Second order dynamical systems associated to variational inequalities

Radu Ioan Boţ *   Ernö Robert Csetnek †

February 17, 2016

Abstract. We investigate the asymptotic convergence of the trajectories generated by the second order dynamical system $\ddot{x}(t) + \gamma\dot{x}(t) + \nabla\phi(x(t)) + \beta(t)\nabla\psi(x(t)) = 0$, where $\phi, \psi : \mathcal{H} \to \mathbb{R}$ are convex and smooth functions defined on a real Hilbert space $\mathcal{H}$, $\gamma > 0$ and β is a function of time which controls the penalty term. We show weak convergence of the trajectories to a minimizer of the function ϕ over the (nonempty) set of minima of ψ as well as convergence for the objective function values along the trajectories, provided a condition expressed via the Fenchel conjugate of ψ is fulfilled. When the function ϕ is assumed to be strongly convex, we can even show strong convergence of the trajectories. The results can be seen as the second order counterparts of the ones given by Attouch and Czarnecki (Journal of Differential Equations 248(6), 1315–1344, 2010) for first order dynamical systems associated to constrained variational inequalities. At the same time we give a positive answer to an open problem posed in [11] by the same authors.

Key Words. dynamical systems, Lyapunov analysis, convex optimization, nonautonomous systems

AMS subject classification. 34G25, 47J25, 47H05, 90C25

1 Introduction

Consider the optimization problem

$$\inf_{x \in \operatorname{argmin}\psi} \phi(x), \tag{1}$$

where $\phi, \psi : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ are proper, convex and lower semicontinuous functions defined on a real Hilbert space $\mathcal{H}$ endowed with inner product $\langle\cdot,\cdot\rangle$ and associated norm $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$. If $\partial\phi + N_{\operatorname{argmin}\psi}$ is maximally monotone, then determining an optimal solution $x \in \mathcal{H}$ of (1) means nothing else than solving the subdifferential inclusion problem

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in \partial\phi(x) + N_{\operatorname{argmin}\psi}(x) \tag{2}$$

or, equivalently, solving the variational inequality

$$\text{find } x \in \mathcal{H} \text{ and } p \in \partial\phi(x) \text{ such that } \langle p, y - x \rangle \geq 0 \ \forall y \in \operatorname{argmin}\psi. \tag{3}$$

Here,

---

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

†University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, email: ernoe.robert.csetnek@univie.ac.at.

- $\partial\phi : \mathcal{H} \rightrightarrows \mathcal{H}$ is the convex subdifferential of ϕ: $\partial\phi(x) = \{p \in \mathcal{H} : \phi(y) \geq \phi(x) + \langle p, y - x\rangle \ \forall y \in \mathcal{H}\}$ for $\phi(x) \in \mathbb{R}$ and $\partial\phi(x) = \emptyset$ for $\phi(x) \notin \mathbb{R}$;

- $\operatorname{argmin}\psi$ denotes the set of minimizers of ψ;

- $N_{\operatorname{argmin}\psi}$ is the normal cone to the set $\operatorname{argmin}\psi$: $N_{\operatorname{argmin}\psi}(x) = \{p \in \mathcal{H} : \langle p, y - x\rangle \leq 0 \ \forall y \in \operatorname{argmin}\psi\}$ for $x \in \operatorname{argmin}\psi$ and $N_{\operatorname{argmin}\psi}(x) = \emptyset$ for $x \notin \operatorname{argmin}\psi$.

Attouch and Czarnecki investigated in [10] the asymptotic behavior of the trajectories of the nonautonomous first order dynamical system

$$0 \in \dot{x}(t) + \partial\phi(x(t)) + \beta(t)\partial\psi(x(t)) \tag{4}$$

to a solution of (2), where $\beta : [0, +\infty) \to (0, +\infty)$ is a function of time assumed to tend to $+\infty$ as $t \to +\infty$. Several ergodic and nonergodic convergence results have been reported in [10] under the key assumption

$$\forall p \in \operatorname{ran} N_{\operatorname{argmin}\psi} \quad \int_0^{+\infty} \beta(t)\left[\psi^*\left(\frac{p}{\beta(t)}\right) - \sigma_{\operatorname{argmin}\psi}\left(\frac{p}{\beta(t)}\right)\right] dt < +\infty, \tag{5}$$

where

- $\psi^* : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is the Fenchel conjugate of ψ: $\psi^*(p) = \sup_{x \in \mathcal{H}}\{\langle p, x\rangle - \psi(x)\} \ \forall p \in \mathcal{H}$;

- $\sigma_{\operatorname{argmin}\psi} : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is the support function of the set $\operatorname{argmin}\psi$: $\sigma_{\operatorname{argmin}\psi}(p) = \sup_{x \in \operatorname{argmin}\psi}\langle p, x\rangle$ for all $p \in \mathcal{H}$;

- $\operatorname{ran} N_{\operatorname{argmin}\psi}$ is the range of the normal cone $N_{\operatorname{argmin}\psi}$: that is $p \in \operatorname{ran} N_{\operatorname{argmin}\psi}$ if and only if there exists $x \in \operatorname{argmin}\psi$ such that $p \in N_{\operatorname{argmin}\psi}(x)$.

We mention that $N_{\operatorname{argmin}\psi} = \partial\delta_{\operatorname{argmin}\psi}$, where $\delta_{\operatorname{argmin}\psi} : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is the indicator function of $\operatorname{argmin}\psi$, which takes the value 0 on the set $\operatorname{argmin}\psi$ and $+\infty$, otherwise. Moreover, for $x \in \operatorname{argmin}\psi$ one has $p \in N_{\operatorname{argmin}\psi}(x)$ if and only if $\sigma_{\operatorname{argmin}\psi}(p) = \langle p, x\rangle$.

Let us present a situation where the above condition (5) is fulfilled. According to [10], if we take $\psi(x) = \frac{1}{2}\inf_{y \in C}\|x - y\|^2$, for a nonempty, convex and closed set $C \subseteq \mathcal{H}$, then the condition (5) is fulfilled if and only if

$$\int_0^{+\infty}\frac{1}{\beta(t)}dt < +\infty,$$

which is trivially satisfied for $\beta(t) = (1 + t)^\alpha$ with $\alpha > 1$.

The paper of Attouch and Czarnecki [10] was the starting point of a considerable number of research articles devoted to this subject, including those addressing generalizations to variational inequalities involving maximal monotone operators (see [8, 10, 12, 13, 15, 17, 20–22, 26, 27]). In the literature enumerated above, the monotone inclusions problems have been approached either through continuous dynamical systems or through their discrete counterparts formulated as splitting algorithms, both of penalty type. We speak in both cases about methods of penalty type, as the operator describing the underlying set of the addressed variational inequality is evaluated as a penalty functional. We refer also to the above-listed references for more general formulations of the key assumption (5) and for further examples for which these conditions are satisfied.

In this paper we are concerned with the asymptotic behavior of the second order dynamical system

$$\ddot{x}(t) + \gamma\dot{x}(t) + \nabla\phi(x(t)) + \beta(t)\nabla\psi(x(t)) = 0, \tag{6}$$

when the functions $\phi, \psi : \mathcal{H} \to \mathbb{R}$ are supposed to be convex and (Fréchet) differentiable with a Lipschitz continuous gradient.

We show weak convergence of the trajectories to an optimal solution to (1) as well as convergence for the objective function values along the trajectories to the optimal objective value, provided (5) is fulfilled and the function β satisfies some mild conditions. In case the function ϕ is strongly convex, the trajectories converge even strongly to the unique optimal solution of (1).

Let us also mention that the solving of (1) has been approached by means of second order dynamical system also in [15, Section 3] by penalizing the gradient of the objective function with a vanishing coefficient. Strong asymptotic convergence of the trajectories to an optimal solution has been proved under the assumption that this gradient is strongly monotone (see also [9]).

During the reviewing process of the initial version of our paper we became aware of a new preprint by Attouch and Czarnecki [11], where a second order dynamical system connected to (1) is investigated, by penalizing the objective function of the latter through a nonincreasing function. On page 11 in [11] the authors leave as an open problem the approach of (1) via the second order dynamical system (6), which this time penalizes the function describing the constraint set. The results presented in this paper provide a positive answer to this question.

2 Preliminaries

In this section we present some preliminary results and tools that will be useful throughout the paper.

The following results can be interpreted as continuous counterparts of the convergence statements for quasi-Fejér monotone sequences. For their proofs we refer the reader to [1, Lemma 5.1] and [3, Lemma 2.2 and 2.3], respectively.

Lemma 1 *Suppose that $F : [0, +\infty) \to \mathbb{R}$ is locally absolutely continuous and bounded from below and that there exists $G \in L^1([0, +\infty))$ such that for almost every $t \in [0, +\infty)$*

$$\dot{F}(t) \leq G(t).$$

Then there exists $\lim_{t\to\infty} F(t) \in \mathbb{R}$.

Lemma 2 *Suppose that $F : [0, +\infty) \to \mathbb{R}$ is locally absolutely continuous such that F is bounded from below, $\dot{F}$ is locally absolutely continuous and there exist $\gamma > 0$ and $G \in L^1([0, +\infty))$ fulfilling for almost every $t \in [0, +\infty)$*

$$\ddot{F}(t) + \gamma\dot{F}(t) \leq G(t).$$

Then there exists $\lim_{t\to\infty} F(t) \in \mathbb{R}$.

We are interested in the asymptotic analysis of the following second order nonautonomous dynamical system:

$$\begin{cases} \ddot{x}(t) + \gamma\dot{x}(t) + \nabla\phi(x(t)) + \beta(t)\nabla\psi(x(t)) = 0 \\ x(0) = u_0, \dot{x}(0) = v_0, \end{cases} \tag{7}$$

where $\mathcal{H}$ is a real Hilbert space, $\gamma > 0$, $u_0, v_0 \in \mathcal{H}$, $\beta : [0, +\infty) \to (0, +\infty)$ is a locally integrable function and the following conditions hold:

(H_ψ) $\psi : \mathcal{H} \to [0, +\infty)$ is convex, (Fréchet) differentiable with Lipschitz continuous gradient and $\operatorname{argmin}\psi = \psi^{-1}(0) \neq \emptyset$;

(H_ϕ) $\phi : \mathcal{H} \to \mathbb{R}$ is convex, (Fréchet) differentiable with Lipschitz continuous gradient, bounded from below and $S := \{z \in \operatorname{argmin}\psi : \phi(z) \leq \phi(x) \ \forall x \in \operatorname{argmin}\psi\} \neq \emptyset$.

We call $x : [0, +\infty) \to \mathcal{H}$ (strong global) solution of (7) if x and $\dot{x}$ are locally absolutely continuous (in other words, absolutely continuous on each interval $[0, b]$ for $0 < b < +\infty$), $x(0) = u_0, \dot{x}(0) = v_0$ and $\ddot{x}(t) + \gamma \dot{x}(t) + \nabla \phi(x(t)) + \beta(t) \nabla \psi(x(t)) = 0$ for almost every $t \in [0, +\infty)$.

Due to the Lipschitz continuity of $\nabla \psi$ and $\nabla \phi$ assumed in (H_ψ) and (H_ϕ), respectively, the existence and uniqueness of (strong global) solutions of (7) is a consequence of the Cauchy-Lipschitz-Picard Theorem (see for example [5, 15, 18, 24]).

In the following sections we provide sufficient conditions which guarantee both ergodic and nonergodic convergence of the trajectory of (7) to an optimal solution of (1).

Remark 3 (a) In case $\psi = 0$ the dynamical system (7) becomes the second order gradient system associated to the the heavy ball method (see also [2, 7]):

$$\begin{cases} \ddot{x}(t) + \gamma \dot{x}(t) + \nabla \phi(x(t)) = 0 \\ x(0) = u_0, \dot{x}(0) = v_0. \end{cases} \tag{8}$$

In this case $\operatorname{argmin} \psi = \mathcal{H}$ and (1) is nothing else than the minimization of the function ϕ over $\mathcal{H}$.

(b) It is well known that time discretization of second order dynamical systems leads to iterative algorithms involving inertial terms, which basically means that the new iterate is defined by making use of the previous two iterates (see for example [3, 4]). We refer the reader to [23] where an inertial algorithm of penalty type has been proposed and investigated from the point of view of its convergence properties.

3 Convergence of the trajectories and of the objective function values

In this section we show that we have weak convergence for the trajectory generated by the dynamical system (7) to an optimal solution of (1) as well as convergence for the objective function values along the trajectory. To this end, we make use of the following supplementary assumptions:

(H_β) $\beta : [0, +\infty) \to (0, +\infty)$ is a C^1-function and it satisfies the growth condition $0 \le \dot{\beta} \le k\beta$, where $0 \le k < \gamma$;

(H) $\forall p \in \operatorname{ran} N_{\operatorname{argmin} \psi}$ $\int_0^{+\infty} \beta(t) \left[\psi^* \left(\frac{p}{\beta(t)} \right) - \sigma_{\operatorname{argmin} \psi} \left(\frac{p}{\beta(t)} \right) \right] dt < +\infty.$

Remark 4 (a) Under (H_ψ), due to $\psi \le \delta_{\operatorname{argmin} \psi}$, we have

$$\psi^* \ge \delta^*_{\operatorname{argmin} \psi} = \sigma_{\operatorname{argmin} \psi}.$$

(b) When $\psi = 0$ (see Remark 3(a)), it holds $N_{\operatorname{argmin} \psi}(x) = \{0\}$ for every $x \in \operatorname{argmin} \psi = \mathcal{H}$, $\psi^* = \sigma_{\operatorname{argmin} \psi} = \delta_{\{0\}}$ and (H) trivially holds.

We start with the following technical result.

Lemma 5 *Assume that (H_ψ), (H_ϕ), (H_β) and (H) hold and let $x : [0, +\infty) \to \mathcal{H}$ be the trajectory generated by the dynamical system (7). Then for every $z \in S$ the following statements are true:*

(i) x and $\dot{x}$ are bounded;

4

(ii) $\int_0^{+\infty} \beta(t)\psi(x(t))dt < +\infty;$

(iii) $\exists \lim_{t \to +\infty} \int_0^t \langle \nabla\phi(z), x(s) - z \rangle ds \in \mathbb{R};$

(iv) $\exists \lim_{t \to +\infty} \int_0^t \left(\phi(x(s)) - \phi(z) + \left(1 - \frac{k}{\gamma}\right) \beta(s)\psi(x(s)) \right) ds \in \mathbb{R};$

(v) $\dot{x} \in L^2([0, +\infty); \mathcal{H});$

(vi) $\exists \lim_{t \to +\infty} \|x(t) - z\| \in \mathbb{R}.$

Proof. Take an arbitrary $z \in S$ and define

$$h_z(t) = \frac{1}{2}\|x(t) - z\|^2$$

and the energy functional

$$E(t) = \frac{1}{2}\|\dot{x}(t)\|^2 + \phi(x(t)) + \beta(t)\psi(x(t)). \tag{9}$$

It immediately follows that for almost every $t \in [0, +\infty)$

$$\dot{h}_z(t) = \langle \dot{x}(t), x(t) - z \rangle, \tag{10}$$

$$\ddot{h}_z(t) = \|\dot{x}(t)\|^2 + \langle \ddot{x}(t), x(t) - z \rangle, \tag{11}$$

hence, due to (7)

$$\ddot{h}_z(t) + \gamma\dot{h}_z(t) = \|\dot{x}(t)\|^2 + \langle -\nabla\phi(x(t)) - \beta(t)\nabla\psi(x(t)), x(t) - z \rangle. \tag{12}$$

Further, from (7) we obtain for almost every $t \in [0, +\infty)$

$$\begin{aligned}
\dot{E}(t) &= \langle \ddot{x}(t), \dot{x}(t) \rangle + \langle \dot{x}(t), \nabla\phi(x(t)) \rangle + \beta(t)\langle \dot{x}(t), \nabla\psi(x(t)) \rangle + \dot{\beta}(t)\psi(x(t)) \\
&= \langle \dot{x}(t), \ddot{x}(t) + \nabla\phi(x(t)) + \beta(t)\nabla\psi(x(t)) \rangle + \dot{\beta}(t)\psi(x(t)) \\
&= -\gamma\|\dot{x}(t)\|^2 + \dot{\beta}(t)\psi(x(t)).
\end{aligned} \tag{13}$$

By combining (12) and (13) we derive for almost every $t \in [0, +\infty)$

$$\ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) - \frac{\dot{\beta}(t)}{\gamma}\psi(x(t)) = \langle -\nabla\phi(x(t)) - \beta(t)\nabla\psi(x(t)), x(t) - z \rangle. \tag{14}$$

The convexity of the functions ϕ and ψ, the fact that $z \in \operatorname{argmin}\psi$ (hence $\psi(z) = 0$) and the non-negativity of β and ψ yield

$$\langle -\nabla\phi(x(t)) - \beta(t)\nabla\psi(x(t)), x(t) - z \rangle \leq \phi(z) - \phi(x(t)) - \beta(t)\psi(x(t)).$$

This inequality in combination with (14) implies

$$\ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) - \frac{\dot{\beta}(t)}{\gamma}\psi(x(t)) \leq \phi(z) - \phi(x(t)) - \beta(t)\psi(x(t)) \tag{15}$$

for almost every $t \in [0, +\infty)$. By using now the growth condition for β we get

$$\ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) + \phi(x(t)) - \phi(z) + \widetilde{\beta}(t)\psi(x(t)) \leq 0 \text{ for almost every } t \in [0, +\infty), \tag{16}$$

5

where

$$\widetilde{\beta}(t) := \left(1 - \frac{k}{\gamma}\right)\beta(t). \tag{17}$$

Since z is an optimal solution for (1), the first order optimality condition delivers us

$$0 \in \partial(\phi + \delta_{\operatorname{argmin}\psi})(z) = \nabla\phi(z) + N_{\operatorname{argmin}\psi}(z), \tag{18}$$

hence

$$-\nabla\phi(z) \in N_{\operatorname{argmin}\psi}(z) \subseteq \operatorname{ran} N_{\operatorname{argmin}\psi}. \tag{19}$$

By using this fact and the Young-Fenchel inequality we obtain for every $t \in [0, +\infty)$

$$
\begin{aligned}
\widetilde{\beta}(t)\psi(x(t)) + \langle \nabla\phi(z), x(t) - z \rangle &= \widetilde{\beta}(t)\left(\psi(x(t)) + \left\langle \frac{\nabla\phi(z)}{\widetilde{\beta}(t)}, x(t) - z \right\rangle\right) \\
&= \widetilde{\beta}(t)\left(\psi(x(t)) - \left\langle \frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}, x(t) \right\rangle + \sigma_{\operatorname{argmin}\psi}\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right)\right) \\
&\geq \widetilde{\beta}(t)\left(-\psi^*\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right) + \sigma_{\operatorname{argmin}\psi}\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right)\right). \tag{20}
\end{aligned}
$$

Furthermore,

$$\phi(x(t)) - \phi(z) \geq \langle \nabla\phi(z), x(t) - z \rangle \ \forall t \in [0, +\infty). \tag{21}$$

Finally, from (16), (20) and (21) we obtain for almost every $t \in [0, +\infty)$

$$
\begin{aligned}
&\ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) + \widetilde{\beta}(t)\left(-\psi^*\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right) + \sigma_{\operatorname{argmin}\psi}\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right)\right) \\
&\leq \ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) + \widetilde{\beta}(t)\psi(x(t)) + \langle \nabla\phi(z), x(t) - z \rangle \\
&\leq \ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) + \phi(x(t)) - \phi(z) + \widetilde{\beta}(t)\psi(x(t)) \\
&\leq 0. \tag{22}
\end{aligned}
$$

Thus

$$\ddot{h}_z(t) + \gamma\dot{h}_z(t) \leq -\frac{1}{\gamma}\dot{E}(t) + \widetilde{\beta}(t)\left(\psi^*\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right) - \sigma_{\operatorname{argmin}\psi}\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right)\right) \tag{23}$$

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

(i) From (22) we have for almost every $t \in [0, +\infty)$

$$\ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) \leq \widetilde{\beta}(t)\left(\psi^*\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right) - \sigma_{\operatorname{argmin}\psi}\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right)\right).$$

By integrating this inequality from 0 to T ($T > 0$) and by taking into account (H), (9) and the fact that ϕ and ψ are bounded from below, it yields that there exists $M > 0$ such that

$$\dot{h}_z(T) + \gamma h_z(T) + \frac{1}{2\gamma}\|\dot{x}(T)\|^2 \leq M \ \forall T \geq 0. \tag{24}$$

We derive

$$\dot{h}_z(T) + \gamma h_z(T) \leq M \ \forall T \geq 0.$$

By multiplying this inequality with $\exp(\gamma T)$ and then integrating from 0 to s, where $s > 0$, one easily obtains

$$h_z(s) \leq h_z(0) \exp(-\gamma s) + \frac{M}{\gamma}(1 - \exp(-\gamma s)) \ \forall s > 0.$$

Thus h_z is bounded, hence x is bounded.

Further, (24) implies

$$\dot{h}_z(T) + \frac{1}{2\gamma}\|\dot{x}(T)\|^2 \leq M \ \forall T \geq 0.$$

Since $t \mapsto x(t)$ is bounded, we deduce from the above inequality and (10) that $\dot{x}$ is bounded as well.

(ii) Let $F : [0, +\infty) \to \mathbb{R}$ be defined by

$$F(t) = \int_0^t \left(-\widetilde{\beta}(s)\psi(x(s)) - \langle \nabla\phi(z), x(s) - z\rangle\right) ds \ \forall t \geq 0.$$

Making again use of (22) we have for almost every $t \in [0, +\infty)$

$$\ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) \leq -\widetilde{\beta}(t)\psi(x(t)) - \langle\nabla\phi(z), x(t) - z\rangle.$$

Since h_z and E are bounded from below and $\dot{h}_z$ is bounded, by integrating the last inequality we derive that F is bounded from below. Moreover, from (20) we have for almost every $t \in [0, +\infty)$

$$\dot{F}(t) \leq \widetilde{\beta}(t)\left(\psi^*\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right) - \sigma_{\operatorname{argmin}\psi}\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right)\right).$$

Since according to (H), the function on the right-hand side of this inequality is L^1-integrable on $[0, +\infty)$, a direct application of Lemma 1 yields that $\lim_{t\to+\infty} F(t)$ exists and is a real number. Hence

$$\exists \lim_{t\to+\infty} \int_0^t \left(\widetilde{\beta}(s)\psi(s) + \langle\nabla\phi(z), x(s) - z\rangle\right) ds \in \mathbb{R} \tag{25}$$

Since $\psi \geq 0$, we obtain for every $t \in [0, +\infty)$

$$\widetilde{\beta}(t)\psi(x(t)) + \langle\nabla\phi(z), x(t) - z\rangle \geq \frac{\widetilde{\beta}(t)}{2}\psi(x(t)) + \langle\nabla\phi(z), x(t) - z\rangle$$

and from here, similarly to (20),

$$\frac{\widetilde{\beta}(t)}{2}\psi(x(t)) + \langle\nabla\phi(z), x(t) - z\rangle \geq \frac{\widetilde{\beta}(t)}{2}\left(-\psi^*\left(\frac{-2\nabla\phi(z)}{\widetilde{\beta}(t)}\right) + \sigma_{\operatorname{argmin}\psi}\left(\frac{-2\nabla\phi(z)}{\widetilde{\beta}(t)}\right)\right).$$

Thus, for almost every $t \in [0, +\infty)$ it holds

$$\ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) + \frac{\widetilde{\beta}(t)}{2}\left(-\psi^*\left(\frac{-2\nabla\phi(z)}{\widetilde{\beta}(t)}\right) + \sigma_{\operatorname{argmin}\psi}\left(\frac{-2\nabla\phi(z)}{\widetilde{\beta}(t)}\right)\right)$$

$$\leq \ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) + \frac{\widetilde{\beta}(t)}{2}\psi(x(t)) + \langle\nabla\phi(z), x(t) - z\rangle$$

$$\leq \ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) + \widetilde{\beta}(t)\psi(x(t)) + \langle\nabla\phi(z), x(t) - z\rangle$$

$$\leq 0.$$

By using the same arguments as in the proof of (25) it yields that

$$\exists \lim_{t\to+\infty} \int_0^t \left(\frac{\widetilde{\beta}(s)}{2} \psi(s) + \langle \nabla\phi(z), x(s) - z \rangle \right) ds \in \mathbb{R}. \tag{26}$$

Finally, from (25), (26) and (17) we obtain (ii).

(iii) Follows from (ii), (25) and (17).

(iv) Follows from (22) and (20), by using the same arguments as for proving statement (25).

(v) From (13) and (H_β) we derive for almost every $t \geq 0$

$$\dot{E}(t) + \gamma\|\dot{x}(t)\|^2 \leq k\beta(t)\psi(x(t)). \tag{27}$$

Since E is bounded from below, integrating the last inequality and taking into account (ii) we conclude that $\dot{x} \in L^2([0, +\infty); \mathcal{H})$.

(vi) Combining (23) with (13) we obtain

$$\ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{\dot{\beta}(t)}{\gamma}\psi(x(t)) \leq \|\dot{x}(t)\|^2 + \widetilde{\beta}(t)\left(\psi^*\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right) - \sigma_{\operatorname{argmin}\psi}\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right)\right) \tag{28}$$

for almost every $t \in [0, +\infty)$. Notice that $\dot{\beta}(t) \geq 0$, hence

$$\ddot{h}_z(t) + \gamma\dot{h}_z(t) \leq \|\dot{x}(t)\|^2 + \widetilde{\beta}(t)\left(\psi^*\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right) - \sigma_{\operatorname{argmin}\psi}\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right)\right) \tag{29}$$

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

According to (v), (H), (19), (17) and the fact that $\operatorname{ran} N_{\operatorname{argmin}\psi}$ is a cone, the function on the right-hand side of the above inequality is L^1-integrable on $[0, +\infty)$. Applying now Lemma 2 we deduce that (i) holds. ∎

Remark 6 The assumption $\dot{\beta} \geq 0$ has been used in the above proof only for showing statement (vi).

For the main result of the paper concerning the weak asymptotic convergence of the trajectory generated by the dynamical system (7), that we formulate and prove afterwards, we will make use of the continuous version of the Opial Lemma.

Lemma 7 *Let S be a nonempty subset of the real Hilbert space $\mathcal{H}$ and $x : [0, +\infty) \to \mathcal{H}$ a given function. Assume that*

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

(ii) *every weak limit point of x belongs to S.*

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

We can state now the main theorem of the paper.

Theorem 8 *Assume that (H_ψ), (H_ϕ), (H_β) and (H) hold under the supplementary condition $\lim_{t\to+\infty} \beta(t) = +\infty$ and let $x : [0, +\infty) \to \mathcal{H}$ be the trajectory generated by the dynamical system (7). Then the following statements are true:*

(i) *$\phi(x(t))$ converges to the optimal objective value of (1) as $t \to +\infty$;*

(ii) *$\lim_{t\to+\infty} \beta(t)\psi(x(t)) = \lim_{t\to+\infty} \psi(x(t)) = 0$;*

8

(iii) $\int_0^{+\infty} \beta(t)\psi(x(t))dt < +\infty;$

(iv) $\int_0^{+\infty} \|\dot{x}(t)\|^2 dt < +\infty;$

(v) $\lim_{t\to+\infty} \dot{x}(t) = 0;$

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

Proof. Consider again the energy functional defined in (9). Taking into account that E is bounded from below, from (27), Lemma 5(ii) and Lemma 1 we derive that

$$\exists \lim_{t\to+\infty} E(t) \in \mathbb{R}. \tag{30}$$

We claim that

$$\lim_{t\to+\infty} \psi(x(t)) = 0. \tag{31}$$

Indeed, notice that

$$\frac{E(t)}{\beta(t)} = \frac{\|\dot{x}(t)\|^2}{2\beta(t)} + \frac{\phi(x(t))}{\beta(t)} + \psi(x(t)). \tag{32}$$

Let $z \in S$ be arbitrary. Since ϕ is bounded from below and by taking into account the inequality

$$\phi(x(t)) \le \phi(z) + \langle \nabla\phi(x(t)), x(t) - z \rangle \forall t \in [0, +\infty),$$

that x is bounded, $\nabla\phi$ is Lipschitz continuous and $\lim_{t\to+\infty} \beta(t) = +\infty$, we can easily conclude that $\lim_{t\to+\infty} \frac{\phi(x(t))}{\beta(t)} = 0$. Furthermore, since $\dot{x}$ is bounded, by using (30), we easily derive (31) by passing to the limit as $t \to +\infty$ in (32).

Now we invoke [10, Lemma 3.4] for the relation

$$\liminf_{t\to+\infty} \phi(x(t)) \ge \phi(z). \tag{33}$$

Indeed, this was stated in [10, Lemma 3.4] for the first order companion of the dynamical system (7). The result remains true for (7), too, since a careful look at the proof of Lemma 3.4 in [10] reveals that ingredients which lead to the conclusion are the statements in Lemma 5(iii), the fact that the trajectory x is bounded, the weak lower semicontinuity of ψ, equation (31), the inequality (21) and the relation (19).

Since $E(t) \ge \phi(x(t))$, from (33) we have $\lim_{t\to+\infty} E(t) \ge \phi(z)$. We claim that

$$\lim_{t\to+\infty} E(t) = \phi(z). \tag{34}$$

Let us assume that $\lim_{t\to+\infty} E(t) > \phi(z)$. Then there exist $\theta > 0$ and $t_0 \ge 0$ such that for every $t \ge t_0$ we have

$$\phi(x(t)) + \beta(t)\psi(x(t)) + \frac{1}{2}\|\dot{x}(t)\|^2 > \phi(z) + \theta. \tag{35}$$

Hence for every $t \ge t_0$

$$\theta < \phi(x(t)) - \phi(z) + \left(1 - \frac{k}{\gamma}\right)\beta(t)\psi(x(t)) + \frac{k}{\gamma}\beta(t)\psi(x(t)) + \frac{1}{2}\|\dot{x}(t)\|^2. \tag{36}$$

Integrating the last inequality and taking into account Lemma 5(ii), (iv) and (v) we obtain a contradiction. Hence (34) holds.

Since $\phi(x(t)) \le E(t)$ we derive from (34) that $\limsup_{t\to+\infty} \phi(x(t)) \le \phi(z)$, which combined with (33) imply (i).

Further, notice that due to $\lim_{t\to+\infty} E(t) = \lim_{t\to+\infty} \phi(x(t)) = \phi(z)$, from (9) we obtain

$$\lim_{t\to+\infty} \left(\beta(t)\psi(x(t)) + \frac{1}{2}\|\dot{x}(t)\|^2 \right) = 0,$$

hence

$$\lim_{t\to+\infty} \beta(t)\psi(x(t)) = 0 \tag{37}$$

and

$$\lim_{t\to+\infty} \|\dot{x}(t)\| = 0.$$

Thus (ii) and (v) hold. The statements (iii) and (iv) have been proved in Lemma 5.

Statement (vi) will be a consequence of the Opial Lemma. Since according to Lemma 5(vi), the first condition in the Opial Lemma is fulfilled, it remains to show that every weak limit point of $x(t)$ as $t \to +\infty$ belongs to S.

Let $(t_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers such that $\lim_{n\to+\infty} t_n = +\infty$ and $x(t_n)$ converges weakly to x_∞ as $n \to +\infty$. By using the weak lower semicontinuity of ψ and (ii) we obtain

$$0 \le \psi(x_\infty) \le \liminf_{n\to+\infty} \psi(x(t_n)) = 0,$$

hence $x_\infty \in \operatorname{argmin}\psi$. Moreover, the weak lower semicontinuity of ϕ and (i) yield

$$\phi(x_\infty) \le \liminf_{n\to+\infty} \phi(x(t_n)) = \phi(z),$$

thus $x_\infty \in S$. ∎

We show in the following that if the objective function of (1) is strongly convex, then the trajectory generated by (7) converges strongly to the unique optimal solution of (1).

Theorem 9 *Assume that (H_ψ), (H_ϕ), (H_β) and (H) hold and let $x : [0,+\infty) \to \mathcal{H}$ be the trajectory generated by the dynamical system (7). If ϕ is strongly convex, then $x(t)$ converges strongly to the unique optimal solution of (1) as $t \to +\infty$.*

Proof. Let $\gamma > 0$ be such that ϕ is γ-strongly convex. It is a well-known fact that in case the optimization problem (1) has a unique optimal solution, which we denote by z.

By combining (20) with the stronger inequality

$$\phi(x(t)) - \phi(z) \ge \langle \nabla\phi(z), x(t) - z \rangle + \frac{\gamma}{2}\|x(t) - z\|^2 \ \forall t \in [0,+\infty), \tag{38}$$

we obtain this time (see the proof of Lemma 5) for almost every $t \in [0,+\infty)$

$$\ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{\gamma}{2}\|x(t) - z\|^2 + \frac{1}{\gamma}\dot{E}(t) + \widetilde{\beta}(t)\left(-\psi^*\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right) + \sigma_{\operatorname{argmin}\psi}\left(\frac{-\nabla\phi(z)}{\widetilde{\beta}(t)}\right)\right)$$

$$\le \ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{\gamma}{2}\|x(t) - z\|^2 + \frac{1}{\gamma}\dot{E}(t) + \widetilde{\beta}(t)\psi(x(t)) + \langle \nabla\phi(z), x(t) - z \rangle$$

$$\le \ddot{h}_z(t) + \gamma\dot{h}_z(t) + \frac{1}{\gamma}\dot{E}(t) + \phi(x(t)) - \phi(z) + \widetilde{\beta}(t)\psi(x(t))$$

$$\le 0.$$

Taking into account (H) and that E and h_z are bounded from below, by integration of the above inequality we obtain that there exists a constant $L \in \mathbb{R}$ such that

$$\dot{h}_z(T) + \frac{\gamma}{2}\int_0^T \|x(t) - z\|^2 dt \le L \ \forall T > 0.$$

10

Notice that due to Lemma 5(i) the function $T \mapsto \dot{h}_z(T)$ is bounded, hence

$$\int_0^{+\infty} \|x(t) - z\|^2 dt < +\infty.$$

Since according to Lemma 5(vi) $\lim_{t \to +\infty} \|x(t) - z\|$ exists, we conclude that $\|x(t) - z\|$ converges to 0 as $t \to +\infty$ and the proof is complete. ∎

Remark 10 The results presented in this paper remain true even if the assumed growth condition is satisfied starting with a $t_0 \geq 0$, that is, if there exists $t_0 \geq 0$ such that

$$0 \leq \dot{\beta}(t) \leq k\beta(t) \ \forall t \geq t_0.$$

References

[1] B. Abbas, H. Attouch, B.F. Svaiter, *Newton-like dynamics and forward-backward methods for structured monotone inclusions in Hilbert spaces*, Journal of Optimization Theory and its Applications 161(2), 331–360, 2014

[2] F. Alvarez, *On the minimizing property of a second order dissipative system in Hilbert spaces*, SIAM Journal on Control and Optimization 38(4), 1102–1119, 2000

[3] F. Alvarez, *Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space*, SIAM Journal on Optimization 14(3), 773–782, 2004

[4] F. Alvarez, H. Attouch, *An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping*, Set-Valued Analysis 9, 3–11, 2001

[5] F. Alvarez, H. Attouch, J. Bolte, P. Redont, *A second-order gradient-like dissipative dynamical system with Hessian-driven damping. Application to optimization and mechanics*, Journal de Mathématiques Pures et Appliquées 81(8), 747–779, 2002

[6] A.S. Antipin, *Minimization of convex functions on convex sets by means of differential equations*, (Russian) Differentsial'nye Uravneniya 30(9), 1475–1486, 1994; translation in Differential Equations 30(9), 1365–1375, 1994

[7] H. Attouch, F. Alvarez, *The heavy ball with friction dynamical system for convex constrained minimization problems*, in: Optimization (Namur, 1998), 25–35, in: Lecture Notes in Economics and Mathematical Systems 481, Springer, Berlin, 2000

[8] H. Attouch, A. Cabot, M.-O. Czarnecki, *Asymptotic behavior of nonautonomous monotone and subgradient evolution equations*, arXiv:1601.00767, 2016

[9] H. Attouch, M.-O. Czarnecki, *Asymptotic control and stabilization of nonlinear oscillators with non-isolated equilibria*, Journal of Differential Equations 179(1), 278–310, 2002

[10] H. Attouch, M.-O. Czarnecki, *Asymptotic behavior of coupled dynamical systems with multiscale aspects*, Journal of Differential Equations 248(6), 1315–1344, 2010

[11] H. Attouch, M.-O. Czarnecki, *Asymptotic behavior of gradient-like dynamical systems involving inertia and multiscale aspects*, arXiv:1602.00232, 2016

[12] H. Attouch, M.-O. Czarnecki, J. Peypouquet, *Prox-penalization and splitting methods for constrained variational problems*, SIAM Journal on Optimization 21(1), 149–173, 2011

[13] H. Attouch, M.-O. Czarnecki, J. Peypouquet, *Coupling forward-backward with penalty schemes and parallel splitting for constrained variational inequalities*, SIAM Journal on Optimization 21(4), 1251–1274, 2011

[14] H. Attouch, X. Goudou, P. Redont, *The heavy ball with friction method. I. The continuous dynamical system: global exploration of the local minima of a real-valued function by asymptotic analysis of a dissipative dynamical system*, Communications in Contemporary Mathematics 2(1), 1–34, 2000

[15] H. Attouch, P.-E. Maingé, *Asymptotic behavior of second-order dissipative evolution equations combining potential with non-potential effects*, ESAIM. Control, Optimisation and Calculus of Variations 17(3), 836–857, 2011

[16] H.H. Bauschke, P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, New York, 2011

[17] S. Banert, R.I. Boţ, *Backward penalty schemes for monotone inclusion problems*, Journal of Optimization Theory and Applications 166(3), 930–948, 2015

[18] R.I. Boţ, E.R. Csetnek, *Second order forward-backward dynamical systems for monotone inclusion problems*, arXiv:1503.04652, 2015

[19] R.I. Boţ, E.R. Csetnek, *A dynamical system associated with the fixed points set of a non-expansive operator*, Journal of Dynamics and Differential Equations, DOI: 10.1007/s10884-015-9438-x, 2015

[20] R.I. Boţ, E.R. Csetnek, *Approaching the solving of constrained variational inequalities via penalty term-based dynamical systems*, Journal of Mathematical Analysis and Applications 435(2), 1688-1700, 2016

[21] R.I. Boţ, E.R. Csetnek, *Forward-backward and Tseng's type penalty schemes for monotone inclusion problems*, Set-Valued and Variational Analysis 22, 313–331, 2014

[22] R.I. Boţ, E.R. Csetnek, *A Tseng's type penalty scheme for solving inclusion problems involving linearly composed and parallel-sum type monotone operators*, Vietnam Journal of Mathematics 42(4), 451–465, 2014

[23] R.I. Boţ, E.R. Csetnek, *Penalty schemes with inertial effects for monotone inclusion problems*, arXiv:1512.04428, 2015

[24] A. Cabot, H. Engler, S. Gadat, *On the long time behavior of second order differential equations with asymptotically small dissipation*, Transactions of the American Mathematical Society 361(11), 5983–6017, 2009

[25] A. Cabot, H. Engler, S. Gadat, *Second-order differential equations with asymptotically small dissipation and piecewise flat potentials*, Proceedings of the Seventh Mississippi State-UAB Conference on Differential Equations and Computational Simulations, 33–38, Electronic Journal of Differential Equations Conference 17, 2009

[26] N. Noun, J. Peypouquet, *Forward-backward penalty scheme for constrained convex minimization without inf-compactness*, Journal of Optimization Theory and Applications, 158(3), 787–795, 2013

[27] J. Peypouquet, *Coupling the gradient method with a general exterior penalization scheme for convex minimization*, Journal of Optimizaton Theory and Applications 153(1), 123-138, 2012