AN ENHANCED BAILLON-HADDAD THEOREM FOR CONVEX FUNCTIONS ON CONVEX SETS

PEDRO PÉREZ-AROS AND EMILIO VILCHES

Abstract. In this paper, we prove the Baillon-Haddad theorem for Gâteaux differentiable convex functions defined on open convex sets of arbitrary Hilbert spaces. Formally, this result establishes that the gradient of a convex function defined on an open convex set is $\beta$-Lipschitz if and only if it is $1/\beta$-cocoercive. An application to convex optimization through dynamical systems is given.

1. Introduction

Let $\mathcal{H}$ be a Hilbert space endowed with a scalar product $\langle \cdot, \cdot \rangle$, induced norm $\| \cdot \|$ and unit ball $B$. Given a nonempty set $\Omega \subset \mathcal{H}$ and $\beta > 0$, we say that an operator $T: \Omega \to \mathcal{H}$ is $1/\beta$-cocoercive if for all $x, y \in \Omega$

$$\beta \langle Tx - Ty, x - y \rangle \geq \| Tx - Ty \|^2,$$

and $T$ is $\beta$-Lipschitz continuous if for all $x, y \in \Omega$

$$\| Tx - Ty \| \leq \beta \| x - y \|.$$

If $\beta = 1$, then (1) means that $T$ is firmly nonexpansive and (2) that $T$ is nonexpansive (see, e.g., [5, Chapter 4]). It is clear that (1) implies (2), while the converse, in general, is false (take for example $T = -\text{Id}$). Despite of this negative result, the Baillon-Haddad theorem ([3, Corollaire 10]) states that if $T$ is the gradient of a convex function, then (1) and (2) are equivalent. The precise statement is the following:

Theorem 1.1 (Baillon-Haddad). Let $f: \mathcal{H} \to \mathbb{R}$ be convex, Fréchet differentiable on $\mathcal{H}$, and such that $\nabla f$ is $\beta$-Lipschitz continuous for some $\beta > 0$. Then $\nabla f$ is $1/\beta$-cocoercive.

This prominent result provides an important link between convex optimization and fixed-point iteration [8]. Moreover, it has many applications in optimization and numerical functional analysis (see, e.g., [2, 5, 9, 10, 17]). An improved version of Theorem 1.1 appeared in [4] (see also [8, Theorem 1.2]), where the authors relate the Lipschitzianity of the gradients of a convex function with the convexity and Moreau envelopes of associated functions (see [4, Theorem 2.1]). Furthermore, they provided the following Baillon-Haddad theorem for twice continuously differentiable convex functions defined on open convex sets.

Theorem 1.2. [4, Theorem 3.3] Let $\Omega$ be a nonempty open convex subset of $\mathcal{H}$, let $f: \Omega \to \mathbb{R}$ be convex and twice continuously Fréchet differentiable on $\Omega$, and let $\beta > 0$. Then $\nabla f$ is $\beta$-Lipschitz continuous if and only if it is $1/\beta$-cocoercive.

Finally, the authors left as an open question the validity of Theorem 1.2 for Gâteaux differentiable convex functions (see [4, Remark 3.5]).
The aim of this paper is to extend Theorem 1.2 to merely Gâteaux differentiable convex functions (see Theorem 3.1). To do that, we first establish the result in finite-dimensions and then we use a finite dimensional reduction.

We emphasize that extend Theorem 1.2 is of interest because it provides an important link between the gradient of convex functions defined on convex sets and cocoercive operators defined on convex sets.

Cocoercivity arises in various areas of optimization and nonlinear analysis (see, e.g., [1, 5, 6, 11, 14]). In particular, it plays an important role in the design of algorithms to solve structured monotone inclusions (which includes fixed points of non-expansive operators). Indeed, let us consider the structured monotone inclusion: find $x \in H$ such that

$$0 \in \partial \Phi(x) + Bx,$$

where $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ is a convex lower semicontinuous function and $B : H \to H$ is a monotone operator. It is well known that (see, e.g., [1]) the problem (3) is equivalent to the fixed point problem: find $x \in H$ such that

$$x = \text{prox}_{\mu \Phi}(x - \mu Bx),$$

where $\mu > 0$ and $\text{prox}_{\mu \Phi} : H \to H$ is the proximal mapping of $\Phi$ (see, e.g., [5] Definition 12.23) defined by

$$\text{prox}_{\mu \Phi}(x) := \text{argmin}_{y \in H} \left\{ \Phi(y) + \frac{1}{2\mu} \|y - x\|^2 \right\}.$$

To solve the fixed point problem (4), Abbas and Attouch [1] introduces the following dynamical system

$$\dot{x}(t) + x(t) = \text{prox}_{\mu \Phi}(x(t) - \mu Bx(t)),$$

whose equilibrium points are solutions of (4). They proved the following result (see [1] Theorem 5.2)

**Proposition 1.3.** Let $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ be a convex lower semicontinuous proper function, and $B$ a maximal monotone operator which is $\beta$-cocoercive. Suppose that $\mu \in (0, 2\beta)$ and $\text{zer} (\partial \Phi + B) := \{ z \in H : 0 \in \partial \Phi(z) + Bz \} \neq \emptyset$.

Then the unique solution of (5) weakly converges to some element $x_\infty \in \text{zer} (\partial \Phi + B)$.

The previous result was extended by Bot and Csetnek (see [6] Theorem 12) to solve the monotone inclusion: find $x \in H$ such that

$$0 \in Ax + Bx,$$

where $A : H \rightrightarrows H$ is a maximal monotone operator and $B : H \to H$ is $\beta$-cocoercive. These two results were extended by the authors in [14], where we proved the strong convergence of a Tikhonov regularization for the dynamical system (5). It is important to emphasize that in order to solve the problems (4) and (5), it is enough that the operator $B$ is defined in $\text{dom} \partial \Phi$ and $\text{dom} A$, respectively. Therefore, it is interesting to have characterizations of cocoercive operators defined on open convex subsets of $H$. Thus, it is important to extend Theorem 1.2 to merely Gâteaux differentiable functions (see [4] Remark 3.5).

The paper is organized as follows. After some preliminaries, in Section 3 we state and prove the main result of the paper (Theorem 3.1). Then, in Section 4 we give an application to convex optimization.
2. Notation and Preliminaries

Let $\mathcal{H}$ be a Hilbert space endowed with a scalar product $\langle \cdot, \cdot \rangle$, induced norm $\| \cdot \|$ and unit ball $B$. We denote by $\mathcal{L}(H,H)$ the set of continuous linear operators from $\mathcal{H}$ into $\mathcal{H}$. The norm of an operator $A \in \mathcal{L}(H,H)$ is defined by

$$\|A\|_{\mathcal{L}(H,H)} := \sup_{h \in B} \|Ah\|.$$ 

Given an open convex set $\Omega \subset H$, we denote by $C^1_1(\Omega)$ the class of Fréchet differentiable functions $f : \Omega \subset H \rightarrow \mathbb{R}$ whose gradient $\nabla f$ is locally Lipschitz (see, e.g., [15, Chapter 9]).

Given $\Omega \subset H$ and $\beta > 0$, we say that $T : \Omega \rightarrow H$ is $\beta$-cocoercive on $\Omega$ if for all $x, y \in \Omega$

$$\beta \langle Tx - Ty, x - y \rangle \geq \|Tx - Ty\|^2.$$

Example 1. The following list provides some examples of cocoercive operators (see refer to [5, Chapter 4] for further properties on cocoercive operators):

(i) $T : \Omega \rightarrow H$ is nonexpansive if and only if $I - T$ is $1/2$-cocoercive.

(ii) $T : \Omega \rightarrow H$ is $1$-cocoercive if and only if $2T - I$ is $1$-Lipschitz.

(iii) A matrix $M$ is psd-plus (that is, $M = E^tAE$ for some $A$ positive definite) if and only if the mapping $x \mapsto Mx$ is cocoercive (see [17, Proposition 2.5]).

(iv) The Yosida approximation $A_\lambda := \frac{1}{\lambda}(I - (I + A)^{-1})$ of a maximal monotone operator $A : H \rightrightarrows H$ is $\lambda$-cocoercive (see [5, Corollary 23.11]).

For a convex function $f : \Omega \subset H \rightarrow \mathbb{R}$ we consider the convex subdifferential of $f$ at $x \in \Omega$ as

$$\partial f(x) := \{y \in H : f(x) + \langle \nabla f(x), y - x \rangle \leq f(y) \text{ for all } y \in H\}.$$ 

It is well-known that for two functions $f, g : \Omega \subset H \rightarrow \mathbb{R}$ the following equality holds (see, e.g., [5, Corollary 16.48]):

$$\partial(f + g)(x) = \partial f(x) + \partial g(x).$$ (7)

To prove our main result, we will use finite dimensional reduction arguments, thus, some elements of generalized differentiation in finite dimensions will be needed. We refer to [15] for more details.

Let $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be a $C^{1,+}(\Omega)$ function. For $\bar{x} \in \Omega$, we define the Generalized Hessian of $f$ at $\bar{x}$ (see, e.g., [15, Theorem 9.62] and [12]) as the set of matrices

$$\nabla^2 f(\bar{x}) := \{A \in \mathbb{R}^{n \times n} \mid \exists x_n \rightarrow \bar{x}, x_n \in D, \nabla^2 f(x_n) \rightarrow A\},$$

where $D \subset \Omega$ is the dense set of points where $f$ is twice differentiable (by virtue of Rademacher’s theorem the set $D$ exists). The following result (see [15, Theorem 13.52]) establishes some properties of the Generalized Hessian $\nabla^2 f(\bar{x})$.

Proposition 2.1. Let $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be a $C^{1,+}(\Omega)$ function, where $\Omega \subset \mathbb{R}^n$ is an open set. Then $\nabla^2 f(\bar{x})$ is a nonempty, compact set of symmetric matrices.

The following result gives a known characterization of convexity and Lipschitzianity of functions (see, e.g., [15, 12]). We give a proof for completeness.

Proposition 2.2. Let $f : \Omega \rightarrow \mathbb{R}$ be a $C^{1,+}(\Omega)$ function with $\Omega \subset \mathbb{R}^n$ convex. Then
(i) \( f \) is convex if and only if for all \( x \in \Omega \) and all \( A \in \nabla^2 f(x) \) one has
\[
\langle Au, u \rangle \geq 0 \quad \text{for all } u \in \mathbb{R}^n.
\]

(ii) \( \nabla f \) is 1-Lipschitz on \( \Omega \) if and only if for all \( x \in \Omega \) and all \( A \in \nabla^2 f(x) \) the inequality \( ||A|| \leq 1 \) holds.

Proof. (i) follows from [12, Example 2.2]. The necessity in (ii) is direct. To prove the sufficiency in (ii), it is enough to assume that for all \( x \in \Omega \) and all \( A \in \nabla^2 f(x) \) the inequality \( ||A|| \leq 1 \) holds. Fix \( y \in \mathbb{B} \) and consider the function \( g_y(x) := \langle \nabla f(x), y \rangle \). Then, \( g_y \) is locally Lipschitz on \( \Omega \) and
\[
\nabla g_y(x) = \langle \nabla^2 f(x), y \rangle \quad \text{a.e. } x \in \Omega.
\]
Thus,
\[
\sup_{w \in \nabla g_y(x)} ||w|| \leq ||y|| \quad \text{for all } x \in \Omega.
\]
Hence, according to [13, Theorem 3.5.2], the map \( g_y \) is \( ||y|| \)-Lipschitz on \( \Omega \). Finally, by virtue of [15, Exercise 9.9], we conclude that \( \nabla f \) is 1-Lipschitz on \( \Omega \). \( \square \)

3. An enhanced Baillon-Haddad theorem

In this section, we state and prove the main result of the paper, that is, the Baillon-Haddad theorem for convex functions defined on convex sets, which extends [3, Theorem 3.3] and solves the question posed in [4, Remark 3.5].

Theorem 3.1. Let \( \Omega \) be a nonempty open convex subset of a Hilbert space \( \mathcal{H} \), let \( f: \Omega \to \mathbb{R} \) be a convex function and \( \beta \in [0, +\infty] \). Then the following are equivalent.

(a) \( \nabla f \) is \( \beta \)-Lipschitz continuous on \( \Omega \).

(b) the map \( x \mapsto \frac{\beta}{2} ||x||^2 - f(x) \) is convex on \( \Omega \).

(c) \( \nabla f \) is Gâteaux differentiable on \( \Omega \) and \( \nabla f \) is \( 1/\beta \)-cocoercive.

Moreover, if any of the above conditions hold, then \( f \in C^{1,\beta}(\Omega) \).

To prove Theorem 3.1, we show first the result in finite dimension under the additional assumption that \( f \in C^{1,\beta}(\Omega) \) (see the next lemma). Then, we obtain Theorem 3.1 in finite dimensional spaces (see Lemma 3.3). Finally, the proof of Theorem 3.1 follows from finite dimensional reductions and Lemma 3.3.

Lemma 3.2. Let \( \Omega \) be a nonempty open convex subset of \( \mathbb{R}^n \), let \( f: \Omega \to \mathbb{R} \) be a \( C^{1,\beta}(\Omega) \) convex function and \( \beta \in [0, +\infty] \). Then the following are equivalent.

(a) \( \nabla f \) is \( \beta \)-Lipschitz continuous on \( \Omega \).

(b) the map \( x \mapsto \frac{\beta}{2} ||x||^2 - f(x) \) is convex on \( \Omega \).

(c) \( \nabla f \) is \( 1/\beta \)-cocoercive.

Proof. Let us consider the functions \( g(x) := \frac{1}{2} ||x||^2 - \frac{1}{\beta} f(x) \) and \( h(x) := \frac{2}{\beta} f(x) - \frac{1}{2} ||x||^2 \). It is clear that
\[
A \in \overline{H}_{f/\beta}(x) \iff B := I - A \in \overline{H}_g(x).
\]
On the one hand,
\[ \nabla f \text{ is } \beta\text{-Lipschitz continuous} \]
\[ \iff (\forall x \in \Omega)(\forall A \in \overline{T}_{f/\beta}(x))(\forall u \in \mathbb{R}^n) 0 \leq \langle u, Au \rangle \leq \|A\| \leq 1 \]  
(by Proposition 2.2 (ii))
\[ \iff (\forall x \in \Omega)(\forall A \in \overline{T}_{f/\beta}(x))(\forall u \in \mathbb{R}^n) 0 \leq \|u\|^2 \]  
(by Proposition 2.2 (i))
\[ \iff (\forall x \in \Omega)(\forall A \in \overline{T}_{f/\beta}(x))(\forall u \in \mathbb{R}^n) 0 \leq \|u\|^2 - \langle u, Au \rangle \]
\[ \iff (\forall x \in \Omega)(\forall B \in \overline{T}_g(x))(\forall u \in \mathbb{R}^n) 0 \leq \langle u, Bu \rangle \]  
(by 3)
\[ \iff g \text{ is convex} \]  
(by Proposition 2.2 (i)),
which shows that (a) is equivalent to (b).

On the other hand,
\[ g \text{ is convex} \]
\[ \iff (\forall x \in \Omega)(\forall B \in \overline{T}_g(x))(\forall u \in \mathbb{R}^n) 0 \leq \langle u, Bu \rangle \]  
(by Proposition 2.2 (i))
\[ \iff (\forall x \in \Omega)(\forall A \in \overline{T}_{f/\beta}(x))(\forall u \in \mathbb{R}^n) 0 \leq \|u\|^2 - \langle u, Au \rangle \]
\[ \iff (\forall x \in \Omega)(\forall A \in \overline{T}_{f/\beta}(x))(\forall u \in \mathbb{R}^n) - \|u\|^2 \leq 2\langle u, Au \rangle - \|u\|^2 \leq \|u\|^2 \]
\[ \iff (\forall x \in \Omega)(\forall B \in \overline{T}_h(x)) \|B\| \leq 1 \]
\[ \iff \text{the map } x \mapsto \nabla h(x) = \frac{2}{\beta} \nabla f(x) - x \text{ is } 1\text{-Lipschitz} \]  
(by Proposition 2.2 (ii))
\[ \iff \nabla f \text{ is } 1/\beta\text{-cocoercive} \]  
(by Example 1 (ii)),
which proves that (b) is equivalent to (c). \[ \square \]

Now we proceed to delete the hypothesis \( f \in C^{1,+}(\Omega) \) from Lemma 3.2.

**Lemma 3.3.** Let \( \Omega \) be a nonempty open convex subset of \( \mathbb{R}^n \), let \( f : \Omega \to \mathbb{R} \) be a convex function and \( \beta \in [0, +\infty[ \). Then the following are equivalent.

(a) \( f \) is Gâteaux differentiable on \( \Omega \) and \( \nabla f \) is \( \beta \)-Lipschitz continuous on \( \Omega \).

(b) the map \( x \mapsto \frac{\beta}{2} \|x\|^2 - f(x) \) is convex on \( \Omega \).

(c) \( f \) is Gâteaux differentiable on \( \Omega \) and \( \nabla f \) is \( 1/\beta\)-cocoercive.

**Proof.** According to [7, Theorem 2.2.1], for functions defined on subsets of \( \mathbb{R}^n \), Gâteaux differentiability is equivalent to Fréchet differentiability. We proceed to show that any of the above conditions imply that \( f \in C^{1,+}(\Omega) \). Indeed, it is clear that (a) and (c) implies that \( f \in C^{1,+}(\Omega) \). To prove that (b) implies that \( f \in C^{1,+}(\Omega) \), we follow some ideas from [8]. Let us define \( h(x) := \frac{\beta}{2} \|x\|^2 - f(x) \). Thus,
\[ \frac{\beta}{2} \|x\|^2 = f(x) + h(x) \quad x \in \Omega, \]
which implies that \( \beta x = \partial f(x) + \partial h(x) \) for all \( x \in \Omega \). Therefore, \( \partial f(x) \) and \( \partial h(x) \) are non-empty and contain a single element. Hence, by virtue of [7, Corollary 4.2.5], the function \( f \) is Gâteaux differentiable on \( \Omega \) and, thus, Fréchet differentiable on \( \Omega \) and continuously differentiable on \( \Omega \) (see [7, Theorem 2.2.2]). It is not difficult to prove that (b) implies the following inequality:
\[ \frac{1}{2} \|x - y\|^2 \geq D_f(x, y) := f(x) - f(y) - \langle \nabla f(y), x - y \rangle \geq 0 \quad \text{for all } x, y \in \Omega. \]
Fix \( y \in \Omega \) and define \( d(x) := D_f(x, y) \). Then \( \nabla d(x) = \nabla f(x) - \nabla f(y) \) and \( D_f(z, x) = D_d(z, x) \) for all \( z \) and \( x \). Thus, we obtain

\[
\frac{1}{2} \| z - x \|^2 \geq D_d(z, x) = d(z) - d(x) - \langle \nabla d(x), z - x \rangle \quad \text{for all } z, x \in \Omega.
\] (9)

Fix \( \bar{x} \in \Omega \) and \( \delta > 0 \) such that \( \bar{x} + \delta \mathbb{B} \subset \Omega \) such that

\[
\sup_{z \in \bar{x} + \delta \mathbb{B}} \| \nabla f(z) \| < +\infty.
\]

Let \( x, y \in \bar{x} + \frac{\delta}{2} \mathbb{B} \) and \( t \in (0, 1) \) such that \( z := x + t(\nabla f(y) - \nabla f(x)) \in \bar{x} + \delta \mathbb{B} \). Therefore, by taking \( z = x + t(\nabla f(y) - \nabla f(x)) \) in (9) and using that \( d(z) \geq 0 \), we obtain

\[
D_f(x, y) \geq t \left( 1 - \frac{t}{2} \right) \| \nabla f(x) - \nabla f(y) \|^2.
\]

Analogously, we get

\[
D_f(y, x) \geq t \left( 1 - \frac{t}{2} \right) \| \nabla f(x) - \nabla f(y) \|^2.
\]

Thus, for all \( x, y \in \bar{x} + \frac{\delta}{2} \mathbb{B} \)

\[
\langle \nabla f(x) - \nabla f(y), x - y \rangle = D_f(x, y) + D_f(y, x) \geq t (2 - t) \| \nabla f(x) - \nabla f(y) \|^2;
\]

which shows that \( \nabla f \) is Lipschitz on \( \bar{x} + \frac{\delta}{2} \mathbb{B} \). Therefore, \( f \in C^{1,1}(\Omega) \). \( \square \)

Now we are ready to prove Theorem 3.1

**Proof of Theorem 3.1**

(a) \( \Rightarrow \) (b): Let \( x, y \in \Omega \) and define \( F := \text{span}\{x, y\} \). We observe that \( (F, \langle \cdot, \cdot \rangle) \) is a finite dimensional Hilbert space. Thus the restriction of \( f \) to \( F \), \( f|_F \), is Gâteaux differentiable in \( F \) and for all \( a, b, h \in F \)

\[
\langle \nabla f|_F(a) - \nabla f|_F(b), h \rangle = \langle \nabla f(a) - \nabla f(b), h \rangle.
\]

Hence, for all \( a, b, h \in F \)

\[
\| \nabla f|_F(a) - \nabla f|_F(b) \| = \sup_{h \in \mathbb{B} \cap F} \langle \nabla f|_F(a) - \nabla f|_F(b), h \rangle
\]

\[
\leq \| \nabla f(a) - \nabla f(b) \|
\]

\[
\leq \beta \| a - b \|,
\]

which shows that \( f|_F \) is \( \beta \)-Lipschitz on \( \Omega \cap F \). Therefore, according to Lemma 3.3, the map

\[
x \mapsto h(x) := \frac{\beta}{2} \| x \|^2 - f|_F(x),
\]

is convex on \( \Omega \cap F \). In particular, for all \( x, y \in F \), for all \( \lambda \in [0, 1] \)

\[
h(\lambda x + (1 - \lambda)y) \leq \lambda h(x) + (1 - \lambda)h(y).
\]

Since \( x, y \) are arbitrary, it follows that the map \( x \mapsto \frac{\beta}{2} \| x \|^2 - f(x) \) is convex on \( \Omega \).

(b) \( \Rightarrow \) (a): We first observe that \( x \mapsto h(x) := \frac{\beta}{2} \| x \|^2 - f(x) \) is convex (with finite values) and
for all \( x \in \Omega \)

\[
\beta \frac{1}{2} \|x\|^2 = f(x) + g(x).
\]

Hence, by virtue of (7), for all \( x \in \Omega \)

\[
\beta x = \partial f(x) + \partial h(x),
\]

which implies that \( \partial f(x) \) and \( \partial h(x) \) are non-empty and contain a single element. Therefore, according to [7, Corollary 4.2.5], the function \( f \) and \( g \) are Gâteaux differentiable on \( \Omega \). Thus, if \( F \subset H \) is finite dimensional, then \( h|_F \) is convex on \( U \cap F \). Hence, by virtue of Lemma 3.3, \( f|_F \) is \( \beta \)-Lipschitz on \( \Omega \), i.e., for all \( x,y \in F \)

\[
\left\langle \nabla f(x) - \nabla f(y), h \right\rangle = \left\| \nabla f|_F(x) - \nabla f|_F(y) \right\|_{L(F,F)} \leq \beta \|x - y\|.
\]

Let us consider

\[
F_{x,y} := \{ F \subset H : F \text{ is a linear subspace of } H \text{ with } x,y \in F \text{ and } \dim F < +\infty \}.
\]

Hence, since (10) holds for any \( F \subset H \) finite dimensional, we obtain

\[
\sup_{h \in B \cap F} \left\langle \nabla f(x) - \nabla f(y), h \right\rangle = \left\| \nabla f|_F(x) - \nabla f|_F(y) \right\|_{L(F,F)} \leq \beta \|x - y\|.
\]

Therefore, by taking supremum in (10), we conclude that for all \( x,y \in \Omega \)

\[
\left\| \nabla f(x) - \nabla f(y) \right\| \leq \beta \|x - y\|,
\]

which proves (a).

(c) \( \Rightarrow \) (a): It is straightforward.

(a) \( \Rightarrow \) (b): Let \( F \subset H \) with \( \dim F < +\infty \). Then \( (F, \langle \cdot, \cdot \rangle) \) is a Hilbert space and for all \( x,y \in \Omega \cap F \)

\[
\beta \left\langle \nabla f|_F(x) - \nabla f|_F(y), x - y \right\rangle = \beta \left( \nabla f(x) - \nabla f(y), x - y \right) \\
\geq \left\| \nabla f(x) - \nabla f(y) \right\|_{L(H,H)} \\\n\geq \left\| \nabla f(x) - \nabla f(y) \right\|_{L(F,F)}.
\]

Hence, as a result of Lemma 3.3 for all \( x,y \in \Omega \cap F \)

\[
\left\| \nabla f|_F(x) - \nabla f|_F(y) \right\|_{L(F,F)} \leq \beta \|x - y\|_F = \|x - y\|_H.
\]

Since \( F \) is arbitrary, by using (11), we conclude that for all \( x,y \in \Omega \)

\[
\left\| \nabla f(x) - \nabla f(y) \right\|_{L(H,H)} \leq \beta \|x - y\|,
\]

which shows the equivalence between (a), (b) and (c). Finally, the fact that any of the above conditions imply that \( f \in C^{1,+}(\Omega) \) follows from Smulian’s theorem (see, e.g., [7, Theorem 4.2.10]).

\( \square \)
4. Application to Convex Optimization

In this section, we present an application of Theorem 3.1 to convex optimization. Let $H$ be a Hilbert space, $\varphi \in \Gamma_0(H)$ and $\psi: \Omega \to \mathbb{R}$ be a convex function defined over an open convex set $\Omega \subset H$ with $\text{dom } \varphi \subset \Omega$

We study the Tikhonov regularization for the projected dynamical system (5) (see [10, 14] for more details on Tikhonov regularization). Let us consider the following assumptions:

**Assumption $\mathcal{E}$** Let $\varepsilon: \mathbb{R}_+ \to \mathbb{R}_+$ be a positive function satisfying

(a) $\varepsilon$ is absolutely continuous, nonincreasing and $\lim_{t \to +\infty} \varepsilon(t) = 0$;
(b) $\int_0^{+\infty} \varepsilon(s)ds = +\infty$;
(c) $\lim_{t \to +\infty} \frac{\dot{\varepsilon}(t)}{\varepsilon^2(t)} = 0$.

We observe that, for example, the function $\varepsilon(t) = \frac{1}{(1 + t)^\beta}$ with $\beta \in (0, 1)$ satisfies the previous assumption.

**Theorem 4.1.** Assume, in addition to Assumption $\mathcal{E}$, that $\nabla \psi$ is $\beta$-Lipschitz on $\Omega$ and $\text{zer}(\partial \varphi + \nabla \psi) \neq \emptyset$. Let $x_0, y \in \overline{\text{dom } \varphi}$ and $\varepsilon$ be a function satisfying Assumption $\mathcal{E}$. Let $x \in [0, +\infty[ \to H$ be the unique solution of

\[
\begin{cases}
-\dot{x}(t) = x(t) - \text{prox}_{\mu \varphi}(x(t) - \mu \nabla \psi(x(t))) + \varepsilon(t)(x(t) - y), \\
x(0) = x_0,
\end{cases}
\]

where $\mu \in (0, 2/\beta)$. Then $x(t)$ converges strongly to $\text{proj}_{\text{zer}(\partial \varphi + \nabla \psi)}(y)$, as $t \to +\infty$.

**Proof.** According to Theorem 3.1 the operator $\nabla \psi: \overline{\text{dom } \varphi} \to H$ is $\frac{1}{\beta}$-cocoercive. Thus, we can apply [14] Theorem 5.2 to obtain the desired result. $\square$

**References**

[1] Abbas, B., Attouch, H.: Dynamical systems and forward-backward algorithms associated with the sum of a convex subdifferential and a monotone cocoercive operator. Optimization 64(10), 2223–2252 (2015)
[2] Attouch, H., Briceno Arias, L., Combettes, P.L.: A parallel splitting method for coupled monotone inclusions. SIAM J. Control Optim. 48(5), 3246–3270 (2009/10)
[3] Baillon, J.B., Haddad, G.: Quelques propriétés des opérateurs angle-bornés et $n$-cycliquement monotones. Israel Journal of Mathematics 26(2), 137–150 (1977)
[4] Bauschke, H.H., Combettes, P.L.: The Baillon-Haddad theorem revisited. J. Convex Anal. 17(3-4), 781–787 (2010)
[5] Bauschke, H.H., Combettes, P.L.: Convex analysis and monotone operator theory in Hilbert spaces, second edn. CMS Books Math./Ouvrages Math. SMC. Springer, Cham (2017)
[6] Boţ, R.I., Csetne, E.R.: A dynamical system associated with the fixed points set of a nonexpansive operator. J. Dynam. Differential Equations 29(1), 155–168 (2017)
[7] Borwein, J., Vanderwerff, J.: Convex functions: constructions, characterizations and counterexamples, Encyclopedia Math. Appl., vol. 109. Cambridge University Press, Cambridge (2010)
[8] Byrne, C.: On a generalized Baillon-Haddad theorem for convex functions on Hilbert space. J. Convex Anal. 22(4), 963–967 (2015)
[9] Combettes, P. L.: Solving monotone inclusions via compositions of nonexpansive averaged operators. Optimization 53(5-6), 475–504 (2004)

[10] Cominetti, R., Peypouquet, J., Sorin, S.: Strong asymptotic convergence of evolution equations governed by maximal monotone operators with Tikhonov regularization. J. Differential Equations 245(12), 3753–3763 (2008)

[11] Contreras, A., Peypouquet, J.: Asymptotic equivalence of evolution equations governed by cocoercive operators and their forward discretizations. J. Optim. Theory Appl. (2018). DOI 10.1007/s10957-018-1332-3. URL https://doi.org/10.1007/s10957-018-1332-3

[12] Hiriart-Urruty, J. B., Strodiot, J. J., Nguyen, V. H.: Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data. Appl. Math. Optim. 11(1), 43–56 (1984)

[13] Mordukhovich, B.: Variational analysis and generalized differentiation. I, Grundlehren Math. Wiss., vol. 330. Springer-Verlag, Berlin (2006). Corrected, 2nd printing 2013

[14] Pérez-Aros, P., Vilches, E.: Tikhonov regularization of dynamical systems associated with nonexpansive operators defined in closed and convex sets. Submitted. (2019)

[15] Rockafellar, R. T., Wets, R.: Variational analysis, Grundlehren Math. Wiss., vol. 317. Springer-Verlag, Berlin (1998). Corrected 3rd printing 2009.

[16] Tseng, P.: Applications of a splitting algorithm to decomposition in convex programming and variational inequalities. SIAM J. Control Optim. 29(1), 119–138 (1991)

[17] Zhu, D., Marcotte, P.: Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities. SIAM J. Optim. 6(3), 714–726 (1996)

(Pedro Pérez-Aros) Universidad de O’Higgins, Instituto de Ciencias de la Ingeniería, Av. Libertador Bernardo O’Higgins 611, Rancagua, 2841959, Chile
E-mail address: pedro.perez@uoh.cl

(Emilio Vilches) Universidad de O’Higgins, Instituto de Ciencias de la Educación and Instituto de Ciencias de la Ingeniería, Av. Libertador Bernardo O’Higgins 611, Rancagua, 2841959, Chile
E-mail address: emilio.vilches@uoh.cl