

A forward-backward-forward differential equation and its asymptotic properties

Sebastian Banert* Radu Ioan Boţ†

March 26, 2015

Abstract. In this paper, we approach the problem of finding the zeros of the sum of a maximally monotone operator and a monotone and Lipschitz continuous one in a real Hilbert space via an implicit forward-backward-forward dynamical system with nonconstant relaxation parameters and stepsizes of the resolvents. Besides proving existence and uniqueness of strong global solutions for the differential equation under consideration, we show weak convergence of the generated trajectories and, under strong monotonicity assumptions, strong convergence with exponential rate. In the particular setting of minimizing the sum of a proper, convex and lower semicontinuous function with a smooth convex one, we provide a rate for the convergence of the objective function along the ergodic trajectory to its minimum value.

Key Words. implicit dynamical system, continuous forward-backward-forward method, Lyapunov analysis, monotone inclusions, convex optimization

AMS subject classification. 34G25, 47H05, 90C25

1 Introduction

In this paper, we address the monotone inclusion problem

$$\text{find } \bar{x} \in \mathcal{H} \text{ such that } 0 \in A\bar{x} + B\bar{x}, \quad (1)$$

where \mathcal{H} is a real Hilbert space, $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is a maximally monotone operator and $B : \mathcal{H} \rightarrow \mathbb{R}$ is a monotone and $\frac{1}{\beta}$ -Lipschitz continuous operator for $\beta > 0$, by means of the dynamical system of equations

$$\begin{cases} z(t) = J_{\gamma(t)A}(x(t) - \gamma(t)Bx(t)) \\ 0 = \dot{x}(t) + x(t) - z(t) - \gamma(t)Bx(t) + \gamma(t)Bz(t) \\ x(0) = x_0, \end{cases} \quad (2)$$

*University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria, sebastian.banert@univie.ac.at

†University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria, radu.bot@univie.ac.at.

where $\gamma : [0, +\infty) \rightarrow (0, \beta)$ is a Lebesgue measurable function, $x_0 \in \mathcal{H}$ and $J_{\gamma(t)A}$ denotes the resolvent of the operator $\gamma(t)A$ for every $t \in [0, +\infty)$.

The pioneering work [16] of Crandall and Pazy represented a cornerstone in the study of dynamical systems governed by maximally monotone operators in Hilbert spaces, as it addressed questions like the existence and uniqueness of solution trajectories and it related the latter to the theory of semi-groups of nonlinear contractions. Brezis has studied in [14] the asymptotic behavior of the trajectories whenever the underlying operator is the convex subdifferential and Bruck proved in [15] that a similar asymptotic convergence analysis can be made also in the general case involving an arbitrary maximally monotone operator.

Dynamical systems governed by maximally monotone operators are recognized as valuable tools for studying numerical algorithms for monotone inclusions and optimization problems obtained by time discretization of the continuous dynamics (cf. [19]). In this context we want to refer to the discrete forward-backward-forward algorithm (see [7, 20]) which generates for an initial point $x_0 \in \mathcal{H}$ and a sequence of stepsizes $(\gamma_n)_{n \geq 0} \subseteq (0, \beta)$, via the iterative scheme

$$(\forall n \geq 0) \begin{cases} z_n := J_{\gamma_n A}(x_n - \gamma_n Bx_n) \\ x_{n+1} := z_n + \gamma_n(Bx_n - Bz_n), \end{cases} \quad (3)$$

two sequences $(x_n)_{n \geq 0}$ and $(z_n)_{n \geq 0}$ that converge to a solution of the monotone inclusion problem (1).

Since they provide a deep understanding of the related discrete iterative schemes, dynamical systems assuming backward (implicit) evaluations of the governing operators have enjoyed much attention in the last years. Abbas and Attouch addressed in [1] a forward-backward dynamical system associated to the solving of (1) for A the convex subdifferential of a proper, convex and lower semicontinuous function and B a cocoercive operator, extending in this way the investigations made by Bolte in [8] on a gradient-projected dynamical system associated to the constrained minimization of a smooth convex function. The study in [1] has been further extended in [9], this time for an arbitrary maximally monotone operator A and also by utilizing variable relaxation parameters, a fact which permitted the derivation of convergence rates for the fixed point residual of the generated trajectories. Recently, in [11], the monotone inclusion problem (1) for B cocoercive has been approached in terms of a second order dynamical system of forward-backward type with variable relaxation parameters and anisotropic damping/variable damping parameters (see also [3, 4]). For more literature addressing dynamical systems of implicit type we refer the reader to [2, 5, 6, 10].

In the first part of the present manuscript we prove the existence of strong global solutions for the dynamical system (2) by making use the classical Cauchy–Lipschitz–Picard Theorem. This is followed by a convergence analysis for the generated trajectories. We show that that $x(t)$ converges weakly, as $t \rightarrow +\infty$, to a solution of the monotone inclusion problem (1) under mild assumptions. We also show that, whenever $A + B$ is strongly monotone, the trajectories converge strongly with exponential rate.

In the last part of the work we deal with the optimization problem

$$\text{minimize } f(x) + h(x),$$

where $f : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ is a proper, convex and lower semicontinuous function and $h : \mathcal{H} \rightarrow \mathbb{R}$ is a convex differentiable one with Lipschitz continuous gradient, by taking into consideration that its set of minimizers is nothing else than the solution set of the monotone inclusion problem

$$\text{find } \bar{x} \in \mathcal{H} \text{ such that } 0 \in \partial f(\bar{x}) + \nabla h(\bar{x}).$$

We provide a rate of convergence for the objective function $f + h$ along the ergodic trajectories generated by (2) (for $A = \partial f$ and $B = \nabla h$) to its minimum value.

2 Preliminaries

In this section we introduce some preliminary notions and recall some fundamental results that we will use throughout the paper. Let \mathcal{H} be a real Hilbert space. A *set-valued operator* $M : \mathcal{H} \rightrightarrows \mathcal{H}$ maps points of \mathcal{H} to subsets of \mathcal{H} . We denote by

$$\begin{aligned} \text{Dom } M &:= \{x \in \mathcal{H} \mid Mx \neq \emptyset\}, \\ \text{Ran } M &:= \{y \in \mathcal{H} \mid \exists x \in \mathcal{H} : y \in Mx\}, \\ \text{Graph } M &:= \{(x, y) \in \mathcal{H} \times \mathcal{H} \mid y \in Mx\}, \\ \text{Zer } M &:= \{x \in \mathcal{H} \mid 0 \in Mx\} \end{aligned}$$

its *domain*, *range*, *graph* and *zeros*, respectively. The *inverse operator* of M is defined by $M^{-1}y = \{x \in \mathcal{H} \mid y \in Mx\}$, the *multiplication by a scalar* $\lambda \in \mathbb{R}$ by $(\lambda M)x = \{\lambda y \mid y \in Mx\}$, and the *sum* with another operator $B : \mathcal{H} \rightrightarrows \mathcal{H}$ via Minkowski sums by $(M + N)x = \{a + b \mid a \in Mx \text{ and } b \in Nx\}$.

A set-valued operator $M : \mathcal{H} \rightrightarrows \mathcal{H}$ is called *monotone* if

$$\langle x - y, x^* - y^* \rangle \geq 0 \quad \text{for all } x, y \in \mathcal{H} \text{ and } x^* \in Mx, y^* \in My.$$

It is called *maximally monotone* if it is monotone and there is no monotone operator whose graph contains $\text{Graph } M$ properly. It is said to be *ρ -strongly monotone* with $\rho > 0$

$$\langle x - y, u - v \rangle \geq \gamma \|x - y\|^2 \quad \text{for all } x, y \in \mathcal{H} \text{ and } x^* \in Mx, y^* \in My.$$

Notice that if M is maximally monotone and strongly monotone, then $\text{Zer } M$ is a singleton, thus nonempty (see [7, Corollary 23.37]).

The *resolvent* $J_{\gamma M} = (\text{Id} + \gamma M)^{-1}$ of the maximally monotone operator γM for $\gamma > 0$ is a single-valued operator with $\text{Dom } J_{\gamma M} = \mathcal{H}$ and it is *firmly nonexpansive*, i.e.,

$$\|J_{\gamma M}x - J_{\gamma M}y\|^2 \leq \langle J_{\gamma M}x - J_{\gamma M}y, x - y \rangle \quad \text{for all } x, y \in \mathcal{H}.$$

Here, $\text{Id} : \mathcal{H} \rightarrow \mathcal{H}$ denotes the *identity operator* on \mathcal{H} . The *Yosida approximation* of a maximally monotone operator M with parameter $\gamma > 0$ is defined by $M_{\gamma} := \frac{1}{\gamma}(\text{Id} - J_{\gamma M})$. It is $\frac{1}{\gamma}$ -Lipschitz continuous, and it holds

$$x \in \text{Zer } M \iff J_{\gamma M}x = x \iff x \in \text{Zer } M_{\gamma}.$$

According to [7, Proposition 23.28] we have the relation

$$\|J_{\lambda M}x - J_{\mu M}x\| \leq |\lambda - \mu| \|M_{\lambda}x\| \quad \forall \lambda, \mu > 0 \quad \forall x \in \mathcal{H}. \quad (4)$$

Let $\beta > 0$ be arbitrary. A single-valued operator $M : \mathcal{H} \rightarrow \mathcal{H}$ is said to be β -cocoercive, if $\langle x - y, Mx - My \rangle \geq \beta \|Mx - My\|^2$ for all $(x, y) \in \mathcal{H} \times \mathcal{H}$, and $\frac{1}{\beta}$ -Lipschitz continuous, if $\|Mx - My\| \leq \frac{1}{\beta} \|x - y\|$ for all $(x, y) \in \mathcal{H} \times \mathcal{H}$. Obviously, every β -cocoercive operator is monotone and $\frac{1}{\beta}$ -Lipschitz continuous, however, the opposite implication is not true.

A function $f : \mathcal{H} \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ is said to be *proper* if it does not take the value $-\infty$ and $\text{dom } f := \{x \in \mathcal{H} \mid f(x) < +\infty\} \neq \emptyset$. It is called *convex* if

$$f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y) \quad \forall \lambda \in [0, 1] \quad \forall x, y \in \mathcal{H}.$$

The *conjugate function* $f^* : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ is defined by

$$f^*(x^*) = \sup \{ \langle x^*, x \rangle - f(x) \mid x \in \mathcal{H} \}$$

and it is convex and lower semicontinuous. If f is proper, convex and lower semicontinuous, then f^* is also proper.

The *convex subdifferential* of f is defined by

$$\partial f(x) = \{x^* \in \mathcal{H} \mid \forall y \in \mathcal{H} : f(y) \geq f(x) + \langle x^*, y - x \rangle\}$$

for $f(x) \in \mathbb{R}$ and $\partial f(x) = \emptyset$, otherwise. It is a set-valued monotone operator $\partial f : \mathcal{H} \rightrightarrows \mathcal{H}$, which is maximally monotone if f is proper, convex and lower semicontinuous.

We close this section by stating the solution concept we consider for the dynamical system (2).

Definition 1. (see for instance [6, 2]) A function $x : [0, b] \rightarrow \mathcal{H}$ (where $b > 0$) is said to be absolutely continuous if one of the following equivalent properties holds:

(i) there exists an integrable function $y : [0, b] \rightarrow \mathcal{H}$ such that

$$x(t) = x(0) + \int_0^t y(s) ds \quad \forall t \in [0, b];$$

(ii) x is continuous and its distributional derivative is Lebesgue integrable on $[0, b]$;

(iii) for every $\varepsilon > 0$, there exists $\eta > 0$ such that for any finite family of intervals $I_k = (a_k, b_k) \subseteq [0, b]$ we have the implication

$$\left(I_k \cap I_j = \emptyset \text{ and } \sum_k |b_k - a_k| < \eta \right) \implies \sum_k \|x(b_k) - x(a_k)\| < \varepsilon.$$

Remark 1. (a) It follows from the above definition that an absolutely continuous function on $[0, b]$ is differentiable almost everywhere, its derivative coincides with its distributional derivative almost everywhere and one can recover the function from its derivative $\dot{x} = y$ by the integration formula (i).

(b) If $x : [0, b] \rightarrow \mathcal{H}$ (where $b > 0$) is absolutely continuous and $M : \mathcal{H} \rightarrow \mathcal{H}$ is a γ -Lipschitz continuous operator for $\gamma > 0$, then the function $z = M \circ x$ is absolutely continuous, too. This follows from the characterization of absolute continuity given in Definition 1(iii). Moreover, z is almost everywhere differentiable and the inequality $\|\dot{z}(\cdot)\| \leq \gamma \|\dot{x}(\cdot)\|$ holds almost everywhere.

Definition 2. We say that $x : [0, +\infty) \rightarrow \mathcal{H}$ is a strong global solution of (2) if the following properties are satisfied:

(i) $x : [0, +\infty) \rightarrow \mathcal{H}$ is *locally absolutely continuous*, that is, absolutely continuous on each interval $[0, b]$ for $0 < b < +\infty$;

(ii) For almost every $t \in [0, +\infty)$ it holds $\dot{x}(t) + x(t) - z(t) - \gamma(t)Bx(t) + \gamma(t)Bz(t) = 0$, where $z(t) = J_{\gamma(t)A}(x(t) - \gamma(t)Bx(t))$;

(iii) $x(0) = x_0$.

3 Existence and uniqueness of trajectories

In this section we investigate the existence and uniqueness of the trajectories generated by the dynamical system (2). To this end we notice that the latter can be written as a non-autonomous differential equation

$$\dot{x}(t) = J_{\gamma(t)A}(x(t) - \gamma(t)Bx(t)) - x(t) + \gamma(t)Bx(t) - \gamma(t)(B \circ J_{\gamma(t)A})(x(t) - \gamma(t)Bx(t))$$

or, equivalently, as

$$\dot{x}(t) = f(\gamma(t), x(t)),$$

with $f : (0, +\infty) \times \mathcal{H} \rightarrow \mathbb{R}$,

$$\begin{aligned} f(\gamma, x) &:= J_{\gamma A}(x - \gamma Bx) - x + \gamma Bx - \gamma(B \circ J_{\gamma A})(x - \gamma Bx) \\ &= ((\text{Id} - \gamma B) \circ J_{\gamma A} \circ (\text{Id} - \gamma B) - (\text{Id} - \gamma B))x. \end{aligned}$$

Lemma 1. *Let $x \in \mathcal{H}$ be fixed. Then the function $\gamma \mapsto f(\gamma, x)$ is continuous on $(0, +\infty)$. Moreover, if $x \in \text{Dom } A$,*

$$\lim_{\gamma \downarrow 0} f(\gamma, x) = 0.$$

Proof. The first statement is a direct consequence of (4). Let $x \in \text{Dom } A$. By nonexpansiveness of $J_{\gamma A}$ we have

$$\|J_{\gamma A} \circ (\text{Id} - \gamma B)x - J_{\gamma A}x\| \leq \gamma \|Bx\| \quad \forall \gamma > 0.$$

On the other hand, $J_{\gamma A}x \rightarrow \text{Proj}_{\text{cl}(\text{Dom } A)} x = x$ as $\gamma \rightarrow 0$ by [13, Théorème 2.2], where Proj denotes the projection operator and one uses that $\text{cl}(\text{Dom } A)$ is a convex and closed set. Hence $J_{\gamma A} \circ (\text{Id} - \gamma B)x \rightarrow x$ as $\gamma \rightarrow 0$ and the assertion follows from the Lipschitz continuity of B . \square

Lemma 2. *For each $\gamma \in (0, \beta)$ and $x, y \in \mathcal{H}$ it holds*

$$\|f(\gamma, x) - f(\gamma, y)\| \leq \sqrt{6} \|x - y\|.$$

Proof. For the sake of brevity, let us write $C := \text{Id} - \gamma B$ and $J := J_{\gamma A}$. By using the firm nonexpansiveness of the resolvent and the monotonicity and Lipschitz continuity of B we get

$$\begin{aligned}
& \|f(\gamma, x) - f(\gamma, y)\|^2 \\
&= \|C \circ J \circ Cx - Cx - C \circ J \circ Cy + Cy\|^2 \\
&= \|C \circ J \circ Cx - C \circ J \circ Cy\|^2 + \|Cx - Cy\|^2 - 2 \langle C \circ J \circ Cx - C \circ J \circ Cy, Cx - Cy \rangle \\
&= \|J \circ Cx - J \circ Cy\|^2 + \gamma^2 \|B \circ J \circ Cx - B \circ J \circ Cy\|^2 \\
&\quad - 2\gamma \langle J \circ Cx - J \circ Cy, B \circ J \circ Cx - B \circ J \circ Cy \rangle \\
&\quad + \|Cx - Cy\|^2 - 2 \langle C \circ J \circ Cx - C \circ J \circ Cy, Cx - Cy \rangle \\
&\leq \left(1 + \frac{\gamma^2}{\beta^2}\right) \langle Cx - Cy, J \circ Cx - J \circ Cy \rangle \\
&\quad - 2\gamma \langle J \circ Cx - J \circ Cy, B \circ J \circ Cx - B \circ J \circ Cy \rangle \\
&\quad + \|Cx - Cy\|^2 - 2 \langle C \circ J \circ Cx - C \circ J \circ Cy, Cx - Cy \rangle \\
&= \left(1 + \frac{\gamma^2}{\beta^2} - 2\right) \langle Cx - Cy, J \circ Cx - J \circ Cy \rangle \\
&\quad - 2\gamma \langle J \circ Cx - J \circ Cy, B \circ J \circ Cx - B \circ J \circ Cy \rangle \\
&\quad + \|Cx - Cy\|^2 + 2\gamma \langle B \circ J \circ Cx - B \circ J \circ Cy, Cx - Cy \rangle \\
&\leq \|Cx - Cy\|^2 + 2\gamma \|B \circ J \circ Cx - B \circ J \circ Cy\| \|Cx - Cy\| \\
&\leq \left(1 + \frac{2\gamma}{\beta}\right) \|Cx - Cy\|^2 \\
&= \left(1 + \frac{2\gamma}{\beta}\right) (\|x - y\|^2 + \gamma^2 \|Bx - By\|^2 - 2\gamma \langle x - y, Bx - By \rangle) \\
&\leq \left(1 + \frac{2\gamma}{\beta}\right) \left(1 + \frac{\gamma^2}{\beta^2}\right) \|x - y\|^2 \\
&\leq 6 \|x - y\|^2.
\end{aligned}$$

□

Lemma 3. *There exists a constant $K > 0$ such that*

$$\|f(\gamma, x)\| \leq K(1 + \|x\|)$$

for every $\gamma \in (0, \beta)$ and $x \in \mathcal{H}$.

Proof. We fix an element $\bar{x} \in \text{Dom } A$ in the domain of A , which is evidently nonempty. According to Lemma 1 the mapping $\gamma \mapsto f(\gamma, \bar{x})$ can be continuously extended to $\gamma = 0$, therefore the image of $[0, \beta]$ under this extension is compact, hence bounded, say,

$\|f(\gamma, \bar{x})\| \leq r$ for all $\gamma \in (0, \beta)$. Furthermore, by Lemma 2 and the triangle inequality,

$$\begin{aligned} \|f(\gamma, x)\| &\leq \|f(\gamma, x) - f(\gamma, \bar{x})\| + \|f(\gamma, \bar{x})\| \\ &\leq \sqrt{6} \|x - \bar{x}\| + r \\ &\leq \sqrt{6} \|\bar{x}\| + r + \sqrt{6} \|x\|. \end{aligned} \quad \square$$

Now we can state the existence and uniqueness statement.

Theorem 1. *Let $\gamma : [0, +\infty) \rightarrow (0, \beta)$ be measurable. Then, for each $x_0 \in \mathcal{H}$, there exists a unique function $x : [0, +\infty) \rightarrow \mathcal{H}$ with $x(0) = x_0$, which is locally absolutely continuous and $\dot{x}(t) = f(\gamma(t), x(t))$ for almost every $t \in [0, +\infty)$.*

Proof. The statement follows as a consequence of the the Cauchy–Lipschitz–Picard Theorem (see [17, Proposition 6.2.1]) applied in connection with the previous two lemmas. We recall that $\gamma \mapsto f(\gamma, x)$ is continuous on $(0, +\infty)$ for each $x \in \mathcal{H}$, so $t \mapsto f(\gamma(t), x)$ is measurable, and it is bounded by Lemma 3, thus locally integrable. \square

4 Convergence analysis

In order to investigate the asymptotic properties of (2) we need some inequalities which we derive in the next subsection.

4.1 Some fundamental inequalities

Lemma 4. *If x and z are given by (2), then, for almost every $t \in [0, +\infty)$, the following statements are true:*

$$\begin{aligned} (a) \quad &\frac{x(t) - z(t)}{\gamma(t)} - Bx(t) \in Az(t); \\ (b) \quad &\frac{x(t) - z(t)}{\gamma(t)} + Bz(t) - Bx(t) = -\frac{\dot{x}(t)}{\gamma(t)} \in (A + B)z(t). \end{aligned}$$

Proof. The statement in (a) is a reformulation of the first equation in (2), while the one in (b) follows by adding $Bz(t)$ to (a) and by using the second equation in (2). \square

Lemma 5. *Let x and z be given by (2) and $\bar{x} \in \text{Zer}(A + B)$. Then, for almost every $t \in [0, +\infty)$, we have*

$$0 \leq \|x(t) - \bar{x}\|^2 - \|x(t) - z(t)\|^2 - \|z(t) - \bar{x}\|^2 + 2\gamma(t) \langle B\bar{x} - Bx(t), z(t) - \bar{x} \rangle.$$

Proof. As $-B\bar{x} \in A\bar{x}$, it holds $-\gamma(t)B\bar{x} \in \gamma(t)A\bar{x}$ for every $t \in [0, +\infty)$. By Lemma 4(a) and the monotonicity of A , for almost every $t \in [0, +\infty)$ it holds

$$\begin{aligned} 0 &\leq 2 \langle x(t) - \gamma(t)Bx(t) - z(t) + \gamma(t)B\bar{x}, z(t) - \bar{x} \rangle \\ &= \|x(t) - \bar{x}\|^2 - \|x(t) - z(t)\|^2 - \|z(t) - \bar{x}\|^2 + 2\gamma(t) \langle B\bar{x} - Bx(t), z(t) - \bar{x} \rangle. \end{aligned} \quad \square$$

Lemma 6. *Let x and z be given by (2), and let $\gamma : [0, +\infty) \rightarrow (0, \beta)$ be locally absolutely continuous. Then z is locally absolutely continuous, and*

$$\|\dot{z}(t)\| \leq \left(\sqrt{\left(1 + \frac{\dot{\gamma}(t)}{\gamma(t)}\right)^2 + \frac{(\gamma(t))^2}{\beta^2}} + \frac{\gamma(t)}{\beta} \sqrt{1 + \frac{(\gamma(t))^2}{\beta^2}} \right) \|x(t) - z(t)\|$$

for almost every $t \in [0, +\infty)$.

Proof. Let $b > 0$ be fixed. Since x, Bx and γ are absolutely continuous on $[0, b]$, the mapping $t \mapsto y(t) := x(t) - \gamma(t)Bx(t)$ is absolutely continuous on $[0, b]$. We show that $t \mapsto J_{\gamma(t)}(y(t))$ is absolutely continuous on $[0, b]$, as well.

For every $s, t \in [0, b]$, by using (4) and the nonexpansiveness of the resolvent, we get

$$\begin{aligned} \|z(t) - z(s)\| &= \|J_{\gamma(t)}(y(t)) - J_{\gamma(s)}(y(s))\| \\ &= \|J_{\gamma(t)}(y(t)) - J_{\gamma(t)}(y(s))\| + \|J_{\gamma(t)}(y(s)) - J_{\gamma(s)}(y(s))\| \\ &\leq \|y(t) - y(s)\| + |\gamma(t) - \gamma(s)| \|A_{\gamma(t)}(y(t))\|. \end{aligned}$$

Since γ is continuous on $[0, b]$, there exist $\gamma_{\min}, \gamma_{\max} \in (0, \beta)$ such that $\gamma_{\min} \leq \gamma(\cdot) \leq \gamma_{\max}$ on $[0, b]$. Using that $\gamma \mapsto \|A_{\gamma}(y(t))\|$ is nonincreasing and the Lipschitz continuity of the Yosida approximation, it yields for every $s, t \in [0, b]$

$$\begin{aligned} \|z(t) - z(s)\| &\leq \|y(t) - y(s)\| + |\gamma(t) - \gamma(s)| \|A_{\gamma(t)}(y(t))\| \\ &\leq \|y(t) - y(s)\| + |\gamma(t) - \gamma(s)| \|A_{\gamma_{\min}}(y(t))\| \\ &\leq \|y(t) - y(s)\| + |\gamma(t) - \gamma(s)| \left(\|A_{\gamma_{\min}}(0)\| + \frac{1}{\gamma_{\min}} \|y(t)\| \right). \end{aligned}$$

From here the absolute continuity of z on $[0, b]$ follows.

Applying Lemma 4 (a) for $s, t \in [0, b]$, $s \neq t$, we obtain by the monotonicity of B

$$0 \leq \left\langle z(s) - z(t), \frac{x(s) - z(s)}{\gamma(s)} - Bx(s) - \frac{x(t) - z(t)}{\gamma(t)} + Bx(t) \right\rangle,$$

which is equivalent to

$$\begin{aligned} \left\| \frac{z(s) - z(t)}{s - t} \right\|^2 &\leq \\ &\left\langle \frac{z(s) - z(t)}{s - t}, \frac{x(s) - x(t)}{s - t} + \frac{\gamma(t) - \gamma(s)}{s - t} \cdot \frac{x(t) - z(t)}{\gamma(t)} + \gamma(s) \frac{Bx(t) - Bx(s)}{s - t} \right\rangle, \end{aligned}$$

so, by the Cauchy–Schwarz inequality,

$$\left\| \frac{z(s) - z(t)}{s - t} \right\| \leq \left\| \frac{x(s) - x(t)}{s - t} + \frac{\gamma(t) - \gamma(s)}{s - t} \cdot \frac{x(t) - z(t)}{\gamma(t)} + \gamma(s) \cdot \frac{Bx(t) - Bx(s)}{s - t} \right\|.$$

By taking the limit $s \rightarrow t$, it follows that for almost every $t \in [0, +\infty)$

$$\begin{aligned} \|\dot{z}(t)\| &\leq \left\| \dot{x}(t) - \frac{\dot{\gamma}(t)}{\gamma(t)}(x(t) - z(t)) - \gamma(t) \frac{d}{dt} Bx(t) \right\| \\ &= \left\| \left(1 + \frac{\dot{\gamma}(t)}{\gamma(t)}\right)(z(t) - x(t)) + \gamma(t)(Bx(t) - Bz(t)) - \gamma(t) \frac{d}{dt} Bx(t) \right\|. \end{aligned} \quad (5)$$

According to Remark 1(b) we have $\|\frac{d}{dt} Bx(t)\| \leq \frac{1}{\beta} \|\dot{x}(t)\|$ for almost every $t \in [0, +\infty)$. Furthermore, by the monotonicity and the Lipschitz continuity of B , we have for almost every $t \in [0, +\infty)$

$$\begin{aligned} \|\dot{x}(t)\|^2 &= \|x(t) - z(t)\|^2 + (\gamma(t))^2 \|Bx(t) - Bz(t)\|^2 + 2\gamma(t) \langle x(t) - z(t), Bz(t) - Bx(t) \rangle \\ &\leq \left(1 + \frac{(\gamma(t))^2}{\beta^2}\right) \|x(t) - z(t)\|^2 \end{aligned} \quad (6)$$

as well as

$$\begin{aligned} &\left\| \left(1 + \frac{\dot{\gamma}(t)}{\gamma(t)}\right)(z(t) - x(t)) + \gamma(t)(Bx(t) - Bz(t)) \right\|^2 \\ &\leq \left(\left(1 + \frac{\dot{\gamma}(t)}{\gamma(t)}\right)^2 + \frac{(\gamma(t))^2}{\beta^2} \right) \|x(t) - z(t)\|^2 \end{aligned}$$

so, getting back to (5), we obtain

$$\|\dot{z}(t)\| \leq \left(\sqrt{\left(1 + \frac{\dot{\gamma}(t)}{\gamma(t)}\right)^2 + \frac{(\gamma(t))^2}{\beta^2}} + \frac{\gamma(t)}{\beta} \sqrt{1 + \frac{(\gamma(t))^2}{\beta^2}} \right) \|x(t) - z(t)\|. \quad \square$$

When $A + B$ is strongly monotone, we have the following strengthened version of the inequality in Lemma 5.

Lemma 7. *Let $A + B$ be ρ -strongly monotone for $\rho > 0$, x and z be given by (2) and $\bar{x} \in \text{Zer}(A + B)$. Then for almost every $t \in [0, +\infty)$ we have*

$$\begin{aligned} 0 \leq \|x(t) - \bar{x}\|^2 - \|x(t) - z(t)\|^2 - (1 + 2\rho\gamma(t)) \|z(t) - \bar{x}\|^2 \\ + 2\gamma(t) \langle Bz(t) - Bx(t), z(t) - \bar{x} \rangle. \end{aligned}$$

Proof. As $0 \in \gamma(t)(A + B)\bar{x}$, and $(A + B)$ is ρ -strongly monotone, by taking Lemma 4 (b) into consideration, we have for almost every $t \in [0, +\infty)$

$$2\rho\gamma(t) \|z(t) - \bar{x}\|^2 \leq 2 \langle x(t) - z(t) + \gamma(t)Bz(t) - \gamma(t)Bx(t), z(t) - \bar{x} \rangle$$

and the assertion follows by rearranging the terms. \square

4.2 Asymptotic properties of the trajectories

The following result, for the proof of which we refer to [2, Lemma 5.2], is the continuous counterpart of a classical result which states the convergence of quasi-Fejér monotone sequences.

Lemma 8. *If $1 \leq p < \infty$, $1 \leq r \leq \infty$, $F : [0, +\infty) \rightarrow [0, +\infty)$ is locally absolutely continuous, $F \in L^p([0, +\infty))$, $G : [0, +\infty) \rightarrow \mathbb{R}$, $G \in L^r([0, +\infty))$ and for almost every $t \in [0, +\infty)$*

$$\frac{d}{dt}F(t) \leq G(t),$$

then $\lim_{t \rightarrow +\infty} F(t) = 0$.

The next result which we recall here is the continuous version of the Opial Lemma (see, for example, [2, Lemma 5.3], [1, Lemma 1.10]).

Lemma 9. *Let $S \subseteq \mathcal{H}$ be a nonempty set and $x : [0, +\infty) \rightarrow \mathcal{H}$ a given map. Assume that*

(i) *for every $\bar{x} \in S$, $\lim_{t \rightarrow +\infty} \|x(t) - \bar{x}\|$ exists;*

(ii) *every weak sequential cluster point of the map x belongs to S .*

Then there exists $x_\infty \in S$ such that $x(t)$ converges weakly to x_∞ as $t \rightarrow +\infty$.

The following proposition will play an essential role when establishing the asymptotic properties of the trajectories generated by (2).

Lemma 10. *Let $\bar{x} \in \text{Zer}(A + B)$. Then $t \mapsto \|x(t) - \bar{x}\|$ is monotonically decreasing and $\int_0^{+\infty} \left(1 - \frac{\gamma(t)}{\beta}\right) \|x(t) - z(t)\|^2 dt < +\infty$.*

Proof. For almost every $t \in [0, +\infty)$, by using Lemma 5, the monotonicity and the Lipschitz continuity of B , we have

$$\begin{aligned} \frac{d}{dt} \|x(t) - \bar{x}\|^2 &= 2 \langle x(t) - \bar{x}, \dot{x}(t) \rangle \\ &= 2 \langle x(t) - \bar{x}, z(t) - x(t) + \gamma(t)Bx(t) - \gamma(t)Bz(t) \rangle \\ &= \|z(t) - \bar{x}\|^2 - \|x(t) - z(t)\|^2 - \|x(t) - \bar{x}\|^2 \\ &\quad + 2\gamma(t) \langle x(t) - \bar{x}, Bx(t) - Bz(t) \rangle \\ &\leq -2 \|x(t) - z(t)\|^2 + 2\gamma(t) \langle B\bar{x} - Bx(t), z(t) - \bar{x} \rangle \\ &\quad + 2\gamma(t) \langle x(t) - \bar{x}, Bx(t) - Bz(t) \rangle \\ &\leq -2 \|x(t) - z(t)\|^2 + 2\gamma(t) \langle Bz(t) - Bx(t), z(t) - \bar{x} \rangle \\ &\quad + 2\gamma(t) \langle x(t) - \bar{x}, Bx(t) - Bz(t) \rangle \\ &= -2 \|x(t) - z(t)\|^2 + 2\gamma(t) \langle Bz(t) - Bx(t), z(t) - x(t) \rangle \\ &\leq 2 \left(\frac{\gamma(t)}{\beta} - 1 \right) \|x(t) - z(t)\|^2 \leq 0, \end{aligned}$$

which shows the decreasing property. Integrating from 0 to T , for $T > 0$, yields

$$\int_0^T \left(1 - \frac{\gamma(t)}{\beta}\right) \|x(t) - z(t)\|^2 dt \leq \frac{\|x(0) - \bar{x}\|^2 - \|x(T) - \bar{x}\|^2}{2} \leq \frac{\|x(0) - \bar{x}\|^2}{2},$$

which is independent of T . \square

Theorem 2. *Let $\text{Zer}(A + B) \neq \emptyset$ and let γ be locally absolutely continuous such that, for some $\delta, \varepsilon > 0$, we have $\delta \leq \gamma(t) \leq \beta - \varepsilon$ for all $t \in [0, +\infty)$ and $\dot{\gamma} \in L^\infty([0, +\infty))$. Then the trajectories $x(t)$ and $z(t)$ generated by (2) converge weakly to an element in $\text{Zer}(A + B)$ as $t \rightarrow +\infty$.*

Proof. According to Lemma 10 we have that $t \rightarrow \|x(t) - z(t)\|^2$, mapping from $[0, +\infty)$ to $[0, +\infty)$, belongs to $L^1[0, +\infty)$. Furthermore, by the Cauchy–Schwarz inequality, the triangle inequality, (6) and Lemma 6 we have that for almost every $t \in [0, +\infty)$

$$\begin{aligned} \frac{d}{dt} \|x(t) - z(t)\|^2 &= 2 \langle x(t) - z(t), \dot{x}(t) - \dot{z}(t) \rangle \\ &\leq 2(\|\dot{x}(t)\| + \|\dot{z}(t)\|) \|x(t) - z(t)\| \\ &\leq \left(\sqrt{\left(1 + \frac{\dot{\gamma}(t)}{\gamma(t)}\right)^2 + \frac{(\gamma(t))^2}{\beta^2}} + \left(1 + \frac{\gamma(t)}{\beta}\right) \sqrt{1 + \frac{(\gamma(t))^2}{\beta^2}} \right) \|x(t) - z(t)\|^2 \\ &\leq \left(\sqrt{\left(1 + \frac{\|\dot{\gamma}\|_{L^\infty([0, +\infty))}}{\delta}\right)^2 + 1 + 2\sqrt{2}} \right) \|x(t) - z(t)\|^2 \end{aligned}$$

By Lemma 8 we have $\lim_{t \rightarrow +\infty} \|x(t) - z(t)\|^2 = 0$, which implies, via (6), that $\dot{x}(t) \rightarrow 0$ as $t \rightarrow +\infty$. Let $w \in \mathcal{H}$ be a weak sequential cluster point of $x(t)$ as $t \rightarrow +\infty$ and $(t_n)_{n \geq 0}$ be a sequence in $[0, +\infty)$ with $t_n \rightarrow +\infty$ and $x(t_n) \rightharpoonup w$ as $n \rightarrow +\infty$. Since $\lim_{t \rightarrow +\infty} (x(t) - z(t)) = 0$, we also have $z(t_n) \rightharpoonup w$ as $n \rightarrow \infty$. Furthermore, $-\frac{\dot{x}(t_n)}{\gamma(t_n)} \rightarrow 0$ as $n \rightarrow +\infty$, since $\gamma(t_n) \geq \delta$ for all $n \geq 0$.

By Lemma 4 (b) and the fact that the graph of the maximally monotone operator $A + B$ is sequentially weak-strong closed (see [7, Corollary 24.4, Proposition 20.33]), we have $(w, 0) \in \text{Graph}(A + B)$, thus $w \in \text{Zer}(A + B)$. By Lemma 10, $\|x(t) - \bar{x}\|$ converges as $t \rightarrow +\infty$. According to the Opial Lemma, $x(t)$ (and, consequently, $z(t)$) converges weakly to an element of $\text{Zer}(A + B)$ as $t \rightarrow +\infty$. \square

For the important special case of strongly monotone inclusions, we are able to show strong convergence of the trajectories to solutions without any continuity assumptions on the function γ .

Theorem 3. *Let $A + B$ be ρ -strongly monotone for $\rho > 0$. and let $\bar{x} \in \text{Zer}(A + B)$. Then we have for every $t \in [0, +\infty)$ the estimate*

$$\|x(t) - \bar{x}\|^2 \leq \frac{\|x(0) - \bar{x}\|^2}{\exp\left(\int_0^t \frac{2\rho\gamma(s)(\beta - \gamma(s))}{\beta\rho\gamma(s) + \beta - \gamma(s)} ds\right)}.$$

In particular, if

$$\int_0^{+\infty} \frac{2\rho\gamma(s)(\beta - \gamma(s))}{\beta\rho\gamma(s) + \beta - \gamma(s)} ds = +\infty,$$

then $x(t)$ converges in norm to the unique element of $\text{Zer}(A + B)$ as $t \rightarrow +\infty$.

Proof. For almost every $t \in [0, +\infty)$, by using Lemma 7, the monotonicity and the Lipschitz continuity of B , we have

$$\begin{aligned} \frac{d}{dt} \|x(t) - \bar{x}\|^2 &= 2 \langle x(t) - \bar{x}, \dot{x}(t) \rangle \\ &= 2 \langle x(t) - \bar{x}, z(t) - x(t) \rangle + 2\gamma(t) \langle x(t) - \bar{x}, Bx(t) - Bz(t) \rangle \\ &= \|z(t) - \bar{x}\|^2 - \|x(t) - z(t)\|^2 - \|x(t) - \bar{x}\|^2 \\ &\quad + 2\gamma(t) \langle x(t) - \bar{x}, Bx(t) - Bz(t) \rangle \\ &\leq -2 \|x(t) - z(t)\|^2 + 2\gamma(t) \langle x(t) - z(t), Bx(t) - Bz(t) \rangle \\ &\quad - 2\rho\gamma(t) \|z(t) - \bar{x}\|^2 \\ &\leq -2 \left(1 - \frac{\gamma(t)}{\beta}\right) \|x(t) - z(t)\|^2 - 2\rho\gamma(t) \|z(t) - \bar{x}\|^2. \end{aligned}$$

For $\alpha : [0, +\infty) \rightarrow (0, +\infty)$,

$$\alpha(t) = 1 + \frac{\beta - \gamma(t)}{\beta\rho\gamma(t)} > 1,$$

we have

$$\|z(t) - \bar{x}\|^2 \geq (1 - \alpha(t)) \|x(t) - z(t)\|^2 + \left(1 - \frac{1}{\alpha(t)}\right) \|x(t) - \bar{x}\|^2,$$

in other words,

$$-2\rho\gamma(t) \|z(t) - \bar{x}\|^2 \leq 2 \left(1 - \frac{\gamma(t)}{\beta}\right) \|x(t) - z(t)\|^2 - \frac{2\rho\gamma(t)(\beta - \gamma(t))}{\beta\rho\gamma(t) + \beta - \gamma(t)} \|x(t) - \bar{x}\|^2$$

for every $t \in [0, +\infty)$. Consequently,

$$\frac{d}{dt} \|x(t) - \bar{x}\|^2 \leq -\frac{2\rho\gamma(t)(\beta - \gamma(t))}{\beta\rho\gamma(t) + \beta - \gamma(t)} \|x(t) - \bar{x}\|^2$$

for almost every $t \in [0, +\infty)$. By Grönwall's inequality, for every $t \in [0, +\infty)$ we have

$$\|x(t) - \bar{x}\|^2 \leq \frac{\|x(0) - \bar{x}\|^2}{\exp\left(\int_0^t \frac{2\rho\gamma(s)(\beta - \gamma(s))}{\beta\rho\gamma(s) + \beta - \gamma(s)} ds\right)}. \quad \square$$

Corollary 1. *Let $A + B$ be ρ -strongly monotone for $\rho > 0$ and γ let be such that for $\delta, \varepsilon > 0$ we have $\delta \leq \gamma(t) \leq \beta - \varepsilon$ for all $t \in [0, +\infty)$. Then the trajectory $x(t)$ converges to the unique element of $\text{Zer}(A + B)$ with an exponential rate as $t \rightarrow +\infty$.*

Proof. Let \bar{x} be the unique element of $\text{Zer}(A + B)$. According to Theorem 3 it holds for every $t \in [0, +\infty)$

$$\begin{aligned} \|x(t) - \bar{x}\|^2 &\leq \frac{\|x(0) - \bar{x}\|^2}{\exp\left(\int_0^t \frac{2\rho\gamma(s)(\beta-\gamma(s))}{\beta\rho\gamma(s)+\beta-\gamma(s)} ds\right)} \leq \frac{\|x(0) - \bar{x}\|^2}{\exp\left(\int_0^t \frac{2\rho\delta\varepsilon}{\beta\rho(\beta-\varepsilon)+\beta-\delta} ds\right)} \\ &= \|x(0) - \bar{x}\|^2 e^{-\frac{2\rho\delta\varepsilon t}{\beta\rho(\beta-\varepsilon)+\beta-\delta}}, \end{aligned}$$

which leads to the desired conclusion. \square

Remark 2. Corollary 1 can be seen as the continuous-time counterpart of [12, Theorem 3.4].

4.3 Ergodic objective rate for convex minimization problems

Consider the convex minimization problem

$$\text{minimize } f(x) + h(x),$$

where $f : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ is a proper, convex and lower semicontinuous function and $h : \mathcal{H} \rightarrow \mathbb{R}$ a convex and Fréchet differentiable one with a $\frac{1}{\beta}$ -Lipschitz continuous gradient for $\beta > 0$. Since $\arg \min(f + h) = \text{Zer}(\partial f + \nabla h)$, one can approach this set by means of the trajectories of the dynamical system (2) written for $A = \partial f$ and $B = \nabla h$. We notice that, for $\eta > 0$, the resolvent of $\eta\partial f$ is given by $J_{\eta\partial f} = \text{Prox}_{\eta f}$ (see [7]), where $\text{Prox}_{\eta f} : \mathcal{H} \rightarrow \mathcal{H}$,

$$\text{Prox}_{\eta f}(x) = \arg \min_{y \in \mathcal{H}} \left\{ f(y) + \frac{1}{2\eta} \|y - x\|^2 \right\},$$

denotes the proximal point operator of ηf . Thus, the dynamical system (2) becomes

$$\begin{cases} z(t) = \text{Prox}_{\gamma(t)f}(x(t) - \gamma(t)\nabla h(x(t))), \\ 0 = \dot{x}(t) + x(t) - z(t) - \gamma(t)\nabla h(x(t)) + \gamma(t)\nabla h(z(t)) \\ x(0) = x_0. \end{cases} \quad (7)$$

In the following, we are concerned with the asymptotic behavior of the ergodic trajectory

$$\zeta(t) := \frac{1}{\Gamma(t)} \int_0^t \gamma(s)z(s) ds, \quad \text{where} \quad \Gamma(t) := \int_0^t \gamma(s) ds,$$

of z with weight γ as $t \rightarrow +\infty$.

Theorem 4. *Let $f : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ be proper, convex and lower semicontinuous, $h : \mathcal{H} \rightarrow \mathbb{R}$ be convex and Fréchet differentiable one with a $\frac{1}{\beta}$ -Lipschitz continuous gradient for $\beta > 0$, x and z be given by (7) and let $\gamma : [0, +\infty) \rightarrow (0, \beta)$ be Lebesgue measurable. Then*

$$(f + h)(\zeta(t)) \leq (f + h)(x) + \frac{\|x(0) - x\|^2}{2\Gamma(t)} \quad (8)$$

for every $x \in \mathcal{H}$ and every $t > 0$ such that $\zeta(t) \in \text{dom } f$.

Proof. Let $x \in \mathcal{H}$ and $t > 0$ be such that $\zeta(t) \in \text{dom } f$ fixed. We have, by Lemma 4 (b),

$$-\frac{\dot{x}(s)}{\gamma(s)} \in (\partial f + \nabla h)(z(s)) = \partial(f + h)(z(s))$$

and further, by the subdifferential inequality, we obtain

$$(f + h)(x) \geq (f + h)(z(s)) + \left\langle -\frac{\dot{x}(s)}{\gamma(s)}, x - z(s) \right\rangle \quad (9)$$

for almost every $s \in [0, +\infty)$. By the Lipschitz continuity of B and the Cauchy–Schwarz inequality it follows that for almost every $s \in [0, +\infty)$

$$\begin{aligned} \langle \dot{x}(s), x - z(s) \rangle &= \langle x(s) - z(s), \dot{x}(s) \rangle + \langle x - x(s), \dot{x}(s) \rangle \\ &= \langle x(s) - z(s), z(s) - x(s) + \gamma(s)Bx(s) - \gamma(s)Bz(s) \rangle + \langle x - x(s), \dot{x}(s) \rangle \\ &\leq -\|x(s) - z(s)\|^2 + \frac{\gamma(s)}{\beta} \|x(s) - z(s)\|^2 + \langle x - x(s), \dot{x}(s) \rangle \\ &\leq \langle x - x(s), \dot{x}(s) \rangle \\ &= -\frac{1}{2} \frac{d}{ds} \|x(s) - x\|^2. \end{aligned}$$

By Jensen's inequality in integral form, (9) and the previous estimate we have

$$\begin{aligned} (f + h)(\zeta(t)) &= (f + h) \left(\frac{1}{\Gamma(t)} \int_0^t \gamma(s) z(s) ds \right) \\ &\leq \frac{1}{\Gamma(t)} \int_0^t \gamma(s) (f + h)(z(s)) ds \\ &\leq \frac{1}{\Gamma(t)} \int_0^t \gamma(s) \left((f + h)(x) + \frac{1}{\gamma(s)} \langle \dot{x}(s), x - z(s) \rangle \right) ds \\ &\leq (f + h)(x) + \frac{1}{\Gamma(t)} \int_0^t -\frac{1}{2} \frac{d}{ds} \|x(s) - x\|^2 ds \\ &= (f + h)(x) + \frac{\|x(0) - x\|^2 - \|x(t) - x\|^2}{2\Gamma(t)}, \end{aligned}$$

from which the assertion follows by neglecting the nonpositive term $\frac{-\|x(t) - x\|^2}{2\Gamma(t)}$. \square

Remark 3. If \mathcal{H} is finite dimensional or $\text{dom } f$ is closed, then one can use Jensen's inequality in integral form in a less restrictive way (see for instance [18]). Under these premises inequality (8) in Theorem 4 is fulfilled for every $x \in \mathcal{H}$ and every $t > 0$.

Remark 4. The statement of Theorem 4 holds in a similar form for the discrete version given by the forward-backward-forward iterative algorithm, too. Let $x_0 \in \mathcal{H}$ be arbitrary, $(\gamma_n)_{n \geq 0} \subseteq (0, \beta)$, and $(x_n)_{n \geq 0}$ and $(z_n)_{n \geq 0}$ be the sequences generated by

$$(\forall n \geq 0) \begin{cases} z_n := \text{Prox}_{\gamma_n f}(x_n - \gamma_n \nabla h(x_n)) \\ x_{n+1} := z_n + \gamma_n (\nabla h(x_n) - \nabla h(z_n)). \end{cases}$$

Let $x \in \mathcal{H}$ and $n \geq 0$ be fixed. For any $k \geq 0$ we have

$$x_k - z_k - \gamma_k \nabla h(x_k) \in \gamma_k \partial f(z_k),$$

so

$$\frac{x_k - z_k}{\gamma_k} - \nabla h(x_k) + \nabla h(z_k) \in \partial(f + h)(z_k).$$

The subgradient inequality yields for any $k \geq 0$

$$\begin{aligned} (f + h)(z_k) &\leq (f + h)(x) - \left\langle \frac{x_k - z_k}{\gamma_k} - \nabla h(x_k) + \nabla h(z_k), x - z_k \right\rangle \\ &= (f + h)(x) - \frac{1}{\gamma_k} \langle x_k - x_{k+1}, x - z_k \rangle. \end{aligned}$$

On the other hand, for any $k \geq 0$ it holds

$$\begin{aligned} \langle x_{k+1} - x_k, x - z_k \rangle &= \frac{1}{2} \left(\|x_{k+1} - z_k\|^2 + \|x_k - x\|^2 - \|x_{k+1} - x\|^2 - \|x_k - z_k\|^2 \right) \\ &\leq \frac{1}{2} \left(\|x_k - x\|^2 - \|x_{k+1} - x\|^2 \right) - \frac{1 - \frac{\gamma_k^2}{\beta^2}}{2} \|x_k - z_k\|^2 \\ &\leq \frac{1}{2} \left(\|x_k - x\|^2 - \|x_{k+1} - x\|^2 \right), \end{aligned}$$

thus

$$(f + h)(z_k) \leq (f + h)(x) + \frac{1}{2\gamma_k} \left(\|x_k - x\|^2 - \|x_{k+1} - x\|^2 \right).$$

Setting

$$\Gamma_n := \sum_{k=0}^n \gamma_k \text{ and } \zeta_n := \frac{1}{\Gamma_n} \sum_{k=0}^n \gamma_k z_k,$$

we obtain, by Jensen's inequality in discrete form, that

$$\begin{aligned} (f + h)(\zeta_n) &\leq \frac{1}{\Gamma_n} \sum_{k=0}^n \gamma_k (f + h)(z_k) \\ &\leq (f + h)(x) + \frac{1}{2\Gamma_n} \sum_{k=0}^n \left(\|x_k - x\|^2 - \|x_{k+1} - x\|^2 \right) \\ &= (f + h)(x) + \frac{1}{2\Gamma_n} \left(\|x_0 - x\|^2 - \|x_{n+1} - x\|^2 \right) \\ &\leq (f + h)(x) + \frac{\|x_0 - x\|^2}{2\Gamma_n}. \end{aligned}$$

References

- [1] B. Abbas and H. Attouch, *Dynamical systems and forward-backward algorithms associated with the sum of a convex subdifferential and a monotone cocoercive operator*, Optimization, DOI: 10.1080/02331934.2014.971412, 2014.

- [2] B. Abbas, H. Attouch, and B. F. Svaiter, *Newton-like dynamics and forward-backward methods for structured monotone inclusions in Hilbert spaces*, Journal of Optimization Theory and Applications **161**(2), 331–360, 2014.
- [3] A.S. Antipin, *Minimization of convex functions on convex sets by means of differential equations*, (Russian) Differential’nye Uravneniya **30**(9), 1475–1486, 1994; translation in Differential Equations **30**(9), 1365–1375, 1994.
- [4] H. Attouch and F. Alvarez, *The heavy ball with friction dynamical system for convex constrained minimization problems*, in: V.H. Nguyen, J.-J. Strodiot, and P. Tossings (eds.), “Optimization (Namur, 1998)”, Lecture Notes in Economics and Mathematical Systems **481**, Springer, Berlin, 25–35, 2000.
- [5] H. Attouch, M. Marques Alves, and B.F. Svaiter, *A dynamic approach to a proximal-Newton method for monotone inclusions in Hilbert spaces, with complexity $O(1/n^2)$* , arXiv:1502.04286v1.
- [6] H. Attouch and B.F. Svaiter, *A continuous dynamical Newton-like approach to solving monotone inclusions*, SIAM Journal on Control and Optimization **49**(2), 574–598, 2011.
- [7] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, 2011.
- [8] J. Bolte, *Continuous gradient projection method in Hilbert spaces*, Journal of Optimization Theory and its Applications **119**(2), 235–259, 2003.
- [9] R.I. Boğ and E.R. Csetnek, *A dynamical system associated with the fixed points set of a nonexpansive operator*, Journal of Dynamics and Differential Equations, DOI: 10.1007/s10884-015-9438-x, 2015.
- [10] R.I. Boğ and E.R. Csetnek, *Approaching the solving of constrained variational inequalities via penalty term-based dynamical systems*, arXiv:1503.01871
- [11] R.I. Boğ and E.R. Csetnek, *Second order forward-backward dynamical systems for monotone inclusion problems*, arXiv:1503.04652, 2015.
- [12] R.I. Boğ and C. Hendrich, *Convergence analysis for a primal-dual monotone + skew splitting algorithm with applications to total variation minimization*, Journal of Mathematical Imaging and Vision **49**(3), 551–568, 2014.
- [13] H. Brezis, *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, Notas de Matemática **50**, North Holland, 1973.
- [14] H. Brezis, *Propriétés régularisantes de certains semi-groupes nonlinéaires*, Israel Journal of Mathematics **9**, 513–534, 1971.
- [15] R.E. Bruck, *Asymptotic convergence of nonlinear contraction semigroups in Hilbert spaces*, Journal of Functional Analysis **18**, 15–26, 1975.

- [16] M.G. Crandall and A. Pazy, *Semi-groups of nonlinear contractions and dissipative sets*, Journal of Functional Analysis **3**, 376–418, 1969.
- [17] A. Haraux, *Systèmes Dynamiques Dissipatifs et Applications*, Recherches en Mathématiques Appliquées, Masson, 1991.
- [18] M.D. Perlman, *Jensen's inequality for a convex vector-valued function on an infinite-dimensional space*, **4**, 52–65, 1974.
- [19] J. Peypouquet, and S. Sorin, *Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time*, Journal of Convex Analysis **17**(3&4), 1113–1163, 2010.
- [20] P. Tseng, *A modified forward-backward splitting method for maximal monotone mappings*, SIAM Journal on Control and Optimization **38**(2), 431–446, 2000.