Ergodic convergence results for the Arrow–Hurwicz differential system

Simon K. Niederländer

***Abstract*— In a real Hilbert space setting, we investigate the ergodic convergence properties of the solutions of the classical Arrow–Hurwicz differential system in view of solving linearly constrained convex minimization problems. Under the convexity hypothesis on the objective function of the minimization problem, we recover the fact that every solution of the Arrow–Hurwicz differential system weakly converges in average towards its asymptotic center. Moreover, it is shown that the primal-dual gap function relative to an averaged solution obeys the asymptotic estimate $\mathcal{O}(1/t)$ as $t \to +\infty$. If, in addition, the linear operator associated with the constraint function of the minimization problem is bounded from below, we find that the primal-dual gap function obeys the refined asymptotic estimate $\mathcal{O}(1/t^2)$ as $t \to +\infty$. Numerical experiments illustrate our theoretical findings.**

I. Introduction

Let X and Y be real Hilbert spaces equipped with inner products $\langle \cdot\,,\cdot \rangle_X, \langle \cdot\,,\cdot \rangle_Y$ and induced norms $\|\cdot\|_X, \|\cdot\|_Y$. Consider the minimization problem

$$\inf\{f(x) \mid Ax = b\} \tag{P}$$

with $f : X \to \mathbb{R}$ being a convex and continuously differentiable function, $A : X \to Y$ a linear and continuous operator, and $b \in Y$. Let us associate with (P) the Lagrangian

$$\begin{aligned} L : X \times Y &\longrightarrow \mathbb{R} \\ (x, \lambda) &\longmapsto f(x) + \langle \lambda, Ax - b \rangle_Y \end{aligned}$$

which, in view of the above assumptions, is a convex-concave and continuously differentiable bifunction. We recall that a pair $(\bar{x}, \bar{\lambda}) \in X \times Y$ is a saddle point of L if

$$L(\bar{x}, \lambda) \le L(\bar{x}, \bar{\lambda}) \le L(x, \bar{\lambda}) \quad \forall (x, \lambda) \in X \times Y.$$

Classically, $(\bar{x}, \bar{\lambda}) \in X \times Y$ is a saddle point of L if and only if $\bar{x} \in X$ is a minimizer of (P), $\bar{\lambda} \in Y$ is a maximizer of the Lagrange dual problem associated with (P), viz.,

$$\sup\{-f^*(-A^*\lambda) - \langle \lambda, b \rangle_Y \mid \lambda \in Y\}, \tag{D}$$

and the optimal values of (P) and (D) coincide; cf. Ekeland and Témam [2]. Here, $f^* : X \to \mathbb{R} \cup \{+\infty\}$ denotes the Fenchel conjugate of f defined by

$$f^*(u) = \sup\{\langle u, x \rangle_X - f(x) \mid x \in X\}$$

and $A^* : Y \to X$ refers to the adjoint operator of A. Equivalently, $(\bar{x}, \bar{\lambda}) \in X \times Y$ is a saddle point of L if and only if $(\bar{x}, \bar{\lambda})$ solves the system of primal-dual optimality conditions (see, e.g., Bauschke and Combettes [3])

$$\begin{cases} \nabla f(x) + A^*\lambda = 0_X \\ \qquad Ax - b = 0_Y. \end{cases}$$

Throughout, we denote by $S \times M \subset X \times Y$ the (possibly empty) set of saddle points of L. Recall that a saddle point of L exists whenever (P) admits a minimizer and, for instance, the constraint qualification

$$b \in \operatorname{sri} A(X)$$

holds[1]. Here, for a convex set $C \subset Y$, we denote by

$$\operatorname{sri} C = \Big\{x \in C \mid \bigcup_{\mu > 0} \mu(C - x) \text{ is a closed linear subspace}\Big\}$$

its strong relative interior; cf. Bauschke and Combettes [3]. In turn, problem (P) admits a minimizer whenever $b \in A(X)$ and, for instance, f is coercive, that is, $\lim_{\|x\|_X \to +\infty} f(x) = +\infty$.

In this work, we reconsider the classical Arrow–Hurwicz differential system

$$\begin{cases} \dot{x} + \nabla f(x) + A^*\lambda = 0_X \\ \qquad \dot{\lambda} + b - Ax = 0_Y \end{cases} \tag{AH}$$

relative to the convex minimization problem (P). The (AH) evolution system was originated by Arrow and Hurwicz [5] (see also Arrow et al. [6]) and is known to be intimately related to the mini-maximization of the Lagrangian L associated with (P). Indeed, in view of the above system of primal-dual optimality conditions, we readily deduce that the zeros of the operator

$$\begin{aligned} T : X \times Y &\longrightarrow X \times Y \\ (x, \lambda) &\longmapsto (\nabla f(x) + A^*\lambda, b - Ax), \end{aligned}$$

that is, the "generator" of the (AH) dynamics, are precisely the saddle points of L, i.e.,

$$T(\bar{x}, \bar{\lambda}) = (0_X, 0_Y) \iff (\bar{x}, \bar{\lambda}) \in S \times M.$$

Moreover, the operator T is maximally monotone; cf. Rockafellar [7]. As a consequence, $S \times M$ can be interpreted as the set of zeros of the maximally monotone operator T and, as such, it is a closed and convex subset of $X \times Y$; see, e.g., Peypouquet and Sorin [8].

An accompanying article appeared as [1] in *Optimization*. Simon K. Niederländer is with the research group Autonomous Systems and Control, Siemens Technology, Otto-Hahn-Ring 6, 81739 Munich, Germany. Email: `simon.niederlaender@siemens.com`

[1] We remark that, in the finite-dimensional case, the condition amounts to $b \in A(X)$ which is commonly referred to as Slater assumption; see, e.g., Hiriart-Urruty and Lemaréchal [4].

Preliminary facts. As previously emphasized by Rockafellar [9], the general theory for semi-groups of contractions generated by maximally monotone operators (see, e.g., Crandall and Pazy [10], Brézis [11]) applies to the (AH) differential system. These results, dating back to the works of Kato [12] and Kōmura [13] (see also Browder [14]), imply that the Cauchy problem associated with (AH) is well posed and that its solutions $(x, \lambda), (y, \eta) : [0, +\infty) \to X \times Y$ are "non-expansive" in the sense that

$$t \longmapsto \|x(t) - y(t)\|_X^2 + \|\lambda(t) - \eta(t)\|_Y^2$$

is non-increasing. As a direct consequence, every zero of the operator T (and thus, every saddle point of the bifunction L) is stable in the sense of Lyapunov. Moreover, the solutions of (AH) remain bounded if and only if the set $S \times M$ contains at least one element. In this case, the solutions $(x(t), \lambda(t))$ of (AH) weakly converge in average, as $t \to +\infty$, towards an element of $S \times M$ (see Baillon and Brézis [15]), i.e., there exists $(\bar{x}, \bar{\lambda}) \in S \times M$ such that

$$\frac{1}{t} \int_0^t (x(\tau), \lambda(\tau))\, \mathrm{d}\tau \rightharpoonup (\bar{x}, \bar{\lambda}) \text{ as } t \to +\infty.$$

If, on the other hand, the weak sequential cluster points of $(x(t), \lambda(t))_{\geq 0}$ belong to $S \times M$, then $(x(t), \lambda(t))$ converges weakly, as $t \to +\infty$, towards some element of $S \times M$; we refer to Brézis [11] (see also Pazy [16]) for more details on the asymptotic behavior of evolution equations governed by maximally monotone operators.

Contributions. The mini-maximizing properties of the solutions $(x, \lambda) : [0, +\infty) \to X \times Y$ of (AH) with respect to the convex minimization problem (P) and its associated Lagrange dual problem (D) are classically measured in terms of the "primal-dual gap function"

$$t \longmapsto L(x(t), \cdot\,) - L(\,\cdot\,, \lambda(t))$$

relative to the set $S \times M$. Whenever f is convex, we observe that the solutions $(x(t), \lambda(t))$ of (AH) may fail to converge as $t \to +\infty$ even if the set of saddle points of L is comprised of a single element. In consequence, it is natural to first study the average behavior of the solutions of (AH). Utilizing the notion of the Cesàro average $(\sigma, \omega) : (0, +\infty) \to X \times Y$ of a solution (x, λ) of (AH), viz.,

$$(\sigma(t), \omega(t)) = \frac{1}{t} \int_0^t (x(\tau), \lambda(\tau))\, \mathrm{d}\tau,$$

we find that the solutions of (AH) obey in average, for any $(\xi, \eta) \in S \times M$, the asymptotic estimate

$$L(\sigma(t), \eta) - L(\xi, \omega(t)) = \mathcal{O}\Big(\frac{1}{t}\Big) \text{ as } t \to +\infty.$$

In this case, the Cesàro average $(\sigma(t), \omega(t))$ of a solution of (AH) weakly converges, as $t \to +\infty$, towards a saddle point of L. This result is in line with the work by Nemirovski and Yudin [17] on the classical Arrow–Hurwicz method. If, further, the linear operator A is bounded from below, we find that the solutions of (AH) obey in average, for any $(\xi, \eta) \in S \times M$, the refined asymptotic estimate

$$L(\sigma(t), \eta) - L(\xi, \omega(t)) = \mathcal{O}\Big(\frac{1}{t^2}\Big) \text{ as } t \to +\infty.$$

Numerical experiments support our theoretical findings.

II. Basic properties

Let $X \times Y$ be endowed with the Hilbertian product structure $\langle \cdot\,, \cdot \rangle = \langle \cdot\,, \cdot \rangle_X + \langle \cdot\,, \cdot \rangle_Y$ and associated norm $\|\cdot\|$. Throughout, we presuppose that

(A1) $f : X \to \mathbb{R}$ is convex and continuously differentiable;
(A2) $\nabla f : X \to X$ is Lipschitz continuous on the bounded subsets of X;
(A3) $A : X \to Y$ is linear and continuous, and $b \in Y$.

Consider the Arrow–Hurwicz differential system

$$\begin{cases} \dot{x} + \nabla f(x) + A^* \lambda = 0_X \\ \qquad \dot{\lambda} + b - Ax = 0_Y \end{cases} \tag{AH}$$

with initial data $(x_0, \lambda_0) \in X \times Y$. As is standard, we say that $(x, \lambda) : [0, +\infty) \to X \times Y$ is a (classical) solution of (AH) if $(x, \lambda) \in \mathcal{C}^1([0, +\infty); X \times Y)$ and (x, λ) satisfies (AH) on $[0, +\infty)$ with $(x(0), \lambda(0)) = (x_0, \lambda_0)$.

The following "descent property" of the solutions of (AH) is readily obtained as a consequence of the monotonicity of the operator T, that is, for any $(x, \lambda), (y, \eta) \in X \times Y$, we have

$$\langle T(x, \lambda) - T(y, \eta), (x, \lambda) - (y, \eta) \rangle \geq 0.$$

The argument we use to prove the following result relies essentially on Aubin and Cellina [18] (see also Brézis [11]).

Theorem II.1. *For any $(x_0, \lambda_0) \in X \times Y$ there exists a unique solution $(x, \lambda) : [0, +\infty) \to X \times Y$ of* (AH). *Moreover,*

(i) $t \mapsto \|(\dot{x}(t), \dot{\lambda}(t))\|$ *is non-increasing and*

$$\|(\dot{x}(t), \dot{\lambda}(t))\| \leq \|T(x_0, \lambda_0)\| \quad \forall t \geq 0;$$

(ii) $\lim_{t \to +\infty} \|(\dot{x}(t), \dot{\lambda}(t))\|$ *exists;*
(iii) $(\dot{x}, \dot{\lambda}) \in \mathcal{L}^\infty([0, +\infty); X \times Y)$.

Remark II.2. We remark that the assertions of Theorem II.1 remain valid even in the case when $f : X \to \mathbb{R} \cup \{+\infty\}$ is a proper convex lower semi-continuous function. Under this assumption, the (AH) dynamics generalize to the evolution system

$$\begin{cases} \dot{x} + \partial f(x) + A^* \lambda \ni 0_X \\ \qquad \dot{\lambda} + b - Ax = 0_Y \end{cases}$$

with ∂f denoting the convex subdifferential of f. The existence and uniqueness of the (strong) solutions of the above evolution system then follow from the general theory for semi-groups of contractions generated by maximally monotone operators; we refer to Brézis [11] (see also Pazy [16]) for a detailed study of the subject.

In the following, let $S \times M$ denote the (possibly empty) set of saddle points of the Lagrangian

$$\begin{aligned} L : X \times Y &\longrightarrow \mathbb{R} \\ (x, \lambda) &\longmapsto f(x) + \langle \lambda, Ax - b \rangle_Y \end{aligned}$$

associated with the convex minimization problem (P). Given the convexity of f, an immediate derivation reveals that for any $(x, \lambda), (y, \eta) \in X \times Y$, we have

$$\langle T(x, \lambda), (x, \lambda) - (y, \eta) \rangle \geq L(x, \eta) - L(y, \lambda). \quad (1)$$

Anchoring this inequality to the set $S \times M$ yields the following integrability estimate for the primal-dual gap function; cf. Niederländer [1].

Proposition II.3. *Let $S \times M$ be non-empty and let $(x, \lambda) : [0, +\infty) \to X \times Y$ be a solution of* (AH) *initialized from $(x_0, \lambda_0) \in X \times Y$. Then, for any $(\xi, \eta) \in S \times M$,*

(i) *$t \mapsto \|(x(t), \lambda(t)) - (\xi, \eta)\|$ is non-increasing and*

$$\begin{aligned} &\frac{1}{2}\|(x(t), \lambda(t)) - (\xi, \eta)\|^2 \\ &+ \int_0^t L(x(\tau), \eta) - L(\xi, \lambda(\tau)) \, \mathrm{d}\tau \\ &\leq \frac{1}{2}\|(x_0, \lambda_0) - (\xi, \eta)\|^2 \quad \forall t \geq 0; \end{aligned}$$

(ii) $\lim_{t \to +\infty} \|(x(t), \lambda(t)) - (\xi, \eta)\|$ *exists;*

(iii) $(x, \lambda) \in \mathcal{L}^\infty([0, +\infty); X \times Y)$*;*

(iv) *it holds that*

$$\int_0^\infty L(x(\tau), \eta) - L(\xi, \lambda(\tau)) \, \mathrm{d}\tau < +\infty.$$

Remark II.4. Proposition II.3*(iii)* states that the solutions of (AH) remain bounded whenever the set of saddle points $S \times M$ of the Lagrangian L is non-empty. Conversely, it can be shown that $S \times M$ is non-empty whenever (AH) admits a bounded solution; see Pazy [16, Theorem 8.7].

III. Weak ergodic convergence

In this section, we focus on the convergence properties of the solutions of (AH). As we shall see, one can not, in general, deduce the convergence of the solutions of (AH) even though the set $S \times M$ may be reduced to a singleton. It is thus natural to first study the average of a solution of (AH) as it is better behaved than the solution itself. To this end, let the Cesàro average of a solution $(x, \lambda) : [0, +\infty) \to X \times Y$ of (AH) be defined by

$$\begin{aligned} (\sigma, \omega) : (0, +\infty) &\longrightarrow X \times Y \\ t &\longmapsto \frac{1}{t} \int_0^t (x(\tau), \lambda(\tau)) \, \mathrm{d}\tau \, . \end{aligned}$$

The following theorem asserts that the solutions of (AH) weakly converge in average towards a saddle point of L. The technique we use to prove the weak ergodic convergence relies on the Opial–Passty lemma; cf. Lemma A.1.

Theorem III.1. *Let $S \times M$ be non-empty and let $(\sigma, \omega) : (0, +\infty) \to X \times Y$ be the Cesàro average of a solution of* (AH)*. Then the following assertions hold:*

(i) *for all $(\xi, \eta) \in S \times M$, it holds that*

$$L(\sigma(t), \eta) - L(\xi, \omega(t)) = \mathcal{O}\Big(\frac{1}{t}\Big) \text{ as } t \to +\infty;$$

(ii) *$\forall t_n \to +\infty$ such that $(\sigma(t_n), \omega(t_n)) \rightharpoonup (\bar{\sigma}, \bar{\omega})$ weakly in $X \times Y$, it holds that $(\bar{\sigma}, \bar{\omega}) \in S \times M$.*

Moreover, there exists an element $(\bar{\sigma}, \bar{\omega}) \in S \times M$ such that $(\sigma(t), \omega(t)) \rightharpoonup (\bar{\sigma}, \bar{\omega})$ weakly in $X \times Y$ as $t \to +\infty$.

Proof. *(i)* Let $(\xi, \eta) \in S \times M$ and recall from Proposition II.3*(i)* that for any $t \geq 0$, we have

$$\begin{aligned} &\frac{1}{2}\|(x(t), \lambda(t)) - (\xi, \eta)\|^2 \\ &+ \int_0^t L(x(\tau), \eta) - L(\xi, \lambda(\tau)) \, \mathrm{d}\tau \\ &\leq \frac{1}{2}\|(x(0), \lambda(0)) - (\xi, \eta)\|^2. \end{aligned}$$

Successively dividing by $t > 0$ and taking into account that $\|(x(t), \lambda(t)) - (\xi, \eta)\|^2 \geq 0$ yields

$$\frac{1}{t} \int_0^t L(x(\tau), \eta) - L(\xi, \lambda(\tau)) \, \mathrm{d}\tau \leq \frac{C}{t},$$

where $C = \|(x(0), \lambda(0)) - (\xi, \eta)\|^2/2$. Applying Jensen's inequality, as $L(\,\cdot\,, \eta)$ and $-L(\xi, \,\cdot\,)$ are both convex, we get

$$L(\sigma(t), \eta) - L(\xi, \omega(t)) \leq \frac{C}{t}.$$

Multiplying the above inequality by $t > 0$ and subsequently passing to the upper limit as $t \to +\infty$ entails

$$\limsup_{t \to +\infty} t\big(L(\sigma(t), \eta) - L(\xi, \omega(t))\big) < +\infty.$$

(ii) Using similar arguments as above, we observe that for any $(\xi, \eta) \in X \times Y$ there exists a constant $C \geq 0$ such that for any $t > 0$,

$$L(\sigma(t), \eta) - L(\xi, \omega(t)) \leq \frac{C}{t}.$$

Passing to the upper limit as $t \to +\infty$ gives

$$\limsup_{t \to +\infty} \big(L(\sigma(t), \eta) - L(\xi, \omega(t))\big) \leq 0.$$

Suppose now that $(\sigma(t_n), \omega(t_n)) \rightharpoonup (\bar{\sigma}, \bar{\omega})$ weakly in $X \times Y$, as $n \to +\infty$, for a sequence $t_n \to +\infty$. Substituting t by t_n in the above inequality then yields

$$\begin{aligned} 0 &\geq \limsup_{n \to +\infty} \big(L(\sigma(t_n), \eta) - L(\xi, \omega(t_n))\big) \\ &\geq \liminf_{n \to +\infty} L(\sigma(t_n), \eta) + \liminf_{n \to +\infty} \big(-L(\xi, \omega(t_n))\big) \\ &\geq L(\bar{\sigma}, \eta) - L(\xi, \bar{\omega}), \end{aligned}$$

thanks to the weak lower semi-continuity of $L(\,\cdot\,, \eta)$ as well as $-L(\xi, \,\cdot\,)$ (noticing that $L(\,\cdot\,, \eta)$ and $-L(\xi, \,\cdot\,)$ are both convex and lower semi-continuous); see, e.g., Bauschke and Combettes [3, Theorem 9.1]. The above inequalities being true for any $(\xi, \eta) \in X \times Y$, we infer that $(\bar{\sigma}, \bar{\omega}) \in S \times M$.

The weak convergence of $(\sigma(t), \omega(t))$ as $t \to +\infty$ is now an immediate consequence of the Opial–Passty lemma, cf.

Lemma A.1, applied to the set $S \times M$ together with Proposition II.3*(ii)*. $\square$

Remark III.2. We note that the weak ergodic convergence of the solutions of (AH) may also be deduced using the maximal monotonicity of the operator T; cf. Baillon and Brézis [15].

As an immediate consequence of the previous result, we infer that the asymptotic behavior of the average of a solution of (AH) is characterized according to the following cases:

Corollary III.3. *Let* $(\sigma,\omega):(0,+\infty)\to X\times Y$ *be the Cesàro average of a solution of* (AH)*. Then the following assertions hold:*

(i) If $S\times M$ *is non-empty, then there exists* $(\bar\sigma,\bar\omega)\in S\times M$ *such that* $(\sigma(t),\omega(t))\rightharpoonup(\bar\sigma,\bar\omega)$ *weakly in* $X\times Y$ *as* $t\to+\infty$*;*

(ii) If $S\times M$ *is empty, then* $\lim_{t\to+\infty}\|(\sigma(t),\omega(t))\|=+\infty$.

Proof. (i) This is precisely the assertion of Theorem III.1.

(ii) Suppose, to the contrary, that there exists a sequence $t_n\to+\infty$ such that

$$\sup\{\|(\sigma(t_n),\omega(t_n))\| \mid n\in\mathbb{N}\}<+\infty.$$

Then $(\sigma(t),\omega(t))_{t>0}$ admits at least one weak sequential cluster point, that is, there exist $(\bar\sigma,\bar\omega)\in X\times Y$ and a subsequence of $t_n\to+\infty$ (again denoted by $t_n\to+\infty$) such that $(\sigma(t_n),\omega(t_n))\rightharpoonup(\bar\sigma,\bar\omega)$ weakly in $X\times Y$ as $n\to+\infty$. Following the derivations in the proof of Theorem III.1*(ii)*, we know that for any $(\xi,\eta)\in X\times Y$,

$$\limsup_{n\to+\infty}\big(L(\sigma(t_n),\eta)-L(\xi,\omega(t_n))\big)\le 0.$$

Using again that $L(\,\cdot\,,\eta)$ and $-L(\xi,\,\cdot\,)$ are both weakly lower semi-continuous, we obtain

$$L(\bar\sigma,\eta)-L(\xi,\bar\omega)\le 0,$$

implying that $(\bar\sigma,\bar\omega)\in S\times M$. $\square$

Let us conclude this section with a classical localization result of the weak limit of the average of a solution of (AH). To this end, we introduce the notion of the asymptotic center of a bounded solution $(x,\lambda):[0,+\infty)\to X\times Y$ of (AH); cf. Edelstein [19]. Given $(y,\eta)\in X\times Y$, consider

$$\phi(y,\eta)=\limsup_{t\to+\infty}\|(x(t),\lambda(t))-(y,\eta)\|^2$$

such that ϕ is continuous, strictly convex (in fact, 2-strongly convex) and coercive, that is, $\lim_{\|(y,\eta)\|\to+\infty}\phi(y,\eta)=+\infty$. As a consequence, ϕ admits a unique minimizer, denoted by $\mathrm{ac}(x,\lambda)$, which is called the asymptotic center or "shadow limit" of (x,λ); see, e.g., Aubin and Cellina [18].

Following Brézis [20], the next result localizes the weak limit of the average of a solution of (AH) as the asymptotic center of the solution itself. We provide its proof for completeness.

Proposition III.4. *Let* $S\times M$ *be non-empty and let* $(\bar\sigma,\bar\omega)\in S\times M$ *be such that* $(\sigma(t),\omega(t))\rightharpoonup(\bar\sigma,\bar\omega)$ *weakly in* $X\times Y$ *as* $t\to+\infty$*. Then* $(\bar\sigma,\bar\omega)$ *is the asymptotic center* $\mathrm{ac}(x,\lambda)$.

Proof. Let $(\bar\sigma,\bar\omega)\in S\times M$ be such that $(\sigma(t),\omega(t))\rightharpoonup(\bar\sigma,\bar\omega)$ weakly in $X\times Y$ as $t\to+\infty$. Given $(\xi,\eta)\in X\times Y$, we have for any $t\ge0$,

$$\begin{aligned}\|(x(t),\lambda(t))-(\bar\sigma,\bar\omega)\|^2&=\|(x(t),\lambda(t))-(\xi,\eta)\|^2\\&\quad+2\langle(x(t),\lambda(t))-(\xi,\eta),(\xi,\eta)-(\bar\sigma,\bar\omega)\rangle\\&\quad+\|(\xi,\eta)-(\bar\sigma,\bar\omega)\|^2.\end{aligned}$$

Integrating this equality over $[0,t]$ and subsequently dividing by $t>0$ gives

$$\begin{aligned}&\frac1t\int_0^t\|(x(\tau),\lambda(\tau))-(\bar\sigma,\bar\omega)\|^2\,\mathrm{d}\tau\\&=\frac1t\int_0^t\|(x(\tau),\lambda(\tau))-(\xi,\eta)\|^2\,\mathrm{d}\tau\\&\quad+2\langle(\sigma(t),\omega(t))-(\xi,\eta),(\xi,\eta)-(\bar\sigma,\bar\omega)\rangle\\&\quad+\|(\xi,\eta)-(\bar\sigma,\bar\omega)\|^2.\end{aligned}\tag{2}$$

Since $(\bar\sigma,\bar\omega)\in S\times M$, it follows from Proposition II.3*(ii)* that

$$\lim_{t\to+\infty}\|(x(t),\lambda(t))-(\bar\sigma,\bar\omega)\|^2 \text{ exists.}$$

Moreover, in view of the classical Cesàro property, we readily obtain

$$\phi(\bar\sigma,\bar\omega)=\lim_{t\to+\infty}\frac1t\int_0^t\|(x(\tau),\lambda(\tau))-(\bar\sigma,\bar\omega)\|^2\,\mathrm{d}\tau\,.$$

Owing to the fact $(\sigma(t),\omega(t))\rightharpoonup(\bar\sigma,\bar\omega)$ weakly in $X\times Y$ as $t\to+\infty$, taking the upper limit in equation (2) yields

$$\begin{aligned}&\phi(\bar\sigma,\bar\omega)+\|(\xi,\eta)-(\bar\sigma,\bar\omega)\|^2\\&\quad=\limsup_{t\to+\infty}\frac1t\int_0^t\|(x(\tau),\lambda(\tau))-(\xi,\eta)\|^2\,\mathrm{d}\tau\\&\quad\le\limsup_{t\to+\infty}\|(x(t),\lambda(t))-(\xi,\eta)\|^2\\&\quad=\phi(\xi,\eta).\end{aligned}$$

The preceding derivations being true for any $(\xi,\eta)\in X\times Y$, we conclude the result. $\square$

IV. Refined ergodic estimates

In this section, we refine some of our previous results by further exploiting the structure of the linear and continuous operator A. In particular, let us assume that $A:X\to Y$ is bounded from below[2], i.e.,

$$\exists\beta>0\ \forall x\in X,\quad \|Ax\|_Y\ge\beta\|x\|_X.$$

The following proposition provides a refined asymptotic estimate on the primal-dual gap function.

Proposition IV.1. *Let* $S\times M$ *be non-empty, let* $A:X\to Y$ *be bounded from below, and let* $(\sigma,\omega):(0,+\infty)\to X\times Y$ *be the Cesàro average of a solution of* (AH)*. Then, for any* $(\xi,\eta)\in S\times M$*, it holds that*

$$L(\sigma(t),\eta)-L(\xi,\omega(t))=\mathcal{O}\Big(\frac{1}{t^2}\Big)\ \textit{as}\ t\to+\infty;$$

$$\|\sigma(t)-\xi\|_X=\mathcal{O}\Big(\frac1t\Big)\ \textit{as}\ t\to+\infty.$$

[2]We recall that $A:X\to Y$ is bounded from below if and only if it is injective with closed range; see, e.g., Brézis [21].

Proof. Let $(\xi,\eta)\in S\times M$. Since A is bounded from below there exists $\beta>0$ such that for any $t>0$,

$$\begin{aligned}\beta\|\sigma(t)-\xi\|_X &\le \|A(\sigma(t)-\xi)\|_Y\\ &= \frac{1}{t}\Big\| \int_0^t A(x(\tau)-\xi)\,\mathrm{d}\tau \Big\|_Y.\end{aligned}$$

Using (AH) together with $A\xi=b$, we obtain

$$\begin{aligned}\beta\|\sigma(t)-\xi\|_X &\le \frac{1}{t}\Big\| \int_0^t \dot\lambda(\tau)\,\mathrm{d}\tau \Big\|_Y\\ &= \frac{1}{t}\|\lambda(t)-\lambda_0\|_Y.\end{aligned}$$

Observing that λ remains bounded, cf. Proposition II.3*(iii)*, there exists $C\ge 0$ such that for any $t>0$,

$$\beta\|\sigma(t)-\xi\|_X \le \frac{C}{t}.$$

Multiplying the above inequality by $t>0$ and subsequently passing to the upper limit as $t\to+\infty$ entails

$$\limsup_{t\to+\infty} t\|\sigma(t)-\xi\|_X < +\infty.$$

On the other hand, we immediately observe from inequality (1) that for any $t>0$,

$$\begin{aligned}&\langle T(\sigma(t),\omega(t)),(\sigma(t),\omega(t))-(\xi,\eta)\rangle\\ &\quad\ge L(\sigma(t),\eta)-L(\xi,\omega(t)).\end{aligned}$$

Combining this inequality with the basic identity

$$\begin{aligned}&\langle T(\sigma(t),\omega(t)),(\sigma(t),\omega(t))-(\xi,\eta)\rangle\\ &= \langle \nabla f(\sigma(t))-\nabla f(\xi),\sigma(t)-\xi\rangle_X,\end{aligned}$$

we obtain

$$\begin{aligned}&L(\sigma(t),\eta)-L(\xi,\omega(t))\\ &\le \langle \nabla f(\sigma(t))-\nabla f(\xi),\sigma(t)-\xi\rangle_X.\end{aligned}$$

Since x is bounded, cf. Proposition II.3*(iii)*, it follows that σ remains bounded as well. Owing to the fact that ∇f is Lipschitz continuous on the bounded subsets of X, there exists $L\ge 0$ such that for any $t>0$,

$$L(\sigma(t),\eta)-L(\xi,\omega(t)) \le L\|\sigma(t)-\xi\|_X^2.$$

Multiplying the previous inequality by $t^2>0$ and observing that $t^2\|\sigma(t)-\xi\|_X^2$ remains bounded from above, passing to the upper limit as $t\to+\infty$ yields

$$\limsup_{t\to+\infty} t^2\big(L(\sigma(t),\eta)-L(\xi,\omega(t))\big) < +\infty,$$

concluding the desired estimates. ☐

V. Numerical experiments

In this section, we provide two simple, yet representative, numerical experiments that allow for a direct exposition of our main results.

Example 1 (Ergodic convergence). Let $X,Y=\mathbb{R}$. Take $f=0$, $A=\mathrm{Id}$ and $b=0$ so that

$$\begin{aligned}L:\mathbb{R}\times\mathbb{R}&\longrightarrow \mathbb{R}\\ (x,\lambda)&\longmapsto \lambda x\end{aligned}$$

with $S\times M=\{(0,0)\}$. In this case, the (AH) differential system reads

$$\begin{cases}\dot x+\lambda=0\\ \dot\lambda - x = 0.\end{cases}$$

Clearly, the non-stationary solutions $(x(t),\lambda(t))$ of the above (AH) differential system remain bounded but do not admit a limit as $t\to+\infty$. Figure 1 illustrates the trajectories of a solution $(x(t),\lambda(t))$ of (AH) together with its Cesàro average $(\sigma(t),\omega(t))$ as well as the evolution of the quantity $\|(\sigma(t),\omega(t))-(\bar\sigma,\bar\omega)\|$ with $(\bar\sigma,\bar\omega)\in S\times M$. The initial data of the (AH) solution is chosen as $(x_0,\lambda_0)=(1,0)$.

Example 2 (Refined ergodic estimate). Let $X,Y=\mathbb{R}^2$ and consider $f(x)=(x_1+x_2)^2/2$ with $x=(x_1,x_2)$. Further, let $A=\mathrm{Id}_{\mathbb{R}^2}$ and $b=0_{\mathbb{R}^2}$ so that A is bounded from below and

$$\begin{aligned}L:\mathbb{R}^2\times\mathbb{R}^2&\longrightarrow \mathbb{R}\\ (x,\lambda)&\longmapsto f(x)+\langle\lambda,x\rangle\end{aligned}$$

with $S\times M=\{(0_{\mathbb{R}^2},0_{\mathbb{R}^2})\}$. The trajectories of a solution component $x(t)=(x_1(t),x_2(t))$ of (AH) along with its Cesàro average $\sigma(t)=(\sigma_1(t),\sigma_2(t))$ as well as the evolution of the primal-dual gap function $L(\sigma(t),\bar\omega)-L(\bar\sigma,\omega(t))$ with $(\bar\sigma,\bar\omega)\in S\times M$ are depicted in Figure 2. The initial data is set to $x_0=(-1,-1)$ and $\lambda_0=(-2,0)$.

Appendix

For the following classical result, named the Opial–Passty lemma, the reader is referred to Passty [22].

Lemma A.1 (Opial–Passty)**.** *Let X be a real Hilbert space, let S be a non-empty subset of X, and let $x:[0,+\infty)\to X$ be continuous. For any $t>0$, set*

$$\sigma(t)=\frac{1}{t}\int_0^t x(\tau)\,\mathrm{d}\tau$$

and assume that

(i) for all $\xi\in S$, $\lim_{t\to+\infty}\|x(t)-\xi\|_X$ exists;
(ii) $\forall t_n\to+\infty$ such that $\sigma(t_n)\rightharpoonup\bar\sigma$ weakly in X, it holds that $\bar\sigma\in S$.

Then $\sigma(t)$ converges weakly, as $t\to+\infty$, to some element $\bar\sigma\in S$.

References

[1] S. K. Niederländer, “On the Arrow–Hurwicz differential system for linearly constrained convex minimization,” *Optimization*, 2023. [Online]. Available: https://doi.org/10.1080/02331934.2023.2215799
[2] I. Ekeland and R. Témam, *Convex analysis and variational problems*, ser. Classics in applied mathematics. Philadelphia: Society for Industrial and Applied Mathematics, 1999.
[3] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, ser. CMS Books in Mathematics. New York: Springer, 2017.
[4] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex analysis and minimization algorithms I*, ser. Grundlehren der mathematischen Wissenschaften 305. New York: Springer, 1993.
[5] K. J. Arrow and L. Hurwicz, “A gradient method for approximating saddle points and constrained maxima,” *RAND Corp., Santa Monica, CA*, pp. P–223, 1951.
[6] K. J. Arrow, L. Hurwicz, and H. Uzawa, *Studies in linear and non-linear programming*. Stanford, CA: Stanford University Press, 1958.

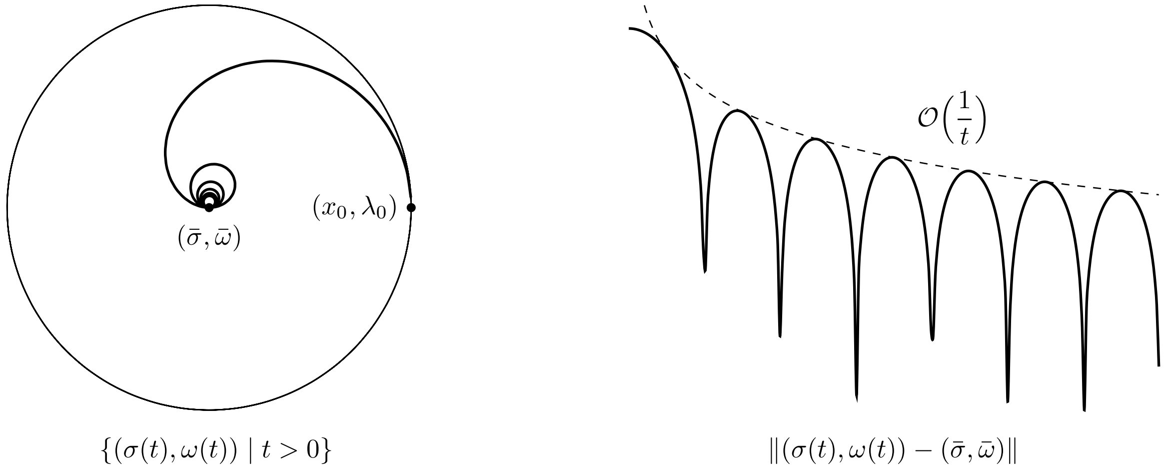

Fig. 1. Graphical view on the trajectories of a solution $(x(t), \lambda(t))$ of (AH) and its Cesàro average $(\sigma(t), \omega(t))$ (left panel) as well as on the quantity $\|(\sigma(t), \omega(t)) - (\bar{\sigma}, \bar{\omega})\|$ plotted in a logarithmic scale (right panel). As asserted by Proposition II.3, the solution $(x(t), \lambda(t))$ of (AH) remains bounded but does not admit a limit as $t \to +\infty$. However, its Cesàro average $(\sigma(t), \omega(t))$ converges, as $t \to +\infty$, towards the unique element $(\bar{\sigma}, \bar{\omega}) \in S \times M$ which is the asymptotic center $\mathrm{ac}(x, \lambda)$; cf. Proposition III.4. Moreover, the error $\|(\sigma(t), \omega(t)) - (\bar{\sigma}, \bar{\omega})\|$ obeys the asymptotic estimate $\mathcal{O}(1/t)$ as $t \to +\infty$ as observed in Theorem III.1 for the primal-dual gap function $L(\sigma(t), \bar{\omega}) - L(\bar{\sigma}, \omega(t))$.

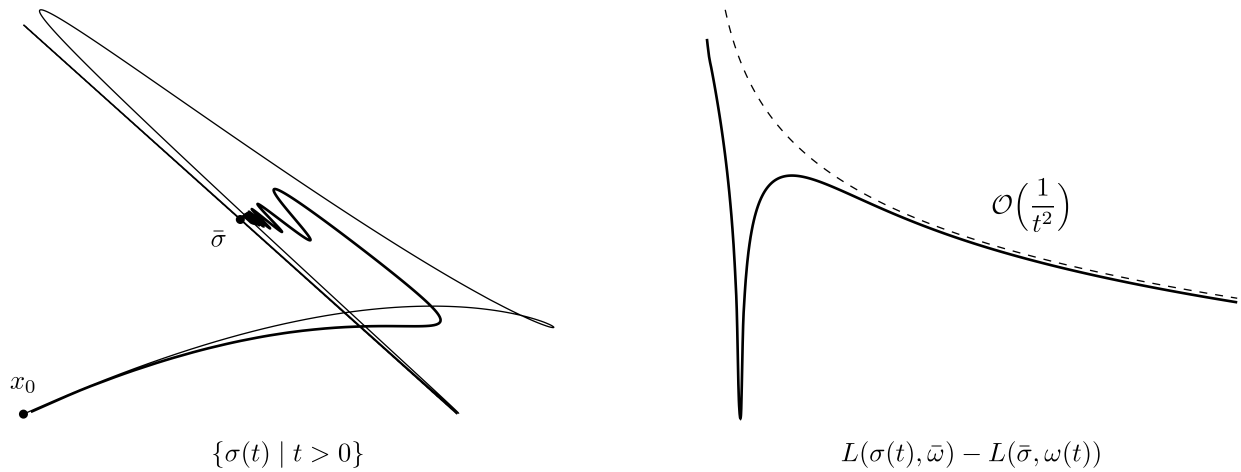

Fig. 2. Graphical view on the trajectories of a solution component $x(t) = (x_1(t), x_2(t))$ of (AH) and its Cesàro average $\sigma(t) = (\sigma_1(t), \sigma_2(t))$ (left panel) as well as on the evolution of the primal-dual gap function $L(\sigma(t), \bar{\omega}) - L(\bar{\sigma}, \omega(t))$ plotted in a logarithmic scale (right panel). Even though the solution $(x(t), \lambda(t))$ of (AH) does not admit a limit as $t \to +\infty$, its Cesàro average $(\sigma(t), \omega(t))$ converges towards the unique element $(\bar{\sigma}, \bar{\omega}) \in S \times M$; cf. Theorem III.1. Moreover, as predicted by Proposition IV.1, the primal-dual gap function $L(\sigma(t), \bar{\omega}) - L(\bar{\sigma}, \omega(t))$ obeys the refined asymptotic estimate $\mathcal{O}(1/t^2)$ as $t \to +\infty$.

[7] R. T. Rockafellar, "Monotone operators associated with saddle-functions and minimax problems," *In Nonlinear Functional Analysis, Proceedings of Symposia in Pure Math., Amer. Math. Soc.*, pp. 241–250, 1969.
[8] J. Peypouquet and S. Sorin, "Evolution equations for maximal monotone operators: Asymptotic analysis in continuous and discrete time," *J. Convex Anal.*, vol. 17, pp. 1113–1163, 2010.
[9] R. T. Rockafellar, "Saddle-points and convex analysis," *In Differential Games and Related Topics, North-Holland*, pp. 109–127, 1971.
[10] M. G. Crandall and A. Pazy, "Semi-groups of nonlinear contractions and dissipative sets," *J. Funct. Anal.*, vol. 3, pp. 376–418, 1969.
[11] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, ser. Mathematics Studies 5. Amsterdam: North-Holland, 1973.
[12] T. Kato, "Nonlinear semigroups and evolution equations," *J. Math. Soc. Japan*, vol. 19, pp. 508–520, 1967.
[13] Y. Kōmura, "Nonlinear semi-groups in Hilbert space," *J. Math. Soc. Japan*, vol. 19, pp. 493–507, 1967.
[14] F. E. Browder, "Nonlinear operators and nonlinear equations of evolution in Banach spaces," *In Nonlinear Functional Analysis, Proceedings of Symposia in Pure Math., Amer. Math. Soc.*, 1976.
[15] J.-B. Baillon and H. Brézis, "Une remarque sur le comportement asymptotique des semigroupes non linéaires," *Houston J. Math.*, vol. 2, pp. 5–7, 1976.
[16] A. Pazy, "Semi-groups of nonlinear contractions and their asymptotic behavior," *In Nonlinear Analysis and Mechanics: Heriot-Watt Symposium III, Pitman, London*, pp. 36–134, 1979.
[17] A. S. Nemirovski and D. B. Yudin, "Cesari convergence of the gradient method of approximating saddle points of convex-concave functions," *Dokl. Akad. Nauk SSSR*, vol. 239, pp. 1056–1059, 1978.
[18] J. P. Aubin and A. Cellina, *Differential inclusions*, ser. Grundlehren der mathematischen Wissenschaften 264. New York: Springer, 1984.
[19] M. Edelstein, "The construction of an asymptotic center with a fixed-point property," *Bull. Amer. Math. Soc.*, vol. 78, pp. 206–208, 1972.
[20] H. Brézis, "Asymptotic behavior of some evolution systems," *In Nonlinear Evolution Equations, Madison, 1977, Acad. Press*, pp. 141–154, 1978.
[21] ——, *Functional analysis, Sobolev spaces and partial differential equations*. New York: Springer, 2011.
[22] G. B. Passty, "Ergodic convergence to a zero of the sum of monotone operators in Hilbert space," *J. Math. Anal. Appl.*, vol. 72, pp. 383–390, 1979.