\rho,x} : X \to X$ by

$$G_{\lambda,\rho,x}(y) := \frac{1}{1+\frac{\lambda}{\rho}} x + \frac{\frac{\lambda}{\rho}}{1+\frac{\lambda}{\rho}} F_\rho y = x + \frac{\frac{\lambda}{\rho}}{1+\frac{\lambda}{\rho}} (F_\rho y - x)$$

where the last two expressions stand for the point $u\left(\frac{\frac{\lambda}{\rho}}{1+\frac{\lambda}{\rho}}\right)$ on the unique constant speed geodesic $u : [0,1] \to X$ with $u(0) = x_1$, $u(1) = F_\rho y$. We will write G when there is no ambiguity about which $\lambda, \rho > 0$ and $x \in X$ we refer to.

Lemma 2.1 *For any $\lambda, \rho > 0$ $x \in X$ the mapping $G_{\lambda,\rho,x}$ is Lipshitz with constant not longer than $\frac{\frac{\lambda}{\rho}}{1+\frac{\lambda}{\rho}} < 1$. Hence by the fixed point argument there is a unique y, s.t. $G_{\lambda,\rho,x}(y) = y$.*

Proof By Lemma 1.2 the definition of $G = G_{\lambda,\rho,x}$ and the fact that F_ρ is contraction

$$d(G(y_1), G(y_2)) \leqslant \frac{\frac{\lambda}{\rho}}{1+\frac{\lambda}{\rho}} d(F_\rho y_1, F_\rho y_2) \leqslant \frac{\frac{\lambda}{\rho}}{1+\frac{\lambda}{\rho}} d(y_1, y_2).$$

∎

Definition For $\lambda, \rho > 0$ $x \in X$ we will denote the unique fixed point of $G_{\lambda,\rho,x}$ by $\mathcal{I}_{\lambda,\rho} x$. We show that $\mathcal{I}_{\lambda,\rho}$ is a contraction.

Lemma 2.2 *For any $\lambda, \rho > 0$ $\mathcal{I}_{\lambda,\rho}$ is a contraction.*

Proof We write $\mathcal{I}$ for $\mathcal{I}_{\lambda,\rho}$ and F for F_ρ for the sake of simplicity. Inductively it is clear why this statement holds as X is nonpositively curved and F is a contraction. We estimate by the NPC property

$$d^2(\mathcal{I}_{x_1}, \mathcal{I}_{x_2}) \leqslant (1-t)d^2(\mathcal{I}_{x_1}, x_2) + td^2(\mathcal{I}_{x_1}, F\mathcal{I}_{x_2}) - t(1-t)d^2(x_2, F\mathcal{I}_{x_2})$$
$$d^2(\mathcal{I}_{x_1}, x_2) \leqslant (1-t)d^2(x_2, x_1) + td^2(x_2, F\mathcal{I}_{x_1}) - t(1-t)d^2(x_1, F\mathcal{I}_{x_1})$$
$$d^2(\mathcal{I}_{x_1}, F\mathcal{I}_{x_2}) \leqslant (1-t)d^2(x_1, F\mathcal{I}_{x_2}) + td^2(F\mathcal{I}_{x_1}, F\mathcal{I}_{x_2}) - t(1-t)d^2(x_1, F\mathcal{I}_{x_2})$$

Combining these inequalities and using that F is a contraction we obtain

$$\begin{aligned} d^2(\mathcal{I}x_1, \mathcal{I}x_2) \leqslant & (1-t)^2 d^2(x_1,x_2) + t(1-t)d^2(x_2, F\mathcal{I}x_1) - t(1-t)^2 d^2(x_1, F\mathcal{I}x_1) \\ & + t(1-t)d^2(x_1, F\mathcal{I}x_2) + t^2 d^2(F\mathcal{I}x_1, F\mathcal{I}x_2) - t^2(1-t)d^2(x_1, F\mathcal{I}x_1) \\ & - t(1-t)d^2(x_2, F\mathcal{I}x_2) \end{aligned}$$

By Thm 1.1. b) Mayer choosing u_0, u_1, v_0, v_1 to be $x_1, F\mathcal{I}x_1, x_2, F\mathcal{I}x_2$ resp. and $t = 1$ we have

$$d^2(F\mathcal{I}x_1, x_2) + d^2(x_1, F\mathcal{I}x_2) \leqslant d^2(x_1,x_2) + d^2(F\mathcal{I}x_1, F\mathcal{I}x_2) + 2d(x_1, F\mathcal{I}x_1)d(x_2, F\mathcal{I}x_2)$$

and inserting this into the previous inequality, together with the fact that F is a contraction we obtain

$$\begin{aligned} d^2(\mathcal{I}x_1, \mathcal{I}x_2) \leqslant & (1-t)^2 d^2(x_1,x_2) + t^2 d^2(F\mathcal{I}x_1, F\mathcal{I}x_2) + \\ & + t(1-t))d^2(x_1,x_2) + d^2(F\mathcal{I}x_1, F\mathcal{I}x_2)) \\ & + 2t(1-t)d(x_1, F\mathcal{I}x_1)d(x_2, F\mathcal{I}x_2) \\ & - t(1-t)(d^2(x_1, F\mathcal{I}x_1) + d^2(x_2, F\mathcal{I}x_2)) \\ & \leqslant (1-t)d^2(x_1,x_2) + td^2(\mathcal{I}x_1, \mathcal{I}x_2) - t(1-t)(d(x_1, F\mathcal{I}x_1) - d(x_2, F\mathcal{I}x_2)) \\ & \leqslant (1-t)d^2(x_1,x_2) + td^2(\mathcal{I}x_1, \mathcal{I}x_2), \end{aligned}$$

hence $d(\mathcal{I}x_1, \mathcal{I}x_2) \leqslant d(x_1, x_2)$ ∎

Lemma 2.3 *For $\lambda, \rho > 0$, $x \in X$ $\frac{d(x, \mathcal{I}_{\lambda,\rho}x)}{\lambda} \leqslant \frac{d(x, F_\rho x)}{\rho}$.*

Proof As $\mathcal{I}_{\lambda,\rho}x$ is the unique fixed point of our strictly contraction $G_{\lambda,\rho,x}$ for any $z \in Z$ the sequence $\{G^n_{\lambda,\rho,x}(x)\}_{n=1}^\infty$ converges to $\mathcal{I}_{\lambda,\rho}x$. In particular choosing $z = x$ yields

$$\lim_n G^n_{\lambda,\rho,x}(x) = \mathcal{I}_{\lambda,\rho}x$$

Hence as $G_{\lambda,\rho,x}$ is a contraction with const. $\frac{\frac{\lambda}{\rho}}{1+\frac{\lambda}{\rho}}$ it is clear that

$$\begin{aligned} d(x, \mathcal{I}_{\lambda,\rho}x) & \leqslant \sum_{n=1}^{\infty} d(G^{n-1}_{\lambda,\rho,x}(x), G^n_{\lambda,\rho,x}(x)) \\ & \leqslant \sum_{n=1}^{\infty} \left(\frac{\frac{\lambda}{\rho}}{1+\frac{\lambda}{\rho}} \right)^{n-1} d(x, G_{\lambda,\rho,x}(x)) \\ & = \frac{1}{1 - \frac{\frac{\lambda}{\rho}}{1+\frac{\lambda}{\rho}}} d(x, F_\rho x) = \frac{\frac{\lambda}{\rho}}{1+\frac{\lambda}{\rho}} d(x, F_\rho x) = \frac{\lambda}{\rho} d(x, F_\rho x) \end{aligned}$$

∎

Next we show the Resolvent identity. If $x \in X$ $0 < \mu < \lambda$; $\rho > 0$ let y be the point $u\left(\frac{\lambda-\mu}{\lambda}\right)$ where $u : [0,1] \to X$ is the unique constant speed geodesic from x to $\mathcal{I}_{\lambda,\rho}x$.

Lemma 2.4 *The Resolvent identity $\mathcal{I}_{\lambda,\rho}x = \mathcal{I}_{\mu,\rho}(y)$ holds.*

Proof As by construction $\mathcal{I}_{\mu,\rho}y$ is the unique point z in X s.t. $z = v\left(\frac{\frac{\mu}{\rho}}{1+\frac{\mu}{\rho}}\right)$ where $v : [0,1] \to X$ is the unique constant speed geodesic joining y and $F_\rho z$, it is enough to show that

$$d(y, \mathcal{I}_{\lambda,\rho}x) = \frac{\frac{\mu}{\rho}}{1+\frac{\mu}{\rho}} d(y, F_\rho \mathcal{I}_{\lambda,\rho}) = \frac{\mu}{\rho+\mu} d(y, F_\rho \mathcal{I}_{\lambda,\rho})$$

Set

$$a := d(x, F_\rho \mathcal{I}_{\lambda,\rho}x)$$
$$b := d(\mathcal{I}_{\lambda,\rho}x, F_\rho \mathcal{I}_{\lambda,\rho}x) = \frac{1}{1+\frac{\lambda}{\rho}} a = \frac{\rho}{\rho+\lambda} a$$
$$c := d(y, \mathcal{I}_{\lambda,\rho}x) = (a-b)\frac{\mu}{\lambda} = \frac{\lambda}{\rho+\lambda} \cdot \frac{\mu}{\lambda} a = \frac{\mu}{\rho+\lambda} a$$

Now

$$\frac{d(y, \mathcal{I}_{\lambda,\rho}x)}{d(y, \mathcal{I}_{\lambda,\rho}x)} = \frac{c}{b+c} = \frac{\frac{\mu}{\rho+\lambda}a}{\frac{\rho+\mu}{\rho+\lambda}a} = \frac{\mu}{\rho+\mu}$$

∎

Corollary 2.5 *For $\lambda, \rho > 0$ $x \in X$, $n \in \mathbb{N}$*

$$d(\mathcal{I}^n_{\lambda,\rho}x, x) \leqslant n\frac{\lambda}{\rho} d(x, F_\rho x)$$

Proof Follows by triangle inequality as $\mathcal{I}_{\lambda,\rho}$ is a contraction.

In order to understand the ... notice the analogy of .... Lemma 2.2 with previous four results above.

We have all the ingredients to execute eventually the same proof as Chandall & Liggef (Thm 1) which in our case yields existence of semigroup generated by "operators $\frac{F_\rho - I}{\rho}$". We start with two preposatory lemmas.

Lemma 2.6 *Let $\rho > 0$, $0 < \mu \leqslant \lambda$, $n, m \in \mathbb{N}$, $n \geqslant m$. Than writing $\mathcal{I}_\mu = \mathcal{I}_{\mu\rho}$ $\mathcal{I}_\lambda = \mathcal{I}_{\lambda,\rho}$ we have*

$$d(\mathcal{I}^n_\mu x, \mathcal{I}^m_\lambda x) \leqslant \sum_{j=1}^{m-1} \alpha^j \beta^{n-j} B(n,j) d(\mathcal{I}^{m-j}_\lambda x, x) + \sum_{j=m}^{n} \alpha^m \beta^{j-m} B(j-1, m-1) d(\mathcal{I}^{n-j}_\mu x, x)$$

where $\alpha = \frac{\mu}{\lambda}$, $\beta = \frac{\lambda-\mu}{\lambda}$ and $B(\cdot,\cdot)$ is a binomial coefficient.

Proof For integers j and k with $0 \leqslant j \leqslant n$, $0 \leqslant k \leqslant m$ put

$$a_{k,j} := d(\mathcal{I}_\mu^j x, \mathcal{I}_\lambda^k x).$$

By Lemma 2.4 and Lemma 2.2 and Lemma 1.3 ($u : [0,1] \to X$ being the unique geodesic joining $\mathcal{I}_\lambda^{k-1}$ and $\mathcal{I}_\lambda^k x$) we have

$$\begin{aligned} a_{k,j} &\leqslant d\left(\mathcal{I}_\mu^j x, \mathcal{I}_\mu\left(u\left(\frac{\lambda-\mu}{\lambda}\right)\right)\right) \leqslant \\ &\leqslant d\left(\mathcal{I}_\mu^{j-1} x, u\left(\frac{\lambda-\mu}{\mu}\right)\right) \\ &\leqslant \frac{\mu}{\lambda} d(\mathcal{I}_\mu^{j-1}x, \mathcal{I}_\lambda^{k-1}x) + \frac{\lambda-\mu}{\lambda} d(\mathcal{I}_\mu^{j-1}x, \mathcal{I}_\lambda^k x) \\ &\leqslant \alpha a_{k-1} j - 1 + \beta a_{k,j-1} \end{aligned}$$

These inequalities can be solved to estimate $a_{m,n}$ as in terms of $a_{k,0}$ and $a_{0,j}$ $0 \leqslant j \leqslant n$, $0 \leqslant k \leqslant m$ as claimed.

When $n \leqslant m$ the estimate becomes

$$d(\mathcal{I}_\mu^n, \mathcal{I}_\lambda^m x) \leqslant \sum_{j=0}^n \alpha^j \beta^{n-j} B(n,j) d(\mathcal{I}_\lambda^{m-j} x, x).$$

The proof goes by induction and left to the reader. ∎

Lemma 2.7 *Let $n \geqslant m > 0$ be integers, and α, β positive numbers such that $\alpha + \beta = 1$. Than*
(i) $\sum_{j=0}^n B(n,j)\alpha^j \beta^{n-j}(m-j) \leqslant \sqrt{(n\alpha - m)^2 + n\alpha\beta}$
and
(ii) $\sum_{j=m}^n B(j-1, m-1)\alpha^m \beta^{j-m}(m-j) \leqslant \sqrt{\frac{m\beta}{\alpha^2} + \left(\frac{m\beta}{\alpha} + m - n\right)^2}$

Proof See Crandall & Ligget Lemma 1.4 ∎

Now we can prove that for each $\rho > 0$ the above defined resolvents $(\mathcal{I}_{\lambda,\rho})_{\lambda \geqslant 0}$ generate a semigroup on X whose paths are Lipschitz on bounded time intervals.

Proposition 2.8 *For each $\rho > 0$ and each $x \in X$ the sequence $\mathcal{I}_{\frac{t}{n},\rho}^n x$ converges. Moreover denoting this limit as $S_\rho(t)x$ we have that $S_\rho(t)$ is a semigroup of contractions on X, and for $t, \tau \geqslant 0$*

$$d(S_\rho(t)x, S_\rho(\tau)x) \leqslant \frac{d(x, F_\rho x)}{\rho} |t - \tau| \tag{1}$$

Proof Pick $x \in X$ fix $\rho > 0$ and write $\mathcal{I}_\lambda = \mathcal{I}_{\lambda,\rho}$ for the sake of simplicity. Let $\lambda \geqslant \mu > 0$, $n, m \in \mathbb{N}$, $n \geqslant m$. By Lemma 2.6 Lemma 2.3 and Cor. 2.5 we have

$$\begin{aligned} d(\mathcal{I}_\mu^n x, \mathcal{I}_\lambda^m x) &\leqslant \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} B(n,j) d(\mathcal{I}_\lambda^{m-j} x, x) \\ &+ \sum_{j=m}^{n} \alpha^m \beta^{j-m} B(j-1, m-1) d(\mathcal{I}_\mu^{n-j} x, x) \\ &\leqslant \lambda \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} B(n,j)(m-j) \frac{d(x, F_\rho x)}{\rho} \\ &+ \mu \sum_{j=m}^{n} \alpha^m \beta^{j-m} B(j-1, m-1)(n-j) \frac{d(x, F_\rho x)}{\rho} \end{aligned}$$

where as in Lemma 2.6 $\alpha = \frac{\mu}{\lambda}$, $\beta = \frac{\lambda-\mu}{\lambda}$. Lemma 2.7 yields

$$\begin{aligned} d(\mathcal{I}_\lambda^m x, \mathcal{I}_\mu^n x) &\leqslant \lambda \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} (m-j) B(n,j) \frac{d(x, F_\rho x)}{\rho} \\ &+ \mu \sum_{j=m}^{n} \alpha^m \beta^{m-j} (n-j) B(j-1, m-1) \frac{d(x, F_\rho x)}{\rho} \\ &\leqslant \Bigg\{ \lambda \sqrt{\left(n\frac{\mu}{\lambda} - m\right)^2 + n \frac{\mu}{\lambda} \frac{\lambda-\mu}{\lambda}} \\ &+ \sqrt{\frac{\lambda^2}{\mu^2} \frac{\lambda-\mu}{\lambda} m + \left(\frac{\lambda}{\mu} \frac{\lambda-\mu}{\lambda} m + m - n\right)^2} \Bigg\} \frac{d(x, F_\rho x)}{\rho} \\ &= \{ \sqrt{(n\mu - \lambda m)^2 + n\mu(\lambda-\mu)} + \sqrt{m\lambda(\lambda-\mu) + (m\lambda - n\mu)^2} \} \frac{d(x, F_\rho x)}{\rho} \end{aligned} \tag{3}$$

Taking $\mu = \frac{t}{n}$ $\lambda = \frac{t}{m}$ above gives

$$d(\mathcal{I}_{\frac{t}{m}}^m x, \mathcal{I}_{\frac{t}{n}}^n x) \leqslant 2t \sqrt{\frac{1}{m} - \frac{1}{n}} \frac{d(x, F_\rho x)}{\rho} \tag{4}$$

hence $\lim_{n\to\infty} \mathcal{I}_{\frac{t}{n}}^n x$ exists indeed. Denote the limit by $S_\rho(t)x$. Further for $\tau > t \geqslant 0$ $n = m$ we can take $\mu = \frac{t}{n}$ $\lambda = \frac{\tau}{n}$ in (3) and take the limit for $n \to \infty$ to obtain

$$d(S_\rho(\tau)x, S_\rho(t)x) \leqslant \frac{d(x, F_\rho x)}{\rho} (\tau - t), \tag{5}$$

that is $t \mapsto S_\rho(t)x$ is Lipschitz cond. S curve.

It remains to show that $(S_\rho(t))_{t\geqslant 0}$ ($S_\rho(t)$ is by definition the identity map on X) is a contraction semigroup, that is for $t, \tau > 0$ $S_\rho(t+\tau) = S_\rho(t)S_\rho(\tau)$, and that $d(S_\rho(t)x, S_\rho(t)y) \leqslant d(x,y)$ for $t \geqslant 0$ $x, y \in X$.

For the later observe that

$$d(S_\rho(t)x, S_\rho(t)y) = \lim_{n\to\infty} d(\mathcal{I}^n_{\frac{t}{n}}x, \mathcal{I}^n_{\frac{t}{n}}y) \leqslant \lim_{n\to\infty} d(x,y)$$

as $\mathcal{I}_{\frac{t}{n}}$ is a contraction for $t > 0$ $n \in \mathbb{N}$.

First of all for $m \in \mathbb{N}$ $x \in X$

$$d\left(S^m_\rho(t)x, \left(\mathcal{I}^n_{\frac{t}{n}}\right)^m x\right) \leqslant d\left(S_\rho(t)S^{m-1}_\rho(t)x, \mathcal{I}^n_{\frac{t}{n}}S^{m-1}_\rho(t)x\right) + d\left(\mathcal{I}^n_{\frac{t}{n}}S^{m-1}_\rho(t)x, \mathcal{I}^n_{\frac{t}{n}}\mathcal{I}^{m(m-1)}_{\frac{t}{n}}x\right)$$

so since $\mathcal{I}^n_{\frac{t}{n}}$ is a contraction, by induction we obtain

$$S^m_\rho(t)x = \lim_{n\to\infty}\left(\mathcal{I}^n_{\frac{t}{m}}\right)^m x = \lim_{n\to\infty}\left(\mathcal{I}^{mn}_{\frac{mt}{mn}}x\right) = S_\rho(mt)x.$$

Next is $l, k, r, s \in \mathbb{N}$ than

$$\begin{aligned} S_\rho\left(\frac{l}{k}+\frac{r}{s}\right) &= S_\rho\left(\frac{ls+rk}{ks}\right) = S_\rho\left(\frac{1}{ks}\right)^{ls+rk} = S_\rho\left(\frac{1}{ks}\right)^{ls} S_\rho\left(\frac{1}{ks}\right)^{rk} = \\ &= S_\rho\left(\frac{ls}{ks}\right) S_\rho\left(\frac{rk}{ks}\right) = S_\rho\left(\frac{l}{k}\right) S_\rho\left(\frac{r}{s}\right) = S_\rho\left(\frac{l}{k}\right) S_\rho\left(\frac{r}{s}\right) \end{aligned}$$

hence $S_\rho(t+\tau) = S_\rho(t)S_\rho(\tau)$ for $t, \tau \in Q_+$.

Finally let $t_n \to t$ $\tau_n \to \tau$, $t_n, \tau_n \in Q_+$, $x \in X$. $S_\rho(t)$ is a contraction each n, and the paths of S_ρ are continuous we have $S_\rho(t_n, \tau_n)x \to S_\rho(t+\tau)x$ and

$$\begin{aligned} d(S_\rho(t_n), S_\rho(\tau_n)x, S_\rho(t)S_\rho(\tau)x) &\leqslant d(S_\rho(t_n)S_\rho(\tau_n)x, S_\rho(t_n)S_\rho(\tau)x) \\ &\quad + d(S_\rho(t_n)S_\rho(\tau)x, S_\rho(t)S_\rho(\tau)x) \\ &\leqslant d(S_\rho(\tau_n)x, S_\rho(\tau)x) \\ &\quad + d(S_\rho(t_n)S_\rho(\tau)x, S_\rho(t)S_\rho(\tau)x) \\ &\to 0 \quad as \quad n \to \infty \end{aligned}$$

and we completed the proof. ∎

The following lemma is a tool that we will need in the ... (see also Thm 1.7 in Bresis).

Lemma 2.9 *Suppose $F_\rho = F$ for all $\rho > 0$. Than for $t \geqslant 0$ $x \in X$ $S_\rho(t) = S_1\left(\frac{t}{\rho}\right)$,*

$$d(S_1(t)x, F^m x) \leqslant \sqrt{(m-t)^2 + t} + d(x, Fx), \tag{6}$$

and

$$d(S_\rho(t)x, F^m x) \leqslant \sqrt{\left(m - \frac{t}{\rho}\right)^2 + \frac{t}{\rho}} d(x, Fx). \tag{7}$$

Proof The last claim is a direct consequence of the first two.

The first claim follows by the simple observation that $\mathcal{I}_{\lambda,\rho} = \mathcal{I}_{\frac{\lambda}{\rho},1}$, hence for $x \in X$

$$S_\rho(t)x \overset{n\to\infty}{\longleftarrow} \mathcal{I}^n_{\frac{t}{n},\rho}x = \mathcal{I}^n_{\frac{t/\rho}{n},1}x \overset{n\to\infty}{\longrightarrow} S_1\left(\frac{t}{\rho}\right)x$$

For the second claim we need to make more efort. Recall that $\mathcal{I}_\lambda x := \mathcal{I}_{\lambda,1}x$ is (unique) point y in X which is a point $\frac{\lambda}{1+\lambda}$ time fraction away from x on the geodesic connecting x and Fy. Hence by the NPC property

$$d^2(F^m x, \mathcal{I}^n_{t/n}x) \leqslant \frac{1}{1+\frac{t}{n}} d^2(F^m x, \mathcal{I}^{n-1}_{\frac{t}{n}}x) + \frac{\frac{t}{n}}{1+\frac{t}{n}} d^2(F^m x, F\mathcal{I}^{n-1}_{\frac{t}{n}}x)$$

Repeating the argument n times we obtain

$$\begin{aligned} d^2(F^m x, \mathcal{I}^n_{\frac{t}{n}}x) &\leqslant \left(1+\frac{t}{n}\right)^{-n} d^2(F^m x, x) + \frac{t}{n}\sum_{k=1}^{n}\left(1+\frac{t}{n}\right)^{-k} d^2(F^m x, F\mathcal{I}^{n-k+1}_{\frac{t}{n}}x) \\ &\leqslant \left(1+\frac{t}{n}\right)^{-n} m^2 d^2(Fx, x) + \frac{t}{n}\sum_{k=1}^{n}\left(1+\frac{t}{n}\right)^{-k} d^2(F^{m-1}x, \mathcal{I}^{n-k+1}_{\frac{t}{n}}x) \\ &= \left(1+\frac{t}{n}\right)^{-n} m^2 d^2(x, Fx) + \frac{t}{n}\sum_{k=1}^{n}\left(1+\frac{t}{n}\right)^{-(n-k+1)} d^2(F^{m-1}x, \mathcal{I}^{k}_{\frac{t}{n}}x) \end{aligned}$$

Set

$$\begin{aligned} f_n(s) &:= \sum_{k=1}^{n}\left(1+\frac{t}{n}\right)^{-(n-k+1)} 1_{\left(\frac{(k-1)t}{n},\frac{kt}{n}\right]}(s) \\ g_n(s) &:= \sum_{k=1}^{n} d^2(F^{m-1}x, \mathcal{I}^k_{\frac{t}{n}}x) 1_{\left(\frac{(k-1)t}{n},\frac{kt}{n}\right]}(s) \end{aligned}$$

With this notation the above inequality becomes

$$d^2(F^m x, \mathcal{I}^n_{\frac{t}{n}}x) \leqslant \left(1+\frac{t}{n}\right)^{-n} m^2 d^2(x, Fx) + \int_0^t f_n(s)g_n(s)ds$$

If we show that $f_n(s) \to e^{-(t-s)}$ $g_n(s) \to d^2(F^{m-1}x, S_{(s)}x)$ for $s \in (0,t]$ and that $\sup_n \sup_{s\in[0,t)} |g_n(s)||f_n(s)| < \infty$ we will have by dominated convergence

$$d^2(F^m x, S_{(t)}x) \leqslant e^{-t}m^2 d^2(x, Fx) + \int_0^t e^{-(t-1)} d^2(F^{m-1}x, S_{(s)}x)ds \qquad (8)$$

Where we write $S(s) := S_1(s)$.

Pick on $n \in \mathbb{N}$ and $s \in (0,t]$. There is a unique $0 < k_{s,n} \leqslant n$ s.t. $\frac{(k_{s,n}-1)t}{n} < s < \frac{k_{s,n}t}{n}$.

Filling in $k_{s,t}$, $\frac{t}{n}$, N, $\frac{k_{s,n}t}{nN}$ for m, λ, n, μ respectively in (3) where $\mathbb{N} \ni N \geqslant n$ gives

$$d\left(\mathcal{I}^{k_{s,n}}_{\frac{t}{n}}x, S_{\left(\frac{k_{s,n}t}{n}\right)}x\right) \leqslant 2\sqrt{\frac{k_{s,n}t^2}{n^2}}d(x, Fx)$$

By (5)

$$d\left(\mathcal{I}_{\frac{t}{n}}^{k_{s,n}}x, \mathcal{I}_{\frac{k_{sn}t}{nN}}^{N}x\right) \leqslant 2\sqrt{\frac{k_{s,n}t}{n}\left(\frac{t}{n}-\frac{k_{s,n}t}{nN}\right)}d(x,Fx)$$

and taking limit for $N \to \infty$ gives

$$\begin{aligned} d\left(\mathcal{I}_{\frac{t}{b}}^{k_{s,n}}x, S_{\left(\frac{k_{s,n}t}{n}\right)}x\right) &\leqslant d(x,Fx)\left(\frac{k_{s,n}t}{n}-s\right) \\ &\leqslant d(x,Fx)\left(\frac{k_{s,n}t-(k_{s,n}-1)t}{n}\right) = d(x,Fx)\frac{t}{n} \end{aligned}$$

So we have indeed

$$\begin{aligned} \left|g_n(s) - d^2(F^{m-1}x, S(s)x)\right| = \Big| &\left(d\left(F^{m-1}x, \mathcal{I}_{\frac{t}{n}}^{k_{s,n}}x\right) + d\left(\mathcal{I}_{\frac{t}{n}}^{k_{s,n}}, S\right)\right) \\ &\cdot \left(d\left(F^{m-1}x, \mathcal{I}_{\frac{t}{n}}^{k_{s,n}}x\right) - d\left(F^{m-1}x, S(s)x\right)\right)\Big| \end{aligned}$$

As we can see by (5) that $\left\{\mathcal{I}_{\frac{t}{n}}^{k}x | n \in \mathbb{N}\ 0 \leqslant k \leqslant n\right\}$ is a bounded sequence, the above argument yield indeed that

$$g_n(s) \to d^2(F^{m-1}x, S(s)x)$$

for $s \in (0,t]$ and that g_n is a uniformly bounded sequence of functions. The sequence f_n is even easier to handle, and we omit the details. Hence we established (8). To complete the proof set

$$\varphi_m(t) := \frac{d(F^m x, S(t)x)}{d(x,Fx)} \quad m \geqslant 0$$

We have by (8) for $m \geqslant 1$

$$\varphi_m^2 \leqslant e^{-t}m^2 + \int_0^t e^{-|t-s|}\varphi_{m-1}^2(s)ds$$

We show by induction that $\varphi_m(t) \leqslant ((m-t)^2+t)$.

By (5) $\varphi_0 \leqslant t$.

Suppose the claim is true for $m-1$. Than

$$\varphi_m^2(t) \leqslant e^{-t}m^2 + \int_0^t e^{-(t-s)}((m-1-s)^2+s)ds =: \psi_m(t)$$

So it's enough to show that

$$m^2 + \int_0^t e^s((m-1-s)^2+s)ds = e^t((m-t)^2+t).$$

For $t=0$ the above is $m^2=m^2$, the derivative of the right side is

$$e^t((m-1-t)^2+t) = e^t((m-t)^2+t) - 2(m-t)e^t + e^t = \frac{d}{dt}e^t((m-t)^2+t)$$

so we are done. ∎

First we introduce some notation.

The two propositions are essentially theorem 3.1 and theorem 3.2 in Bresis & Pazy, adopted to our situation. With the machinery we developed the lack of linear structure ... a problem any more, for these two steps. Let φ^1, φ^2 be proper convex functionals on X satisfying (H2), and denote by $(S^1(t))_{t\geqslant 0}$, $(S^2(t))_{t\geqslant 0}$ the semigroup generated by these functionals and by $\mathcal{I}_t^1, \mathcal{I}_t^2$ the corresponding resolvents. Assume that $\varphi^1 := \varphi(t)\ 0$ in proper for, and let $(S(t))_{t\geqslant 0}$ be the semigroup generated by φ, and let $(\mathcal{I}_\lambda)_{\lambda\geqslant 0}$ be it's resolvents. Assume that $D(|\partial\varphi^1|) \cap D(|\partial\varphi^2|) = X$. Than $S(t), S^1(t), S^2(t)$ are contractions on X, and $F_\rho := S^1(\rho)S^2(\rho)$ as well for $t, \rho > 0$. Finally recall the resolvents $\mathcal{I}_{\lambda,\rho}$ from def **??**, and let $(S_\rho(t))_{t\geqslant 0}$ be the semigroup generated by proposition ...

Proposition 2.10 *Suppose that there is a $\lambda_0 > 0$ s.t. for $0 < \lambda < \lambda_0$ and for $x \in X$*

$$\mathcal{I}_{\lambda,\rho}x \to \mathcal{I}_\lambda x \quad \textit{as} \quad \rho \downarrow 0.$$

Than $\lim_{\rho\downarrow 0} S_\rho(t)x = S(t)x$ for every $x \in X$ and the limit is uniform on bounded time intervals.

Proof Let $x \in D(\partial\varphi^1) \cap D(\partial\varphi^2)$, $t > 0$. We estimate for $\lambda > 0$

$$\begin{aligned} d(S_\rho(t)x, S(t)x) &\leqslant d(S_\rho(t)x, S_\rho(t)\mathcal{I}_{\lambda,\rho}x) + d(S_\rho(t)\mathcal{I}_{\lambda,\rho}x, S(t)x) \\ &\leqslant d(x, \mathcal{I}_{\lambda,\rho}x) + d(S_\rho(t)\mathcal{I}_{\lambda,\rho}x, S(t)x) \\ &\leqslant \frac{\lambda}{\rho} d(x, F_\rho x) + d(S_\rho(t)\mathcal{I}_{\lambda,\rho}x, S(t)x) \end{aligned} \tag{9}$$

by Lemma 2.3. Further for $n \in \mathbb{N}$

$$\begin{aligned} d(S_\rho(t)\mathcal{I}_{\lambda,\rho}x, S(t)x) \leqslant\ & d\left(S_\rho(t)\mathcal{I}_{\lambda,\rho}x, \left(\mathcal{I}_{\frac{t}{n},\rho}\right)^2 \mathcal{I}_{\lambda,\rho}x\right) \\ &+ d\left(\left(\mathcal{I}_{\frac{t}{n},\rho}\right)^n \mathcal{I}_{\lambda,\rho}x, \left(\mathcal{I}_{\frac{t}{n},\rho}\right)^n x\right) \\ &+ d\left(\left(\mathcal{I}_{\frac{t}{n},\rho}\right)^n x, \left(\mathcal{I}_{\frac{t}{n}}\right)^n x\right) + d\left(\left(\mathcal{I}_{\frac{t}{n},\rho}\right)^n x, S(t)x\right) \end{aligned} \tag{10}$$

By (4) in the proof of prop 2.8, the construction of $\mathcal{I}_{\lambda,\rho}x$, and Lemma 2.3 we have

$$\begin{aligned} d\left(S_\rho(t)\mathcal{I}_{\lambda,\rho}x, \left(\mathcal{I}_{\frac{t}{n},\rho}\right)^n \mathcal{I}_{\lambda,\rho}x\right) &\leqslant \frac{2t}{\sqrt{n}} \frac{d(\mathcal{I}_{\lambda,\rho}x.F_\rho\mathcal{I}_{\lambda,\rho}x)}{\rho} \\ &= \frac{2t}{\sqrt{n}} \frac{d(x, \mathcal{I}_{\lambda,\rho}x)}{\rho} \qquad LeidenUniversit \\ &\leqslant \frac{2t}{\sqrt{n}} \frac{d(x, F_\rho x)}{\rho} \end{aligned} \tag{11}$$

As $\mathcal{I}_{\lambda,\rho}$ are contractions, for $n \in \mathbb{N}$

$$d\left(\left(\mathcal{I}_{\frac{t}{n},r}\right)^n \mathcal{I}_{\lambda,\rho}x, \left(\mathcal{I}_{\frac{t}{n},\rho}\right)^n x\right) \leqslant d(\mathcal{I}_{\lambda,\rho}x, x) \leqslant \frac{\lambda d(x, F_\rho x)}{\rho}$$

and by (4.0.15) in AGS (notice that $D(\partial\varphi^1)\cap D(\partial\varphi^2)\subset D(\partial\varphi)$)

$$d\left(\left(\mathcal{I}_{\frac{t}{n}}\right)^n x, S(t)x\right) \leqslant \frac{t}{\sqrt{2}n}|\partial\varphi|(x)$$

By Lemma 2.11 $\frac{d(x,F_\rho x)}{\rho} \leqslant C$ where $C=\sqrt{3|\partial\varphi^2|^2(x)+|\partial\varphi^1|^2(x)}+|\partial\varphi^2|(x)$ so we arrive

$$\begin{aligned} d(S_\rho(t)x, S(t)x) &\leqslant \left(2\lambda+\frac{2t}{\sqrt{n}}\right)C+\frac{t}{2\sqrt{n}}|\partial\varphi|(x)+d\left(\left(\mathcal{I}_{\frac{t}{n},\rho}\right)^n x, \mathcal{I}^n_{\frac{t}{n}}x\right)\\ &\leqslant \left(2\lambda+\frac{2T}{\sqrt{n}}\right)C+\frac{T}{2\sqrt{n}}|\partial\varphi|(x)+d\left(\left(\mathcal{I}_{\frac{t}{n},\rho}\right)^n x, \mathcal{I}^n_{\frac{t}{n}}x\right)\end{aligned}$$

for any $\lambda>0$, $n\in\mathbb{N}$ and $t\leqslant T$ where $T<\infty$ is fixed. Now we can choose λ and n so that the first two terms above are smaller than a fixed $s>0$. Finally as

$$\begin{aligned} d\left(\mathcal{I}^n_{\frac{t}{n},\rho}x, \mathcal{I}^n_{\frac{t}{n}}x\right) &\leqslant d\left(\mathcal{I}^n_{\frac{t}{n},\rho}x, \mathcal{I}^{n-1}_{\frac{t}{n},\rho}\mathcal{I}_{\frac{t}{n}}x\right)+d\left(\mathcal{I}^{n-1}_{\frac{t}{n},\rho}\mathcal{I}_{\frac{t}{n}}x, \mathcal{I}^{n-1}_{\frac{t}{n}}\mathcal{I}_{\frac{t}{n}}x\right)\\ &\leqslant d\left(\mathcal{I}_{\frac{t}{n},\rho}x, \mathcal{I}_{\frac{t}{n}}x\right)+d\left(\mathcal{I}^{n-1}_{\frac{t}{n},\rho}\mathcal{I}_{\frac{t}{n}}x, \mathcal{I}^{n-1}_{\frac{t}{n}}\mathcal{I}_{\frac{t}{n}}x\right)\end{aligned}$$

the assumption of the proposition and induction yield that for ρ small enough $d\left(\mathcal{I}^n_{\frac{t}{n},\rho}x, \mathcal{I}^n_{\frac{t}{n}}x\right)<\epsilon$.

Next, we show that for $x\in D(|\partial\varphi^1|)\cap D(|\partial\varphi^2|)$ the limit is uniform for $t\in(0,T]$. Fix a $\tau\in(0,T]$. Than

$$\begin{aligned} d(S_\rho(t)x, S(t)x) &\leqslant d(S_\rho(t)\mathcal{I}_{\lambda,\rho}x, S(t)x)+\lambda C\\ &\leqslant d(S_\rho(t)\mathcal{I}_{\lambda,\rho}x, S_\rho(\tau)\mathcal{I}_{\lambda,\rho}x)\\ &+d(S_\rho(\tau)\mathcal{I}_{\lambda,\rho}x, S(\tau)x)\\ &+d(S(\tau)x, S(t)x)+\lambda C\\ &\leqslant \left(\frac{d(\mathcal{I}_{\lambda,\rho}x, F_\rho\mathcal{I}_{\lambda,\rho}x)}{\rho}+\frac{d(x,F_\rho x)}{\rho}\right)|t-\tau|\\ &+d(S_\rho(\tau)\mathcal{I}_{\lambda,\rho}x, S(\tau)x)+\lambda C\\ &\leqslant 2C|t-\tau|+d(S_\rho(\tau))\mathcal{I}_{\lambda,\rho}x, S(\tau)x)+\lambda C\end{aligned} \tag{12}$$

and as we already showed that the last term converges to ... as $\rho\downarrow 0$ for any $\lambda>0$, we can finish the argument by compactness of $[0,T]$.

At last let $x\in X$, $\varepsilon>0$ and pick $y\in D(|\partial\varphi^1|)\cap D(|\partial\varphi^2|)$ s.t. $d(x,y)<\epsilon$. Than

$$\begin{aligned} d(S_\rho(t)x, S(t)x) &\leqslant d(S_\rho(t)x, S_\rho(t)y)+d(S_\rho(t)y, S(t)y)\\ &+d(S(t)y, S(t)x)\\ &\leqslant 2d(x,y)+d(S_\rho(t)y, S(t)y)\\ &\leqslant 2\epsilon+d(S_\rho(t)y, S(t)y)\end{aligned} \tag{13}$$

hence by the previous the limit is uniform for $x\in X$. ∎

Lemma 2.11 *Let $D(|\partial\varphi^1|)\cap D(|\partial\varphi^2|) = X$ and $x \in D(|\partial\varphi^1|)\cap D(|\partial\varphi^2|)$. Than*

$$\limsup_{\rho\downarrow} \frac{d(x, \mathcal{I}^2_\rho \mathcal{I}^1_\rho x)}{\rho} < \infty \tag{14}$$

and

$$\limsup_{\rho\downarrow} \frac{d(x, F_\rho x)}{\rho} < \infty \tag{15}$$

Proof We start with (14). By (3.1.20) in AGS

$$\frac{d(\mathcal{I}^2_\rho x, x)}{\rho} \leqslant |\partial\varphi^2|(x),$$

and as

$$\frac{d(\mathcal{I}^2_\rho \mathcal{I}^1_\rho x, x)}{\rho} \leqslant \frac{d(\mathcal{I}^2_\rho \mathcal{I}^1_\rho x, \mathcal{I}^2_\rho x) + d(\mathcal{I}^2_\rho x, x)}{\rho}$$

it is enough to show that $\frac{d(\mathcal{I}^2_\rho \mathcal{I}^1_\rho x, \mathcal{I}^2_\rho x)}{\rho}$ is bounded as $\rho \downarrow 0$. We have

$$\frac{1}{2\rho} d^2(\mathcal{I}^2_\rho \mathcal{I}^1_\rho x, \mathcal{I}^2_\rho x) - \frac{1}{2\rho} d^2(x, \mathcal{I}^2_\rho \mathcal{I}^1_\rho x) + \varphi^2(\mathcal{I}^2_\rho x) \leqslant \varphi^2(\mathcal{I}^2_\rho \mathcal{I}^1_\rho x) \tag{16}$$

$$\frac{1}{2\rho} d^2(x, \mathcal{I}^2_\rho \mathcal{I}^1_\rho x) - \frac{1}{2\rho} d^2(\mathcal{I}^1_\rho x, x) + \varphi^2(\mathcal{I}^2_\rho \mathcal{I}^1_\rho x) \leqslant \varphi^2(x) \tag{17}$$

where the first inequality is obtained by applying (4.1.2) in AGS to φ^2 with $u = x$ aLeiden Universitnd $v = \mathcal{I}^2_\rho \mathcal{I}^1_\rho x$ and the second one with $u = \mathcal{I}^1_\rho x$, $v = x$, and in each equation a positive term in left out. Adding these two inequalities and dividing by ρ yields (with the end of (3.1.20) in AGS)

$$\begin{aligned} \frac{1}{2\rho^2} d^2(\mathcal{I}^2_\rho \mathcal{I}^1_\rho x, \mathcal{I}^2_\rho x) &\leqslant \frac{\varphi^2(x) - \varphi^2(\mathcal{I}^2_\rho x)}{\rho} + \frac{1}{2\rho^2} d^2(x, \mathcal{I}^1_\rho x) \\ &\leqslant |\partial\varphi^2|^2(x) + \frac{1}{2\rho^2} d^2(x, \mathcal{I}^2_\rho x) + \frac{1}{2\rho^2} d^2(x, \mathcal{I}^1_\rho x) \\ &\leqslant \frac{3}{2} |\partial\varphi^2|^2(x) + |\partial\varphi^1|^2(x) \end{aligned} \tag{18}$$

hence

$$\frac{d(\mathcal{I}^2_\rho \mathcal{I}^1_\rho x, x)}{\rho} \leqslant |\partial\varphi^2|(x) + \sqrt{3|\partial\varphi^2|^2(x) + |\partial\varphi^1|^2(x)} \tag{19}$$

for any $\rho > 0$.

To prove (15) observe that (4.0.12) in AGS yields that $t \mapsto \varphi^2(S^2(t)y)$ is non-increasing for any $y \in D(\varphi^2)$, hence by (4.0.13)

$$\frac{1}{2\rho} d^2(S^2(\rho)y, z) - \frac{1}{2\rho} d^2(y, z) + \varphi^2(S(\rho)y) \leqslant \varphi^2(z)$$

holds for any $z \in D(\varphi^2)$ $y \in D(\varphi^2)$. Now as by (4.0.12) in AGS for any $\rho > 0$ $\varphi^2(S^2_\rho S^1_\rho x) < \infty$ and we assumed $x \in D(|\partial\varphi^2|) \subset D(\varphi^2)$ we can argue in exactly same way as we did showing (14) until (18) to obtain (applying 2.4.9 in AGS).

$$\begin{aligned}\frac{1}{2\rho^2} d^2(S^2_\rho S^1_\rho x, S^2_\rho x) &\leqslant \frac{\varphi^2(x) - \varphi^2(S^2_\rho x)}{d(S^2_\rho x, x)} \cdot \frac{d(S^2_\rho x, x)}{\rho} + \frac{1}{2\rho^2} d^2(x, S^x_\rho) \\ &\leqslant |\partial\varphi^2|(x) \frac{d(S^2_\rho x, x)}{\rho} + \frac{d^2(x, S^1_\rho x)}{2\rho^2}\end{aligned}$$

and we can finish applying triangle inequality (4.0.15) and (3.1.20) in (AGS). ∎

Proposition 2.12 *Suppose that $\mathcal{I}_{\lambda,\rho} x \to \mathcal{I}_\lambda x$ for $x \in X$. Than $\left(F_{\frac{t}{n}}\right)^n x \to S(t)x$ for $x \in X$, that is the Trotter product formula holds.*

Proof For $\rho > 0$ $n \in \mathbb{N}$ $\lambda > 0$ $x \in X$ we have

$$\begin{aligned} d(S_\rho(n\rho)x, F^n_\rho x) &\leqslant d(S_\rho(n\rho)x, S_\rho(n\rho)\mathcal{I}_{\lambda,\rho} x) \\ &\quad + d(S_\rho(n\rho)\mathcal{I}_{\lambda,\rho} x, F^n_\rho \mathcal{I}_{\lambda,\rho} x) + d(F^n_\rho \mathcal{I}_{\lambda,\rho} x, F^n_\rho x) \\ &\leqslant 2d(x, \mathcal{I}_{\lambda,\rho} x) + d(S_\rho(n\rho)\mathcal{I}_{\lambda,\rho} x, F^n_\rho \mathcal{I}_{\lambda,\rho} x) \end{aligned} \tag{20}$$

Let $(\tilde{S}_\rho(t)))t \geqslant 0$ be the semigroup generated by the resolvents $\tilde{\mathcal{I}}_{\lambda,\rho} := \mathcal{I}_{\lambda\rho,\rho}$ for $\lambda \geqslant 0$. Than by Lemma 2.9 $\tilde{S}_\rho(t) = S_\rho(t\rho)$, and with the aid of the same lemma (part (i)).

$$\begin{aligned} &d(S_\rho(n\rho)\mathcal{I}_{\lambda,\rho} x, F^n_\rho \mathcal{I}_{\lambda,\rho} x) = d(\tilde{S}_\rho(n)\mathcal{I}_{\lambda,\rho} x, F^n_\rho \mathcal{I}_{\lambda,\rho} x) \leqslant \\ &\leqslant \sqrt{n} d(\mathcal{I}_{\lambda,\rho} x, F_\rho \mathcal{I}_{\lambda,\rho} x) = \sqrt{n}\frac{\rho}{\lambda} d(x, \mathcal{I}_{\lambda,\rho} x) \leqslant c\sqrt{n}\rho \end{aligned} \tag{21}$$

with the aid of Lemma 2.3, Lemma 2.11 and C as in (19). Taking $\rho = \frac{t}{n}$ above we obtain that for any $\lambda > 0$

$$d\left(S_{\frac{t}{n}}(t)x, \left(F_{\frac{t}{n}}\right)^n x\right) \leqslant 2\lambda C + C\frac{t}{\sqrt{n}} \tag{22}$$

hence as $S_{\frac{t}{n}}(t)x \to S(t)x$, uniformly for $t \in [0, T]$ for any fixed $T < \infty$ (by proposition 2.10) we have that $\left(F_{\frac{t}{n}}\right)^n x \to S(t)x$ uniformly on $[0, T]$. ∎

3 The Trotter Product Formula

Let $x \in X$, $t > 0$ and let $y_0(t) := \mathcal{I}_{\lambda,t} x$, $y_1(t) = S^1_t y_0(t)$ $y_2(t) = S^2_t S^1_t \mathcal{I}_{\lambda,t} x ==$ $F_t y_0(t)$. Applying (4.0.13) in AGS yields

$$\frac{1}{2t} d^2(y_1(t), v) - \frac{1}{2t} d^2(y_0(t), v) + \varphi^1(y_1(t)) \leqslant \varphi^1(v) \quad \forall v \in D(\varphi^1)\ \forall t > 0 \tag{23}$$

$$\frac{1}{2t}d^2(y_2(t),v) - \frac{1}{2t}d^2(y_1(t),v) + \varphi^2(y_2(t)) \leqslant \varphi^2(v) \quad \forall v \in D(\varphi^2)\ \forall t > 0 \tag{24}$$

and adding these inequalities yields

$$\frac{1}{2t}d^2(y_2(t),v) - \frac{1}{2t}d^2(y_0(t),v) + \varphi^1(y_1(t)) + \varphi^2(y_2(t)) \leqslant \varphi(v) \quad \forall v \in D(\varphi^1) \cap D(\varphi^2)\ \forall t > 0 \tag{25}$$

Further as $y_0(t) = \frac{1}{1+\frac{\lambda}{t}}x + \frac{\frac{\lambda}{t}}{1+\frac{\lambda}{t}}y_2(t)$ the NPC inequality gives

$$\frac{1}{2t}d^2(y_o(t),v) \leqslant \frac{1}{2t}\frac{1}{1+\frac{\lambda}{t}}d^2(x,v) + \frac{1}{2t}\frac{\frac{\lambda}{t}}{1+\frac{\lambda}{t}}d^2(y_0(t),v) - \frac{1}{2t}\frac{\frac{\lambda}{t}}{\left(1+\frac{\lambda}{t}\right)^2}d^2(x,y_2(t))$$

hence together with (25) and $d^2(x,y_2(x)) = \left(\frac{1+\frac{\lambda}{t}}{\frac{\lambda}{t}}\right)^2 d^2(x,y_0(t))$

$$\frac{1}{2(\lambda+t)}d^2(y_2(t),v) - \frac{1}{2(\lambda+t)}d^2(x,v) + \frac{\lambda}{2(\lambda+t)^2}\left(\frac{\lambda+t}{\lambda}\right)^2 d^2(y_0(t),x) + \varphi^1(y_1(t)) + \varphi^2(y_2(t)) \leqslant \varphi(v) \tag{26}$$

$\forall t > 0\ \forall v \in D(\varphi^1) \cap D(\varphi^2)$.

Further $d^2(y_2(t),v) \geqslant d^2(y_2(t),x) - 2d(y_2(t),x)d(x,v) + d^2(x,v)$ and there are constants $c \in \mathbb{R}$ $b \in \mathbb{R}$ s.t.

$$\varphi^2(y_i(t) \geqslant c - bd(x,y_i(t)) \quad i = 1,2\ t > 0. \tag{27}$$

As

$$d(x,y_1(t)) \leqslant d(x,S_t^1x) + d(S_t^1x,y_1(t)) \leqslant d(x,S_t^1x) + d(x,y_0(t)) \tag{28}$$

and

$$\begin{aligned} d(x,y_2(t)) &\leqslant d(x,S_t^2x) + d(S_t^2x,F_tx) + d(F_tx,y_2(t)) \\ &\leqslant d(x,s_t^2x) + d(x,S_t^1x) + d(x,y_0(t)) \end{aligned} \tag{29}$$

So as $S_t^1x \to x$, $S_t^2x \to x$ for $t \downarrow 0$ there are constants $c',b' \in \mathbb{R}$ $\varphi^1(y_1(t)) + \varphi^2(y_2(t)) \geqslant c' - b'd(x,y_0(t))$ for t small enough. Implementation of the above listed facts in (26) yields

$$\left(\frac{1}{2(\lambda+t)} + \frac{1}{2\lambda}\right)d^2(x,y_0(t)) - \frac{1}{\lambda+t}d(x,y_0(t))d(x,v) + c' - b'd(x,y_0(t)) \leqslant \varphi(v) \tag{30}$$

$\forall v \in D(\varphi^1) \cap D(\varphi^2)$, for t small enough.

We can further estimate by replacing the first t above by 1 and the second one by 0 to obtain

$$\left(\frac{1}{2\lambda} + \frac{1}{2(\lambda+1)}\right)d^2(x,y_0(t)) - \left(\frac{1}{\lambda}d(x,v) + b'\right)d(x,y_0(t)) + c' \leqslant \varphi(v) \tag{31}$$

for $t < 1$ small enough and $\forall v \in D(\varphi^1) \cap D(\varphi^2)$.

Well this implies that $d(x, y_0(t))$ is bounded as $t \downarrow 0$ hence by (28) and (29) so are $d(x, y_1(t))$ and $d(x, y_2(t))$. But then by (27) $\varphi^2(y_i(t))$ is bounded from bellow for $i = 1, 2$, hence by (25) from above as well. We collect the above deduced facts in the following:

Proposition 3.1 *Let $x \in X$ and set $y_0(t) := \mathcal{I}_{\lambda,t}x$, $y_1(t) := S_t^1 y_0(t)$, $y_2(t) := F_t y_0(t)$. Than $y_0(t), y_1(t), \varphi^1(y_1(t))$, $\varphi^2(y_2(t))$ are bounded as $t \downarrow 0$.*

Further $d(y_0(t), y_2(t)) = \frac{t}{t+\lambda} d(x, y_2(t)) \to 0$ as $t \downarrow 0$ as $y_2(t)$ is bounded, and like in (28), (29).

$$
\begin{aligned}
&d(y_1(t), z) \leqslant d(z, S_t^1 z) + d(z, y_0(t)) \\
&d(y_2(t), z) \leqslant d(z, S_t^2 z) + d(S_t^2 z, y_2(t)) \leqslant d(z, S_t^2 z) + d(z, y_1(t))
\end{aligned}
$$

so $|d(y_2(t), z) - d(y_0(t), z)| \leqslant d(y_2(t), y_0(t)) \to 0$ as $t \downarrow 0$ and as $S_t^i z \to z$ for $t \downarrow 0$ $i = 1, 2$, and $y_0(t), y_1(t), y_2(t)$ are bounded as $t \downarrow 0$ we have showed the following

Lemma 3.2 *$d(y_0(t), y_2(t)) \to 0$ as $t \downarrow 0$, and for $z \in X$ $d^2(y_i(t), z) - d^2(y_0(t), z) \to 0$ as $t \downarrow 0$ for $i = 1, 2$.*

As by 1.1

$$
\frac{1}{2(\lambda + t)} d^2(y_i, v) \geqslant \frac{1}{2\lambda} d^2(y_0, v) - \frac{t}{2\lambda(\lambda + t)} d^2(x, v) + \frac{t}{2(\lambda + t)^2} d^2(x, y_0)
$$

inserting this in (26) we obtain

$$
\frac{1}{2\lambda} d^2(y_0^{(t)}, x) + \frac{1}{2\lambda} d^2(y_0^{(t)}, v) + \frac{t}{2(\lambda + t)^2} d^2(x, y_0^{(t)}) - \frac{t}{2\lambda(\lambda + t)} d^2(x, v) + \varphi^1(y_1(t)) + \varphi^2(y_2(t)) \leqslant \frac{1}{2(\lambda + t)} d^2(x, \tag{32}
$$

$\forall v \in D(\varphi)$.

Hence by proposition 3.1 and Lemma 3.2 $\forall \epsilon > 0$ $\exists \delta > 0$ s.t. for $t < \epsilon$

$$
\begin{aligned}
\frac{1}{2\lambda} d^2(y_2(t), x) + \frac{1}{2\lambda} d^2(y_2(t), \mathcal{I}_l x) + \varphi^1(y_1(t)) + \varphi^2(y_2(t)) &\leqslant \frac{1}{2\lambda} d^2(x, \mathcal{I}_\lambda x) + \varphi(\mathcal{I}_l x) + \epsilon \\
&= \varphi_\lambda(x) + \epsilon
\end{aligned} \tag{33}
$$

Now from (26) and Lemma 3.2 we see that for any sequence $t_n \downarrow 0$ the only accumulation point of $y_0(t_n)$ is $\mathcal{I}_\lambda x$. Indeed suppose $y_0(t_{n_k}) \to z$ as $k \to \infty$ for some sequence $\{t_{n_k}\}$. Than by Lemma 3.2 $y_1(t_{n_k}) \to z$ and $y_2(t_{n_k}) \to z$ hence taking lim inf in (26) we obtain

$$
\frac{1}{2\lambda} d^2(z, v) + \frac{1}{2\lambda} d^2(z, x) + \varphi(z) \leqslant \frac{1}{2\lambda} d^2(x, v) + \varphi^2(v) \quad \forall v \in D(\varphi)
$$

hence $z = \mathcal{I}_\lambda x$. Now we can prove the main result of this paper:

Theorem 3.3 *Suppose X is a complete geodesic NPC space and let φ^1, φ^2 be two proper convex functionals on X s.t. $D(|\partial\varphi^1|) \cap D)|\partial\varphi^2|) = X$. Than if the bounded level sets of one of these functionals are relatively compact (i.e. if for $i = 1$ or $i = 2$ we have that $\forall x \in X$ $\forall r > 0$ $\forall c \in \mathbb{R}$ $B_r(x) \cap \{x|\varphi^2(x) \leqslant c\}$ is relatively compact) the Trotter product formula holds:*

$$\left(S^1_{\frac{t}{n}} S^2_{\frac{t}{n}}\right)^n \to S_t x \quad \textit{for} \quad \forall x \in X$$

and uniformly on compact time intervals.

Proof Suppose the assumption holds for $i = 1$. Let $t_n \downarrow 0$. By proposition 3.1 $\{y_1(t_n)\}_n$ is relatively compact. Let $\{y_1(t_{n_k})\}_k$ be subsequence converging to a limit z. By (**??**) $y_0(t_{n_k})$ and $y_2(t_{n_k})$ also converge to z. Hence by the argument before this theorem $z = \mathcal{I}_\lambda x$. But than we must have that $y_0(t) \to \mathcal{I}_\lambda x$ and we can apply proposition 2.12. ∎

Corollary 3.4 *Suppose X is a complete symply connected Riemanian manifold, and φ^1, φ^2 two convex functionals. Than the Trotter product formula holds (see [2])*

Brezis, H.; Pazy, A. Convergence and approximation of semigroups of nonlinear operators in Banach spaces. J. Functional Analysis 9 (1972), 6374.

Proof By the Nesh imbedding thm. X can be nameticaly imbedded in $\mathbb{R}^{2n+1}$ where $n = \dim X$. Hence closed ball are compact. ∎

References

[1] Mayer, Uwe F. Gradient flows on nonpositively curved metric spaces and harmonic maps. Comm. Anal. Geom. 6 (1998), no. 2, 199–253.

[2] Ambrosio, Luigi; Gigli, Nicola; Savar, Giuseppe Gradient flows in metric spaces and in the space of probability measures. Lectures in Mathematics ETH Zrich. Birkhuser Verlag, Basel, 2005. viii+333 pp. ISBN: 978-3-7643-2428-5; 3-7643-2428-7 49-02 (28A33 35K55 35K90 49Q20 60B10)

[3] Jordan, Richard; Kinderlehrer, David; Otto, Felix The variational formulation of the Fokker-Planck equation. (English summary) SIAM J. Math. Anal. 29 (1998), no. 1, 1–17.

[4] Otto, Felix The geometry of dissipative evolution equations: the porous medium equation. Comm. Partial Differential Equations 26 (2001), no. 1-2, 101–174.

[5] MR0287357 (44 4563) Crandall, M. G.; Liggett, T. M. Generation of semi-groups of nonlinear transformations on general Banach spaces. Amer. J. Math. 93 1971 265–298. 47.50

[6] Kato, Tosio; Masuda, Kyya Trotter's product formula for nonlinear semigroups generated by the subdifferentials of convex functionals. J. Math. Soc. Japan 30 (1978), no. 1, 169–178.